Computational Parameters in Simulation of Microscope Images

Egon Marx and James Potzick

National Institute of Standards and Technology, Gaithersburg, MD 20899-8212, USA

Abstract— The simulation of microscope images computed from scattered fields determined using integral equations depends on a number of parameters that are not related to the scatterer or to the microscope but are choices made for the computation method. The effect of different choices on the images and on the estimated line width are determined for two typical configurations.

1. INTRODUCTION

Lines on wafers or photomasks used by the semiconductor industry can be represented by two-dimensional dielectric or conducting strips on a substrate, which is a simple scattering configuration. Image simulation is required in the determination of the width of a feature from its image, especially when this feature is smaller than the wavelength of the light used in the microscope. Simulated optical images of such a strip and other similar cylindrical configurations are obtained from fields computed above the substrate and the strips. This has been done both in the reflection and the transmission modes. The latter involves a light source under a transparent layer of finite thickness.

The scattered fields are determined from unknown boundary functions that satisfy singular integral equations [1]. They can be solved by a variety of methods and we have selected point matching to do so. In this method, the segments representing interfaces shown in Fig. 1 are divided into patches and the equations are satisfied at the center points. Fig. 1 also shows the coordinate system. The effects of computational parameters such as the number of patches on the interfaces and the number of plane waves representing the light illuminating the target are expressed in terms of a quality of fit. In general, periodic functions require at least six or ten points per wavelength of the light when used in numerical integrations. This is not required far away from the scatterer where the boundary functions are negligibly small. These unknown functions diverge where the regions have sharp corners, as has been extensively studied for dielectric and conducting wedges [2]. Corners have to be avoided but points may be chosen to accumulate near them. If one knew how the unknown functions behave near edges one could incorporate this information in the integrals. An approach using hypersingular integral equations [2] leads to constant functions near the edges, but integrals are difficult to evaluate due to the rapid variation of the kernel near the singularity.

In this study, a reference image is computed choosing some parameters such as numbers of points as large as permitted by the available computing resources (*e.g.*, 40 Gb) and selecting other optimum values. This image is compared to images using smaller parameters by determining a quality of fit based on the distances between the curves, either for the same abscissa or perpendicular to the curves. If the latter method is chosen, the plotting parameters should preferably be dimensionless. Alternatively, one can express the discrepancy in the images in terms of the difference in the apparent line width or space width [3]. Visual comparison of graphs can indicate rough differences but is inadequate for high precision measurements.

When simulated images are compared to measured ones, the differences due to imperfections of the instrument and simplifications of the feature characteristics are much larger than those between simulations [4]. Effects differ for different materials, feature size, wavelength, focus height, numerical apertures, and so on. A change in computational parameters may lead to savings in computer run time, required memory, or both.

2. BASIC EQUATIONS

The eight unknown jumps in the normal derivatives of E_z and H_z, η_{ij} and η'_{ij}, respectively, shown in Fig. 1 are defined on the boundaries C_j. One of the eight integral equations is [1]

$$
\begin{aligned}
&\left(\frac{1}{2}G_{11}^2 + b_2 G_{32}^2 N'^2_{11} + N_{32}^2 G_{11}^2\right)\{\eta_{11}\} + \left(\frac{1}{2}G_{12}^2 + \frac{1}{2}b_2 G_{32}^2 + b_2 G_{32}^2 N'^2_{12} + N_{32}^2 G_{12}^2\right)\{\eta_{12}\} \\
&- \left(\bar{b}_3 G_{33}^2 N'^3_{21} + N_{33}^2 G_{21}^3\right)\{\eta_{21}\} + \left(\frac{1}{2}\bar{b}_3 G_{33}^2 - \bar{b}_3 G_{33}^2 N'^3_{23} - N_{33}^2 G_{23}^3\right)\{\eta_{23}\} \\
&+ \beta_2 G_{32}^2 N''^2_{11}\{\eta'_{11}\} + \beta_2 G_{32}^2 N''^2_{12}\{\eta'_{12}\} - \bar{\beta}_3 G_{33}^2 N''^3_{21}\{\eta'_{21}\} - \bar{\beta}_3 G_{33}^2 N''^3_{23}\{\eta'_{23}\} \\
&= -\frac{1}{2}E_z^{h1} - N_{32}^2\left\{E_z^{h1}\right\} - b_2 G_{32}^2\left\{\partial E_z^{h1}/\partial n\right\} + N_{33}^2\left\{E_z^{h2}\right\} \\
&\quad + \bar{b}_3 G_{33}^2\left\{\partial E_z^{h2}/\partial n\right\} - \beta_2 G_{32}^2\left\{\partial H_z^{h1}/\partial s\right\} + \bar{\beta}_3 G_{33}^2\left\{\partial H_z^{h2}/\partial s\right\},
\end{aligned}
\tag{1}
$$

in a reflection problem, where the arguments on the right-hand side of the equation are mainly normal or tangential derivatives of the homogeneous fields on the boundaries; the functionals $G\{\eta\}$ and $N\{\phi\}$, evaluated at the field point $\vec{\xi} = x\hat{e}_x + y\hat{e}_y$, are defined in terms of Hankel functions by

$$
G\{\eta\}\left(\vec{\xi}\right) = -\frac{i}{4}\int_C ds'\,\eta\left(s'\right) H_0^{(1)}\left(k_\perp R\right), \quad N\{\phi\}\left(\vec{\xi}\right) = \frac{i}{4}\int_C ds'\,\phi\left(s'\right) H_1^{(1)}\left(k_\perp R\right) k_\perp \hat{n}' \cdot \hat{R}, \tag{2}
$$

where $k_\perp$ is the magnitude of the component of the wave vector in the xy-plane, $\vec{\xi}'$ is the source point at s', $\vec{R} = \vec{\xi} - \vec{\xi}'$, $R = |\vec{R}|$, $\hat{R} = \vec{R}/R$, and $\hat{n}' = \hat{n}(s')$. We use η for a jump in the normal derivative of a field and ϕ for a jump in a field. The functionals N' and N'' are the normal and tangential derivatives of G. The system of equations is reduced to a set of linear algebraic equations by the point-matching method or by the method of moments. The fields are obtained from

$$
E_x^{\mathrm{sc}}\left(\vec{\xi}\right) = -\frac{1}{4k_\perp}\sum_{j=1}^2\left[\omega\mu_1\int_{C_j} ds'\,\eta'_{1j}(s') H_1^{(1)}\left(k_\perp R\right)\frac{y-y'}{R} + k_z\int_{C_j} ds'\,\eta_{1j}(s') H_1^{(1)}(k_\perp R)\frac{x-x'}{R}\right], \tag{3}
$$

for instance, where $\vec{\xi}$ is in V_1 above the line, once the boundary functions are computed.
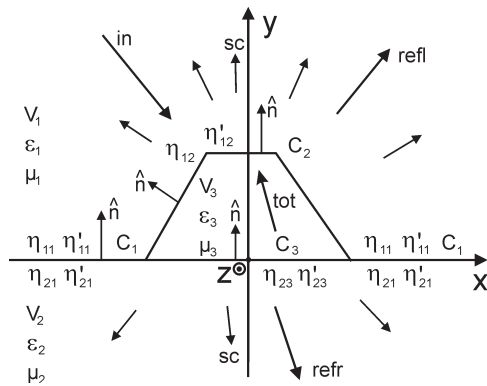

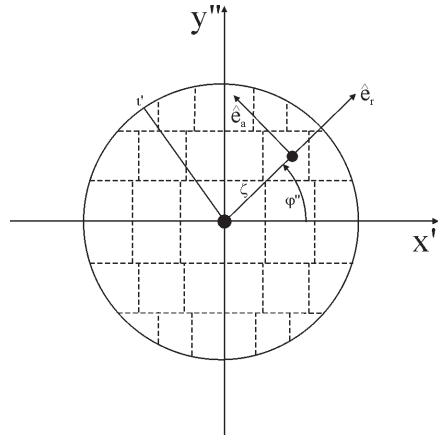Figure 1: Scattering by a strip on a substrate.

Figure 2: Illumination pattern.

3. THE MICROSCOPE

We approximate the illumination in the microscope by a set of plane monochromatic waves with different directions of incidence, as in Köhler illumination, here chosen to be symmetric about the normal to the substrate. The illumination numerical aperture determines a circle in the back focal plane of the lens system. This circle is subdivided into a number of strips of equal height and each strip is subdivided into patches of more or less equal size, as shown in Fig. 2. The intensity of each

plane wave is proportional to the size of the patch, and the intensities of the resulting scattered plus reflected (for the reflection mode) or transmitted (for the transmission mode) fields are added. The image is also determined by the collection numerical aperture and the magnification of the lens.

4. COMPUTATIONAL PARAMETERS

In addition to the parameters defining the scatterer and the microscope, there are other parameters that are related solely to the numerical computations. They are the number of strips into which the circle in the back focal plane is divided, the number of patches on the largest strip, the number of points on each segment of the interfaces in the cross section of the scatterer where the fields are matched, the distance of closest approach to the corner at the end of each segment, the height where the fields are calculated, and the number and distribution of points on this segment for the Fourier transforms used to compute the image. If the field points are not equispaced and do not number 2^n for integer n, we cannot use the fast Fourier transform used here but we can use other numerical Fourier transforms and save on points away from the scatterer.

We selectively reduce or otherwise change some of the parameters and obtain a degraded image. We compute the quality of fit, that is, the difference between the two images, by a the square root of the sum of the squares of the differences or the absolute value of the differences. The former method gives more emphasis to the part of the image with a large slope. Still another method would add the sum of the perpendicular distances between the images, which is more complicated to implement and requires the same choice of dimension for abscissas and ordinates. The images are normalized to the values far away from the scatterer, corresponding to the reflection or transmission of the incident wave in the absence of the scatterer, which makes the intensity values dimensionless. The abscissas would have to be made dimensionless by dividing them, for instance, by the wavelength or by the line width. Since the line width and similar quantities are determined from the image near the scatterer, we limit the evaluation of the quality of fit to a specified region about the scatterer. We divide by the number of points, M, of the first image and define two of the possible quality of fit metrics,

$$Q_{12}^{\mathrm{s}} = \left\{ \frac{1}{M} \sum_{i=1}^{M} \left[I_1\left(x_i\right) - \bar{I}_2\left(x_i\right) \right]^2 \right\}^{1/2}, \quad Q_{12}^{\mathrm{a}} = \frac{1}{M} \sum_{i=1}^{M} \left| I_1(x_i) - \bar{I}_2\left(x_i\right) \right|, \tag{4}$$

where I_1 and $\bar{I}_2$ are the intensities at point x. These definitions tend to emphasize the contributions where the image has a large slope, especially with Q_{12}^{s}. This may be desirable if the purpose of the computations is the determination of the line width represented by the distance between the part of the image with a large slope. It is not necessary for the second image to be defined at the same points as the first one. If it is not, it is necessary to interpolate the values of the image intensity of the second image to match the abscissas for the reference image I_1, indicated by the bar over I_2.

We quantify the errors that are caused by changing some of the computational parameters. These values depend on the scatterers under consideration and the properties of the media. For some configurations we determine the errors in the apparent line width obtained by decreasing the values of the computational parameters to relate them to the quantity of interest in the simulation.

5. NUMERICAL EXPERIMENTS

We have chosen two configurations that are of interest to the semiconductor industry. One set is composed of lines and spaces for a chrome layer on a quartz substrate $6350\,\mu\mathrm{m}$ thick imaged in the transmission mode illuminated with light of a wavelength equal to $0.365\,\mu\mathrm{m}$, and the other of lines and trenches on a Si substrate imaged in the reflection mode at $0.546\,\mu\mathrm{m}$. We have varied the computational parameters one at a time deviating from the reference simulation. The images have been all normalized to 1 far from the scatterer, where only the reflected or transmitted field contributes. We also limit the evaluation of the quality of fit to a region in the vicinity of the scatterer to avoid the contribution from numerical errors in the long tails of the images.

The variation in the number of strips in the illumination circle in the back focal plane was chosen to be 30, 20, 10, and 6 strips and the same maximum number of patches in the widest strip. The quality of fit for chrome on quartz lines and spaces of width $0.125\,\mu\mathrm{m}$, $0.25\,\mu\mathrm{m}$, and $1\,\mu\mathrm{m}$ are shown in Fig. 3. For the $1\,\mu\mathrm{m}$ line and space we also show the quality of fit for a point distribution in geometric progression in which the distance of closest approach is $0.0001\,\mu\mathrm{m}$ at the

ends of each segment and increasing towards the middle. In the other simulations the points are equispaced except for those in the regions outside the scatterer. The differences in the images shown in Fig. 4 are mostly away from the scatterer. This seems to be the case for such images. As this would suggest, the apparent line widths computed from the images [3] do not differ much from the nominal input value. We have not found a correlation between the error in line width and the quality of fit. We have obtained similar results for the Si line on a Si substrate, as shown in Fig. 5, although the quality of fit increases monotonically with a decreasing number of strips, unlike some of the lines in Fig. 3. We have varied the height at which the fields are first computed, which is usually chosen to 5 μm. We have found only small changes in the images for heights of 1 μm, 3 μm, and 10 μm. Getting too close to the line may be a problem because the influence of the induction fields increases, while getting too far away increases the relative contribution of the reflected field. Distorted images were obtained for values of 100 μm and 200 μm. The distribution of points in a geometric progression can lead to problems when this procedure leaves an insufficient number of points in the middle of a segment. It would be better to incorporate the behavior of the unknown functions near an edge in one patch and distribute the others uniformly.
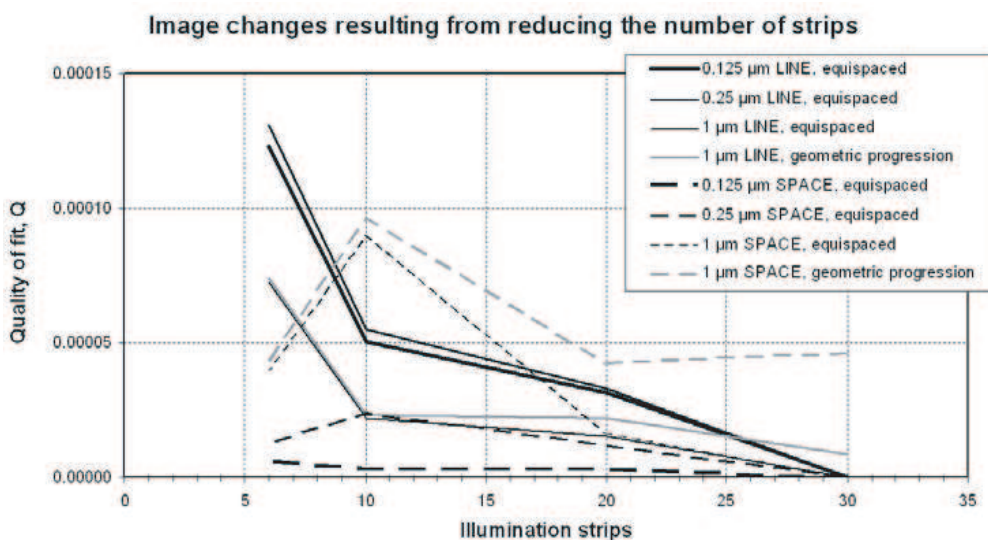

Figure 3: Quality of fit Q^s for lines and spaces (chrome on quartz).
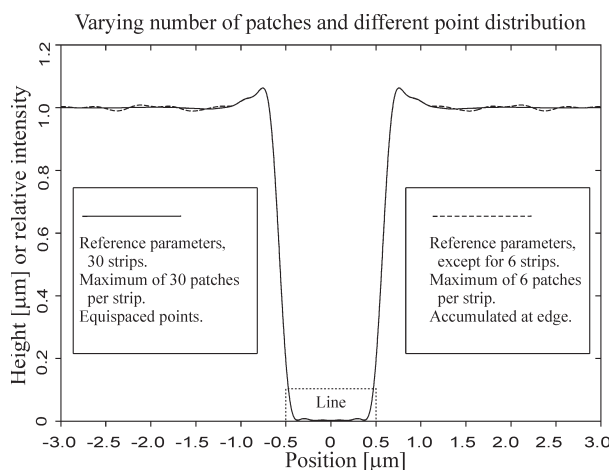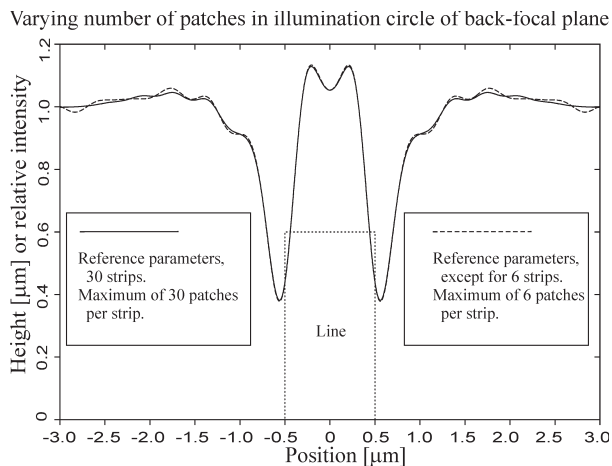

Figure 4: Cr line on quartz substrate (transmission).

Figure 5: Si line on Si substrate (reflection).

6. CONCLUSIONS

There are computational parameters that can be reduced or otherwise changed without significantly affecting the resulting image, such as the length of and the number of points in the segments

representing the interface outside the scatterer. In the transmission mode, the number of points placed on the underside of a thick quartz layer or making the layer infinite and hypothetically placing the source in the substrate does not make much difference. Reducing all the numbers of points to one half the reference values chosen initially does not significantly degrade the image.

The numerical experiments carried out to determine the best values of the computational parameters suggest that memory and time can be saved by choosing them appropriately. On the other hand, the dependence of these values on the experimental setup and the samples being simulated indicate that reaching general conclusions might be a very lengthy process. We do not obtain best parameter values that are valid generally, but we might carry out these numerical experiments until we can make an educated guess of the values chosen for the computational parameters.

We believe that the most profitable areas for further research and experimentation may be the behavior of the unknown functions near sharp edge, the influence of the segments of the interface away from the scatterer, and the type of Fourier transform used in the calculations.

REFERENCES

1. Marx, E., "Images of strips on and trenches in substrates," *Appl. Opt.*, Vol. 46, 5571–5587, 2007.
2. Marx, E., "Scattering of an arbitrary plane wave by a dielectric wedge: Integral equations and fields near the edge," *Radio Sci.*, Vol. 42, RS6S09, doi:10.1029/2006RS003568, 2007.
3. Marx, E. and J. Potzick, "Simulation of optical microscope images for photomask feature size measurements," *2005 Digest of the IEEE Antennas and Propagation Society International Symposium*, 2116–2119, 2005.
4. Silver, R., R. Attota, M. Stocker, J. Jun, E. Marx, R. Larrabee, B. Russo, and M. Davidson, "Comparison of measured optical image profiles of silicon lines with two different theoretical models," *Metrology, Inspection, and Process Control for Microlithography XVI, Proceedings of the SPIE*, Vol. 4689, 409–429, 2002.

E-polarized Diffraction Coefficients of a Composite Wedge Composed of a Perfect Conductor and a Lossy Dielectric

S. Y. Kim

Korea Institute of Science and Technology, P. O. Box 131, Cheongryang, Seoul, Korea

Abstract— The method of hidden rays is applied to the electromagnetic diffraction by composite wedges composed of perfect conductor and lossy dielectric. The E-polarized diffraction coefficients are expressed as a finite series of the cotangent functions, of which the angular period is adjusted to satisfy the edge condition at the wedge tip. The accuracy of the diffraction coefficients is assured by showing how closely the null-field condition is satisfied in the complementary region.

1. INTRODUCTION

Ordinary ray-tracing provides the geometrical optics (GO) field. Keller suggested a generalization of the Fermat principle on diffracted rays produced by an incident ray which hits edges, corners, or vertices of boundary surfaces. But the geometrical theory of diffraction (GTD) could not be implemented by employing only the ray-tracing data because the initial value of diffracted rays should be determined from the exact diffraction coefficients of some canonical structures, e.g., perfectly conducting half-plane, wedge, and cone. Hence the applicability of the GTD scheme was restricted due to the lack of rigorous diffraction coefficients of such canonical structures as penetrable wedges and cones. Recently an approximate but analytical solution to the diffraction by a composite wedge consisting of perfect conductor and lossless dielectric was constructed using the hidden rays of diffraction (HRD) method [1]. The hidden rays obey the usual principle of geometrical optics (GO) but do not exist in the physical region. These rays can be traced only in the complementary region, in which the original media inside and outside of the wedge are exchanged.

In this paper, the HRD method is applied to the E-polarized diffraction by composite wedges composed of perfect conductor and lossy dielectric. The modified propagation constants for non-uniform plane wave transmission through conducting media [2] are employed to account for the multiple reflections inside the lossy dielectric. Then one may easily obtain both trajectories of ordinary rays in the physical region and hidden rays in the complementary region. The one-to-one correspondence between geometrical rays and diffraction coefficients provides the HRD diffraction coefficients consisting of the cotangent functions. The angular period of the diffraction coefficients is adjusted to satisfy the edge condition at the tip of the composite wedge. Both PO and HRD diffraction coefficients are plotted here and compared each other.

2. THEORY

Figure 1 shows a wedge composed of a perfect conductor in S_c and a lossy dielectric with complex relative dielectric constant $\varepsilon_r^* = \varepsilon_r + i\varepsilon_i = \varepsilon_r + i\sigma/\omega\varepsilon_0$ in S_d. Consider an E-polarized unit plane wave $u_i(\rho, \theta)$ in air region S_0 incident on the wedge with an arbitrary angle θ_i.

Then the z-component of the total electric field $u(\rho, \theta)$ may be written into the dual integral equations [1] as

$$u_i(\rho,\,\theta) - F^{-1}\left[\frac{J_c(\alpha,\,\beta) + J_d(\alpha,\,\beta) + M_d(\alpha,\,\beta)}{\alpha^2 + \beta^2 - k_0^2}\right] = \begin{cases} u(\rho,\,\theta), & \text{in} \quad S_0 \\ 0, & \text{in} \quad S_c^{(0)} + S_d^{(0)} \end{cases} \quad (1)$$

$$F^{-1}\left[\frac{J_d(\alpha,\,\beta) + M_d(\alpha,\,\beta) + J_0(\alpha,\,\beta)}{\alpha^2 + \beta^2 - k_d^2}\right] = \begin{cases} u(\rho,\,\theta), & \text{in} \quad S_d \\ 0, & \text{in} \quad S_c^{(d)} + S_0^{(d)} \end{cases} \quad (2)$$

where F^{-1} denotes the operator of two-dimensional inverse Fourier transform. In (1), $S_c^{(0)}$ and $S_d^{(0)}$ are the air-filled regions of S_c and S_d, respectively. In (2), $S_c^{(d)}$ and $S_0^{(d)}$ denote the dielectric-filled
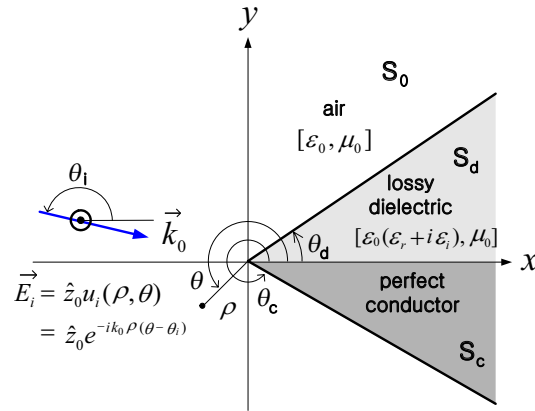
Figure 1: Geometry of composite wedge consisting of perfect conductor and lossy dielectric illuminated by E-polarized plane wave.

regions of S_c and S_0, respectively. The spectral functions J_c, J_d, M_d, and J_0 are given by [1]

$$J_c(\alpha, \beta) = -\int_0^\infty d\rho \left[\frac{1}{\rho} \frac{\partial}{\partial \theta} u(\rho, \theta_c) \right] \exp[-i\rho(\alpha \cos \theta_c + \beta \sin \theta_c)] \tag{3a}$$

$$J_d(\alpha, \beta) = \int_0^\infty d\rho \left[\frac{1}{\rho} \frac{\partial}{\partial \theta} u(\rho, \theta_d) \right] \exp[-i\rho(\alpha \cos \theta_d + \beta \sin \theta_d)] \tag{3b}$$

$$M_d(\alpha, \beta) = -i(\alpha \sin \theta_d - \beta \cos \theta_d) \int_0^\infty d\rho u(\rho, \theta_d) \exp[-i\rho(\alpha \cos \theta_d + \beta \sin \theta_d)] \tag{3c}$$

$$J_0(\alpha, \beta) = -\int_0^\infty d\rho \left[\frac{1}{\rho} \frac{\partial}{\partial \theta} u(\rho, 0) \right] \exp[-i\rho \alpha]. \tag{3d}$$

But there is no systematic way to solve the dual integral equations exactly. Hence the conventional PO approximation is performed at first. The ordinary rays in the physical region are traced using the usual principle of the geometrical optics (GO), as shown in Figure 2. Multiple reflections inside the lossy dielectric can be accounted accurately by determining the modified propagation constants for non-uniform plane wave transmission through the conducting media [2].
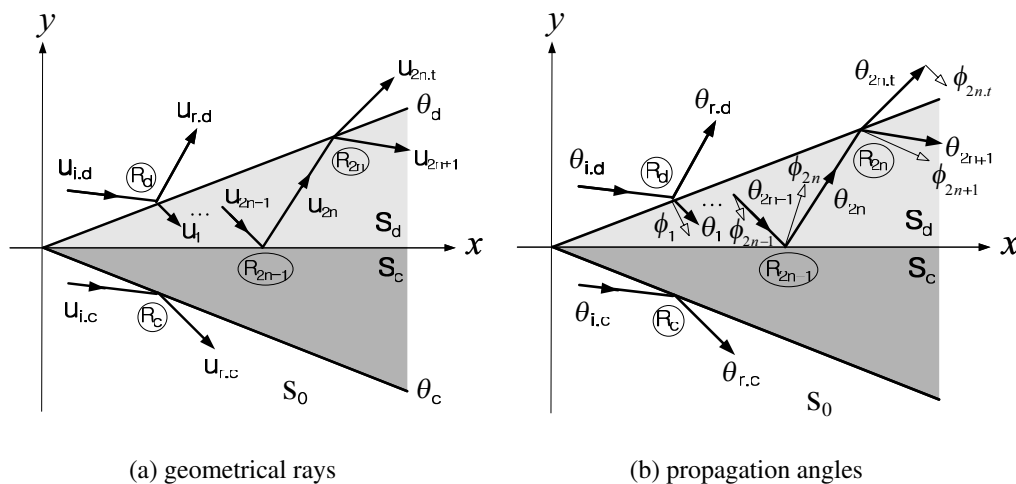
(a) geometrical rays   (b) propagation angles

Figure 2: Ordinary ray-tracing in the physical regions.

After the ordinary ray-tracing is terminated in the physical region, the first hidden ray is generated by extending the last ordinary ray into its complementary region, as shown in Figure 3(a). Figure 3(b) shows the multiple reflections in the complementary dielectric region.

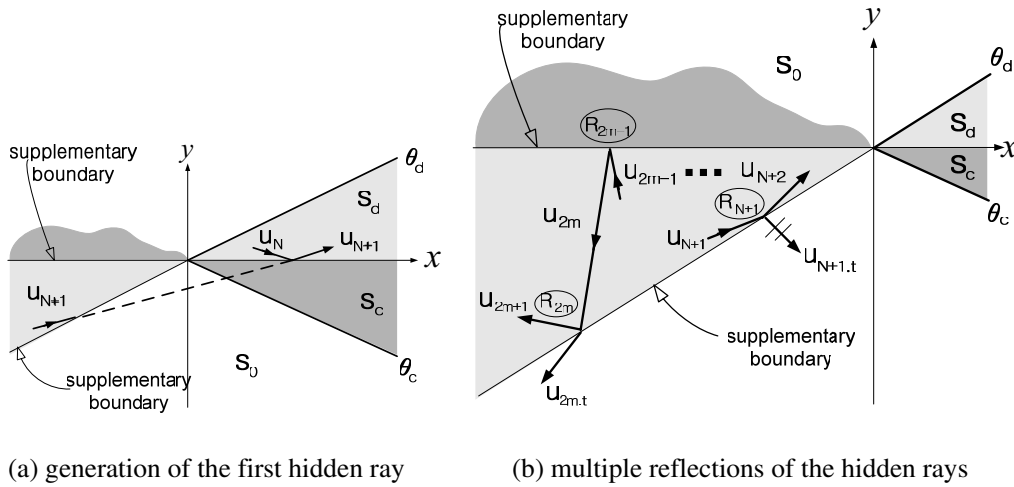(a) generation of the first hidden ray   (b) multiple reflections of the hidden rays

Figure 3: Hidden ray-tracing in the complementary regions.

The amplitudes of hidden rays are obtained routinely by the multiplication of the Fresnel's reflection coefficients. Then the HRD diffraction coefficients may be expressed by the finite series of the cotangent functions, which correspond to not only the ordinary rays in the physical region but also the hidden rays in the complementary region. It should be noted that the angular period of the cotangent functions is adjusted to satisfy the edge condition at the tip of the composite wedge.

3. RESULTS

According to the above procedure, one may easily construct the HRD diffraction coefficients in the physical and complementary air regions, $f_1^{(e)}(w)$ as

$$f_1^{(e)}(w) = \begin{cases} f_{1c}^{(e)}(w) + f_{1d}^{(e)}(w + 2\pi), & \text{in} \quad 0 \le w < \theta_d \\ f_{1c}^{(e)}(w) + f_{1d}^{(e)}(w), & \text{in} \quad \theta_d \le w \le \theta_c \\ f_{1c}^{(e)}(w - 2\pi) + f_{1d}^{(e)}(w), & \text{in} \quad \theta_c < w \le 2\pi \end{cases} \tag{4}$$

$$f_{1c}^{(e)}(w) = -\frac{1}{2\nu_r}\left[\cot\left(\frac{w - \theta_{i,c}}{2\nu_r}\right) + R_c \cot\left(\frac{w - \theta_{i,c}}{2\nu_r}\right)\right] \tag{5a}$$

$$f_{1d}^{(e)}(w) = -\frac{1}{2\nu_r}\left[\cot\left(\frac{w - \theta_{i,d}}{2\nu_r}\right) + R_d \cot\left(\frac{w - \theta_{r,d}}{2\nu_r}\right)\right] + \sum_{m=1}^{M} K_{2m,t} \cot\left(\frac{w - \theta_{2m,t}^*}{2\nu_r}\right) \tag{5b}$$

where M denotes the total number of internal reflections at the conducting boundary. And the constant ν_r is the minimum positive real value satisfying the edge condition at the tip of the composite wedge tip as

$$(\varepsilon_r + i\varepsilon_i)\tan\left(\frac{\theta_d - \theta_c}{\nu_r + i\nu_i}\right) = \tan\left(\frac{\theta_d}{\nu_r + i\nu_i}\right). \tag{6}$$

The typical example in Figure 1 is carried out for $\theta_d = 60°$, $\theta_c = 330°$, $\theta_i = 170°$, $\varepsilon_r = 1.01$, and $\varepsilon_i = 0.1 \sim 100$. The real parts of the PO and HRD diffraction coefficients are plotted in Figures 4(a) and (b), respectively. In Figure 4, the dotted ($\varepsilon_i = 0$) and broken ($\varepsilon_i = \infty$) lines denote the exact solutions to the perfectly conducting wedges with $\theta_d = 0°$ and $60°$, respectively. $S_d^{(0)}$ and $S_c^{(0)}$ denote the complementary air regions, in which the original media in S_d and S_c are replaced by air, respectively. All of the PO curves in Figure 4(a) intersect two exact patterns and cannot become zero on the conducting boundary. In contrast, the HRD diffraction coefficients in Figure 4(b)

approach the corresponding exact diffraction coefficients monotonically as ε_i decreases to 0.1 or increases to 100. According to the formulation of dual integral equations, the exact diffraction coefficients should become zero in the complementary regions. Unlike the PO diffraction coefficients in Figure 4(a), the HRD diffraction coefficients in Figure 4(b) satisfy the null-field condition in the $S_d^{(0)}$ and $S_c^{(0)}$ quite well.
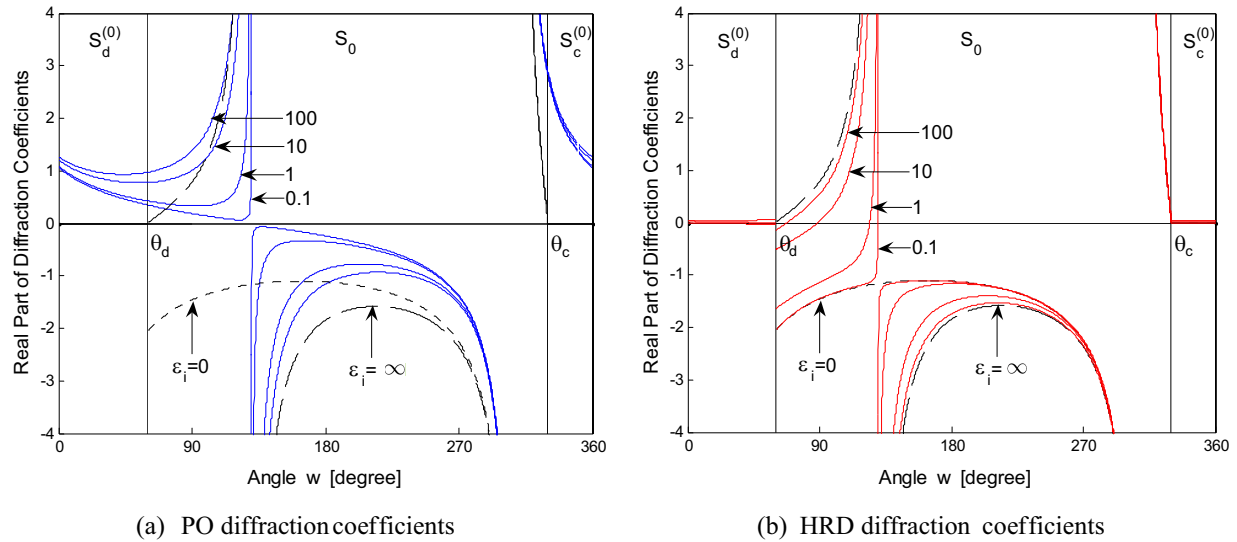

(a) PO diffraction coefficients

(b) HRD diffraction coefficients

Figure 4: Real parts of the diffraction coefficients for $\theta_d = 60°$, $\theta_c = 330°$, $\theta_i = 170°$, $\varepsilon_r = 1.01$, and $\varepsilon_i = 0.1$ to 100 (dotted: exact for $\varepsilon_r = 1$ and $\varepsilon_i = 0$, broken: exact for $\varepsilon_i = \infty$).

4. CONCLUSION

The E-polarized diffraction coefficients of composite wedges composed of perfect conductor and lossy dielectric are expressed in analytic form using the HRD method. Unlike to the conventional PO solution, the HRD diffraction coefficients provide smooth transition between the exact solutions in two limiting cases as ε_i increases from 0.1 to 100.

REFERENCES

1. Kim, S.-Y., "Hidden rays of diffraction," *IEEE Trans. Antennas Propagat.*, Vol. 55, No. 3, 892–906, Mar. 2007.
2. Radcliff, R. D. and C. A. Balanis, "Modified propagation constants for nonuniform plane wave transmission through conducting media," *IEEE Trans. Geosci. Remote Sensing*, Vol. 20, No. 3, 408–411, Jul. 1982.

A Physical De-embedding Method for Silicon-based Device Applications

Hsiao-Tsung Yen[1], Tzu-Jin Yeh[1], and Sally Liu[1, 2]

[1]Taiwan Semiconductor Manufacturing Company Ltd.
9, Creation Road, Hsinchu Science Park, Hsinchu, Taiwan, R.O.C.
[2]RFMP Department of Taiwan Semiconductor Manufacturing Company, Taiwan, R.O.C.

Abstract— In the recent years, de-embedding method of "open-short", and "open-thru" with dummy DUTs are mostly used for on-wafer devices. This paper shows a method, called L-2L, with two metal lines or transmission lines. One is two times the length of the other one. Based on the measurement data of two lines, PAD parasitics and scalable length model with certain width are obtained by L2L. Furthermore, the tape-out DUT numbers for the de-embedding are extremely minimized to 2. The applications for inductors by 65 nm Low-K Si-based process and 0.18 μm FSG Si-based process, and for 0.13 μm transmission lines are also shown for benchmark in this paper.

1. INTRODUCTION

The analysis of on-wafer device characteristics is very important for RF IC designer. In the recent years, on-wafer de-embedding methods with "open-short" and "open-thru" with dummy DUTs are the most used for handling the difficult fabrication on accurate 50 ohms load [1–3]. Both methods can subtract parasitic *resistance-inductance-capacitance* (RLC) up to GHz level. Although they are commonly used with certain confidence level, there is still a large amount of dummy test DUTs needed to gain fairly accuracy for "open-short" [4] and TRL [5, 6]. However, many de-embedding approaches for inductors are published in recent years, such as TRL [7] and THRU model with mutual inductance calculation [8].

In this paper, L-2L method is proposed with just two metal lines for dummy test DUTs, which has been published for transmission lines study [9, 10]. With physical scaling on metal line measurement, L2L can improve the high frequency calibration for small devices without additional parasitics. In this work, two metal lines that one is two times longer than the other are required. Parasitic RLC of probing PAD and metal lines with physical scalable equations will be calculated from L2L, separately. Finally, on wafer devices of 65 nm, 0.18 μm inductors and 0.13 μm transmission lines are taken as benchmark.

2. THEORY AND MEASUREMENT RESULTS

2.1. Theory

The L2L de-embedding method theory is based on the two metal lines with same width on the same wafer, which represents the same process condition. Fig. 1 shows the top-view of transmission line tape-out. They can be decomposed into ABCD matrix components shown in the following formula and figure.

$$[L] = [PAD] \cdot [TLine] \cdot [PAD] \tag{1}$$

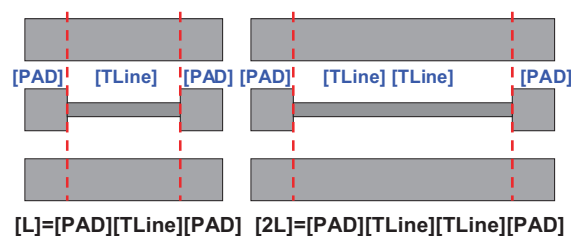$$[2L] = [PAD] \cdot [TLine] \cdot [TLine] \cdot [PAD] \tag{2}$$

Figure 1: Two 0.18 μm Dummy metal lines for example of L2L.

Based on Equations (1) and (2) shown above, we can conclude [PAD] and [Line] with the following formula.

$$[\text{PAD}] \cdot [\text{PAD}] = \left[[\text{L}]^{-1} [2\text{L}] \cdot [\text{L}]^{-1} \right]^{-1} \tag{3}$$

$$[\text{TLine}] = [\text{PAD}]^{-1} [\text{L}] \cdot [\text{PAD}]^{-1} \tag{4}$$

In the Equations (3) and (4), [PAD] can be easily calculated, and its characteristics can show the real parasitic capacitance of GSG PAD, according the same process variation on the same die. Otherwise, ABCD matrix scaling method is already published [4], which can be applied in scaling matrix of [TLine].

2.2. Measurement Results

Agilent 8510 and Agilent PNA are used for measuring from 0.2 GHz to 30 GHz and 0.2 GHz to 50 GHz, respectively. Cadence Microtech microwave GSG probes are used with LRM calibration standard procedure and also with impedance standard substrate (ISS). The probes and cables parasitics are subtracted right after the calibration procedure is finished.
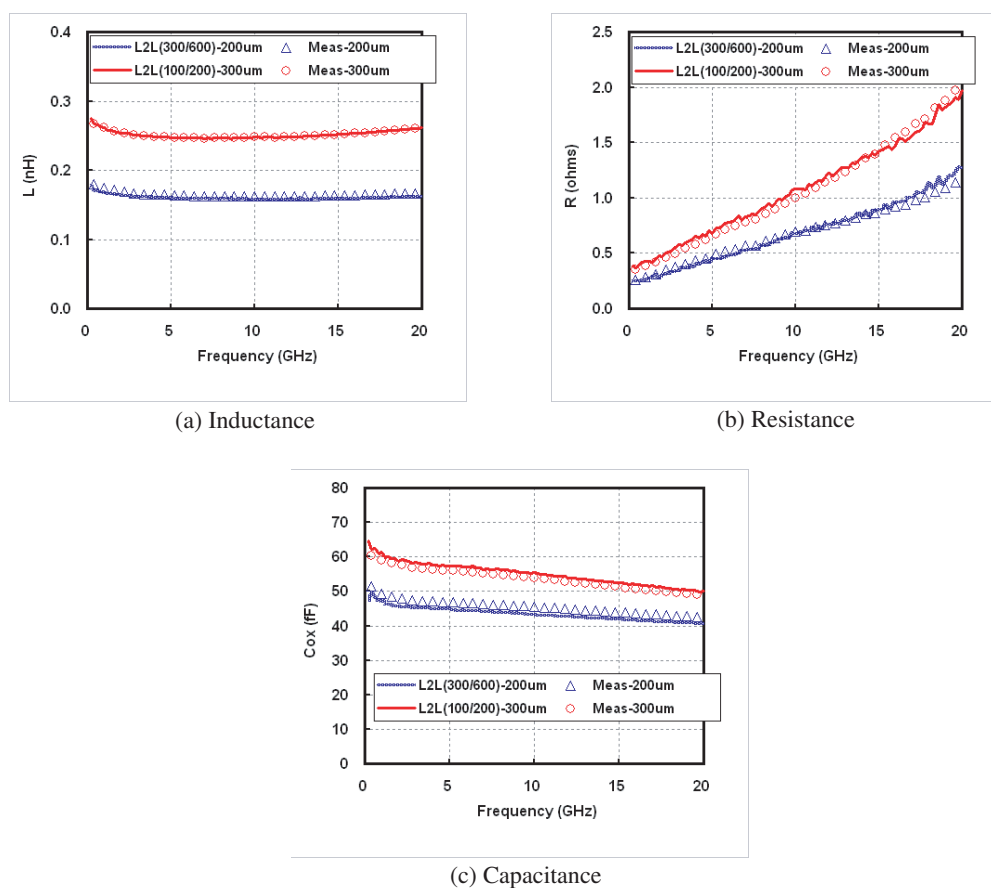

(a) Inductance

(b) Resistance

(c) Capacitance

Figure 2: 0.18 μm 1P6M metal lines DUTs with UTM, compared with 200 μm and L2L calculated from 300 μm/600 μm measurement data, whereas 300 μm and L2L calculated from 200 μm/100 μm measurement data. (a), (b), and (c) is for inductance, resistance and capacitance plots, respectively. They are all measurement data.

The measurement results are with Four different lengths of metal lines, and they are 100 μm, 200 μm, 300 μm, and 600 μm long. 100 μm and 200 μm is one set of L2L, as 300 μm and 600 μm are. Fig. 2 shows two sets of comparison results. Solid lines are for L2L calculation with GSG PAD from the other set of L2L and symbols are exactly the value from the measurement. They are all with GSG PAD. 300 μm and 600 μm metal lines are for the first setup of L2L scaling method to perform 200 μm transmission line with GSG PAD. The calculation of L2L is well-matched 200 μm measurement itself. 100 μm and 200 μm are the other setup of L2L to extrapolate the RLC result of 300 μm metal line with GSG PAD, which has excellent agreement with the actual

300 μm measurement data with PAD. Therefore, accurate results can be calculated by L-2L method according to these consistence results.

The de-embedded results by "open-thru" and L2L can be found in Fig. 3 for one inductor with only 60 μm feed metal lines. "open-thru" is widely used for inductor de-embedding recently. The difference between raw data and intrinsic DUT after L2L de-embedding method is around $47 \sim 50$ pH. This value is very consistence with Fig. 2(a), which shown 300 μm with 250 pH in measurement. On the other hand, solid lines are for "open-thru" de-embedding results. There are almost no difference between L2L and "open-thru" de-embedding within 20 GHz. But in much higher frequency, the THRU model, that is used in "open-thru" method, will be suffered more modeling quality issue and lose some accuracy on substrate resistance.
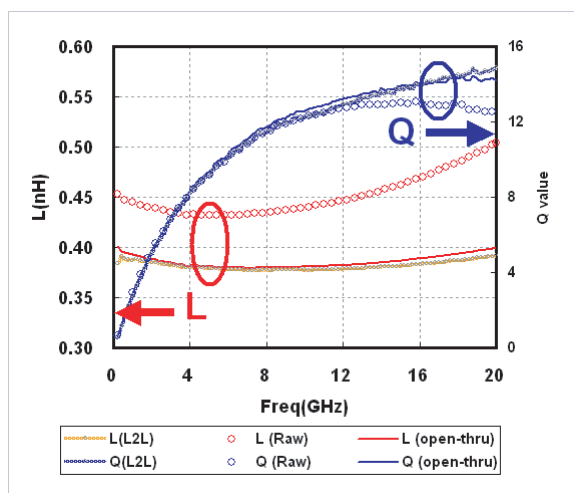

Figure 3: 0.18 μm 1P6M UTM inductor comparison of the single-ended L values between the raw data without de-embedding and the data after L2L de-embedding. The width/spacing/turns/radius is 9 μm/3 μm/1.5/30 μm.

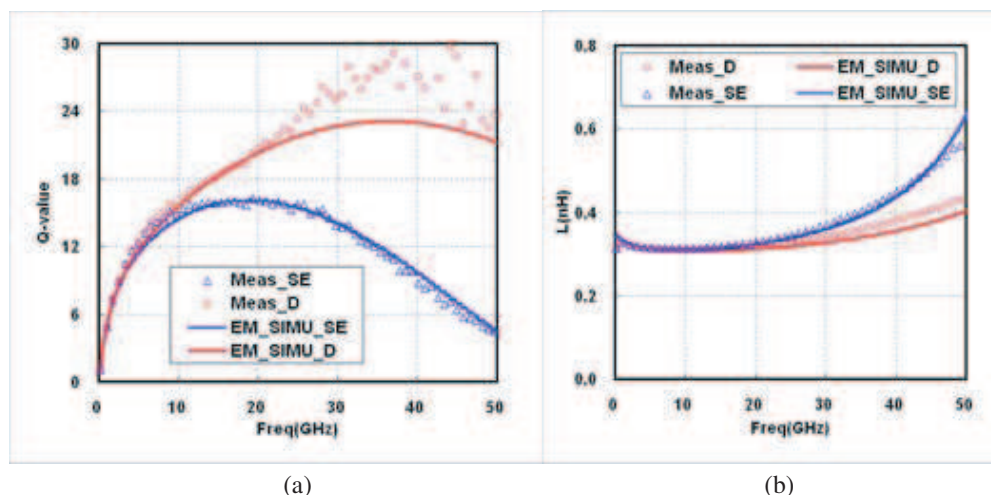
(a)   (b)

Figure 4: 65 μm 1P9M symmetric inductor with differential and single-ended L and Q plots, as D is for differential and SE is for single-ended. These data are measured up to 50 GHz.

2.3. Validation

One symmetric inductor by 65 nm back-end Low-K Si-based process is shown in Fig. 4. The dimension of this small inductor is as follows: Width /spacing /turns /radius $= 9$ μm/3 μm/2/15 μm, with 306 μm feed metal lines. [Ind] is the result of following ABCD matrixes after L2L de-embedding,

$$[\text{Ind}] = [\text{TLine}]^{-1} [\text{PAD}]^{-1} [\text{Ind(raw)}] \cdot [\text{PAD}]^{-1} [\text{TLine}]^{-1} \tag{5}$$

Another inductor case with 460 µm metal feed-lines, counted from each signal PAD, is de-embedded by L2L. In Fig. 5, the comparison is quite well between L2L de-embedded data and EM simulation result. In this paper, the process tech file of 3D EM simulator, Ansoft HFSS, was exactly the same equivalent dioxide and substrate parameters from back-end layer configuration, which dominates resonance frequency and Q value.
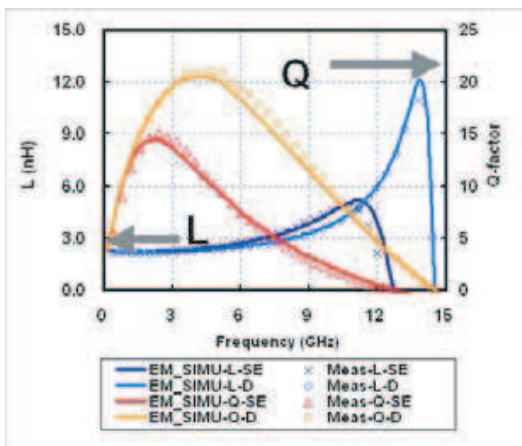

Figure 5: 65 nm 1P9M UTM symmetric inductors Q, L comparisons plot. D is for differential and SE is for single-ended. Width/Spacing/Turns/Radius = 15 µm/3 µm/3/60 µm.

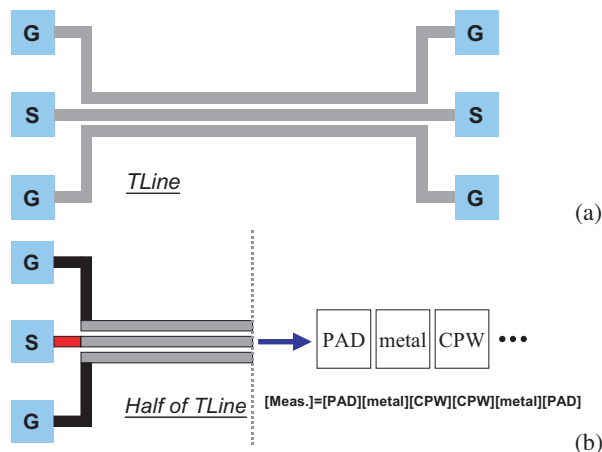
Figure 6: CPW with its top-view (a) and its analysis topology (b).

Figure 6(a) shows the layout of CPW. Three measurements of CPW DUTs are done for 500 µm, 1000 µm, and 2000 µm. If the little discontinuity in black area is neglected in Fig. 6(b), close data can also be found in Fig. 7. Solid line is the results calculated from L2L with 500 µm and 1000 µm CPW lines. It matches with actual measurement of 2000 µm long CPW in inductance and resistance, and performs good benchmark for L2L. They are all compared within GSG PAD structure.
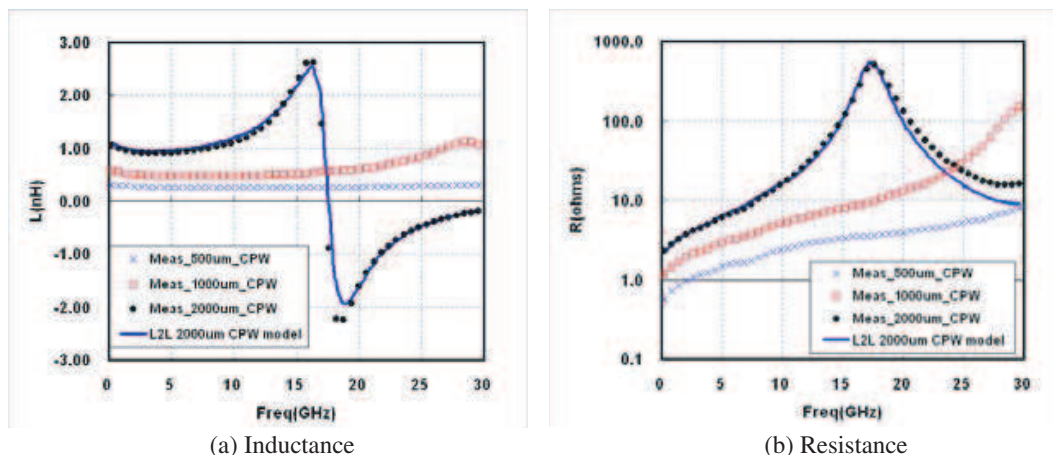

(a) Inductance

(b) Resistance

Figure 7: 0.13 µm CPW with L2L model calculated from 500 µm and 1000 µm CPW to compare with the measurement data of CPW (2000 µm).

3. CONCLUSIONS

L-2L de-embedding method is well examined in this paper with good accuracy, applying up to 50 GHz. Extremely minimized DUT number, just two dummy DUTs, is required and consistence results of inductors and transmission lines are shown. In addition, de-embedded inductors are well matched with the simulation of 3D EM tool, including small inductors.

ACKNOWLEDGMENT

REFERENCES

1. Tiemeijer, L. F. and R. J. Havens, "A calibrated lumped-element de-embedding technique for on-wafer RF characterization of high-quality inductors and high-speed transistors," *IEEE Transactions on Electron Devices*, Vol. 50, No. 3, 822–829, March 2003.

2. Jenner, M. B. and T. E. Kolding, "Test structures and techniques for on-wafer CMOS TRL calibration" *2001 Proc. ICMTS Conf.*, 137–141, 2001.

3. Padmanabhan, S., P. Kirby, J. Daniel, and L. Dunleavy, "Accurate broadband on-wafer SOLT calibrations with complex load and thru models," *ARFTG Conference Digest, Spring 2003. 61st*, 5–10, June 13, 2003.

4. Cho, M.-H., G.-W. Huang, K.-M. Chen, and A.-S. Peng, "A novel cascade-based de-embedding method for on-wafer microwave characterization and automatic measurement," *2004 IEEE MTT-S International Microwave Symposium Digest*, Vol. 2, 1237–1240, 2004.

5. Vandamme, E. P., D. M. M.-P. Schreurs, and G. Van Dinther, "Improved three-step de-embedding method to accurately account for the influence of pad parasitics in silicon on-wafer RF test-structures," *IEEE Transactions on Electron Devices*, Vol. 48, No. 4, 737–742, April 2001.

6. Kolding, T. E., O. K. Jensen, and T. Larsen, "Ground-shielded measuring technique for accurate on-wafer characterization of RF CMOS devices," *2000 Proc. ICMTS Conf.*, 246–251, 2000.

7. Guo, J.-C. and T.-Y. Tan, "A broadband and scalable lumped element model for fully symmetric inductors under single-ended and differentially driven operations," *IEEE Transactions on Electron Devices*, Vol. 54, No. 8, 1878–1888, August 2007.

8. Blaschke, V. and J. Victory, "Accurate inductance de-embedding technique for scalable inductor models," *Microelectronic Test Structures, 2007. IEEE ICMTS Conf.*, 248–252, March 2007.

9. Johansen, T. K., C. Jiang, D. Hadziabdic, and V. Krozer, "EM simulation accuracy enhancement for broadband modeling of on-wafer passive components," *Microwave Conference, 2007. European*, 1245–1248, October 9–12, 2007.

10. Goren, D., S. Shlafman, B. Sheinman, W. Woods, and J. Rascoe, "Silicon-chip single and coupled coplanar transmission line measurements and model verification up to 50 GHz," *IEEE Workshop on Signal Propagation on Interconnects, 2007. SPI 2007*, 21–24, May 13–16, 2007.

Thermal Conductivity of Nanofluid with Magnetic Nanoparticles

T.-H. Tsai[1], **L.-S. Kuo**[1], **P.-H. Chen**[1], and **C.-T. Yang**[2]

[1]Department of Mechanical Engineering, National Taiwan University, Taiwan
[2]Department of Mechanical and Computer-aided Engineering, St. John's University, Taiwan

Abstract— The thermal conductivities of Fe_3O_4 and Al_2O_3 nanofluids with the viscous base fluid composed by various fluids are investigated in this study. In order to change the viscosity of mixed fluid, the volumetric fractions between two fluids in the mixed fluid are varied. Measured values of thermal conductivity of nanofluids gradually approach the values predicted by Maxwell equation with increasing the viscosity. Our measured results demonstrate that Brownian motion of suspended magnetic particles could be an important factor that enhances the thermal conductivity of nanofluids, but it has little effect for suspended Al_2O_3 nanoparticles. It also indicates that the conduction part of prediction model can be obtained from Maxwell prediction.

1. INTRODUCTION

Due the applications of suspended nanoparticles in a base fluid, the thermal properties of nanofluids have been continually studied in the last decade. Various nanoparticles and base fluids are applied to fabricate different kind of nanofluids. Former experimental data have shown that nanofluids have higher thermal conductivity than that predicted by the conventional models like Maxwell model [1–11]. However, the conventional models did not consider the effects associated with Brownian motion. This may cause the underestimation to the thermal conductivity of nanofluids. New models included the effects associated with Brownian motion are hotly investigated in the recent years [11–23].

Some prior studies have indicated that the enhancement of thermal conductivity which directly results from collisions between nanoparticles is not obvious, and the dominant factor of the enhancement in thermal conductivity is the convection-like behavior which indirectly results from Brownian motion [12, 14, 16, 22].

Most new models divide the effective thermal conductivity, k_{eff}, into the conduction part, $k_{conduction}$, which stands for the thermal conductivity due to differences of composition of nanofluids and convection part, $k_{convection}$, which stand for the thermal conductivity due to the effects associated with Brownian motion.

$$k_{eff} = k_{conduction} + k_{convection} \tag{1}$$

However, for the conduction part, some models are based on parallel path of conduction heat transfer [22, 23] and other models are based on Maxwell prediction [13, 20]. They do not provide a consistent and convincing predict model for the conduction part of all nanofluids.

This study tries to clarify the model of conduction part on the enhancement of thermal conductivity with nanoparticles in a base fluid at different values of viscositiy. According to the Einstein-Stokes's equation, the Brownian diffusion coefficient, D_B, is the function of temperature, diameter of nanoparticle and viscosity of fluid:

$$D_B = \frac{k_B T}{3\pi \mu d_p} \tag{2}$$

where k_B is Boltzmann constant, T is temperature, μ is viscosity of nanofluids, and d_p is diameter of nanoparticles. The Brownian diffusion coefficient will decrease with increasing the viscosity of nanofluids, and the effect of Brownian motion or convectionlike behavior will be weakened. Then the thermal conductivity of convection part should decrease and the thermal conductivity of conduction part will be measured solely in the experiment.

2. EXPERIMENTAL SECTION

In this study, Al_2O_3 water-based nanofluid and Fe_3O_4 oil-based nanofluid are investigated. For the Al_2O_3 water-based nanofluid, two kinds of viscous base fluids, the compound of water and EG and the compound of EG and Glycerol, are used. The measured viscosities of water, EG and glycerol are 1cP, 16.8cP and 937cP, respectively. Figure 1(a) shows the viscosity of the water-EG base

fluid as a function of volume fraction of EG at 25°C. Figure 1(b) shows the viscosity of the EG-glycerol base fluid as a function of volume fraction of glycerol at 25°C. The viscosities of both base fluids grow exponentially when the volume fraction of EG or glycerol increases. Al_2O_3 of average diameter of 13 nm is used as nanoparticles dispersed in water-EG or EG-glycerol base fluid. Al_2O_3 nanoparticles are purchased from the Yong-Zhen technomaterial CO., LTD.

For the Fe_3O_4 oil-based nanolfuid, diesel oil and PDMS are used to form viscous base fluid. The viscosity of diesel oil and PDMS are 4.188cP and 5500cP, respectively. Figure 1(c) shows the viscosity of the base fluid as a function of volume fraction of PDMS at 25°C. The viscosity of the base fluid grows exponentially when the volume fraction of PDMS increases. Fe_3O_4 of average diameter of 10 nm used in this study is fabricated by co-precipitation method. The chemical reaction formula is expressed as:

$$Fe^{2+} + 2Fe^{3+} + 8OH^- \rightarrow Fe_3O_4 + 4H_2O$$

The thermal conductivity of such nanofluids is measured by the thermal hot wire method at 25°C.
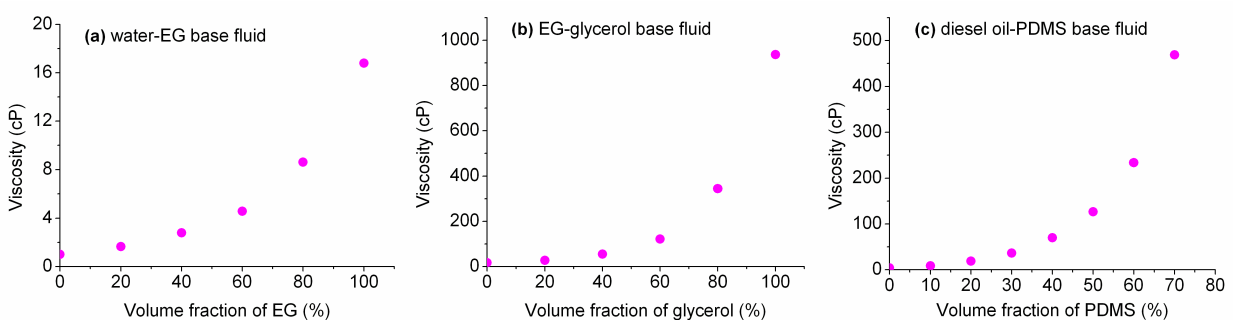

Figure 1: Measured values of viscosity of the mixed base fluid versus volume fraction of (a) EG (b) glycerol (c) PDMS at 25°C.

3. RESULTS AND DISCUSSION

Figure 2 shows the thermal conductivity ratio of Al_2O_3 nanofluid versus the viscosity of base fluid with 2% volume fraction of Al_2O_3 nanoparticles. The experimental results do not have regularities in thermal conductivity. It may result from that the commercial nanoparticles are dry powder and cannot be dispersed in the base fluid well. The aggregation causes the unstable thermal conductivity ratio of nanofluid.
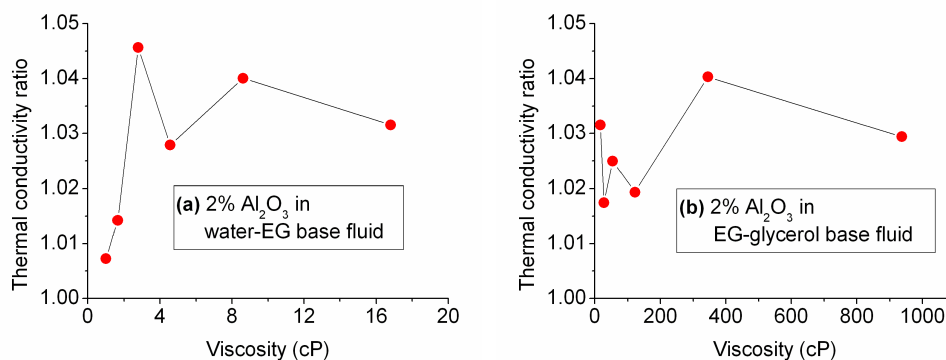

Figure 2: Thermal conductivity ratio of Al_2O_3 nanofluids versus viscosity of base fluid: (a) water-EG base fluid, (b) EG-glycerol base fluid.

Figure 3 shows the effect of viscosity of base fluid on the thermal conductivity ratio, k_{nano}/k_{bf} and $k_{Maxwell}/k_{bf}$, of nanofluids with 1% and 2% volume fraction of Fe_3O_4 nanoparticles. The parameters, k_{nano} and k_{bf}, denote the thermal conductivity of nanofluids and base fluid, respectively. The predicted thermal conductivity of nanofluid, $k_{Maxwell}$, is determined by the Maxwell equation,

given by

$$k_{eff} = \frac{k_p + 2k_{bf} + 2(k_p - k_{bf})\phi}{k_p + 2k_{bf} - (k_p - k_{bf})\phi} \tag{3}$$

where k_p is thermal conductivity of nanoparticles, k_{bf} is thermal conductivity of base fluid, and ϕ is volume fraction of nanoparticles.
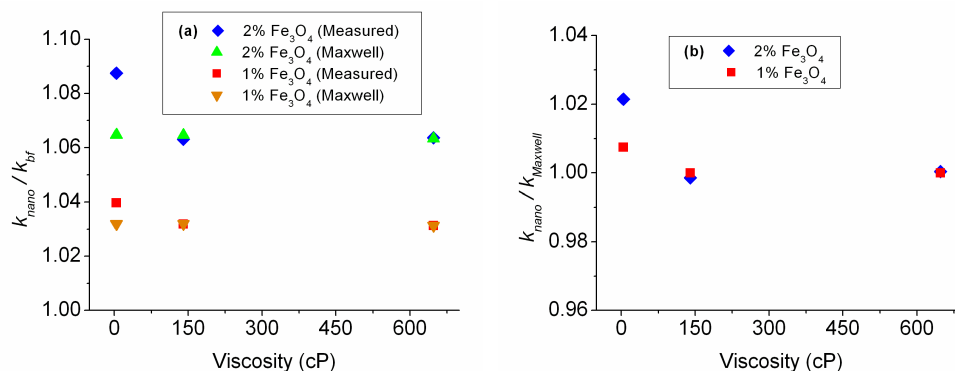

Figure 3: Effect of viscosity of base fluid on thermal conductivity ratio of nanofluids: (a) k_{nano}/k_{bf}, (b) $k_{nano}/k_{Maxwell}$.

With the low viscous base fluid (4.188cP), the thermal conductivity ratio (k_{nano}/k_{bf}) of nanofluids is higher than that predicted by the Maxwell equation. The highest enhancement of thermal conductivity is 8.75% at 2.24% volume fraction of Fe_3O_4 nanoparticles. And $k_{nano}/k_{Maxwell}$ are 1.007 and 1.021 at 1.12% and 2.24% volume fraction of Fe_3O_4 nanoparticles, respectively. The experimental results indicate that Brownian motion and the convectionlike behavior are much active and cause the observable enhancement of thermal conductivity in the low viscous base fluid.

With the high viscous base fluid (140.4cP), the thermal conductivity ratio (k_{nano}/k_{bf}) of nanofluids becomes the same as that predicted by the Maxwell equation. Measured values of $k_{nano}/k_{Maxwell}$ in the high viscous fluid are 1 at 1.12% and 2.24% volume fraction of Fe_3O_4 nanoparticles. At this situation, high viscosity of nanofluids makes Brownian motion and the convectionlike behavior disappear. The measured thermal conductivity of nanofluids only presents the conduction part which Maxwell equation predicts.

The experimental results with the extra high viscous base fluid (648cP) are almost identical to those in the high viscous base fluid (140.4cP).

4. CONCLUSIONS

For the Al_2O_3 nanofluid, the unstable thermal conductivity may result from the aggregation. For the Fe_3O_4 nanofluid, high viscosity of base fluid weakens Brownian motion of suspended nanoparticles. Our experimental results with various viscosities of base fluids indicate that Brownian motion of suspended nanoparticles is one of important factors to enhance the thermal conductivity of nanofluids. In such a highly viscous fluid, Maxwell equation gives a good prediction on the thermal conductivity of nanofluids without Brownian motion of suspended nanoparticles, and the conduction part of prediction model can be obtained from Maxwell prediction, i.e., $k_{conduction} = k_{Maxwell}$.

REFERENCES

1. Lee, S., S. U. S. Choi, and J. A. Eastman, "Measuring thermal conductivity of fluids containing oxide nanoparticles," *J. Heat Transfer*, Vol. 121, No. 2, 280–289, 1999.

2. Wang, X., X. Xu, and S. U. S. Choi, "Thermal conductivity of nanoparticle-fluid mixture," *J. Thermophys. Heat Transfer*, Vol. 13, No. 4, 474–480, 1999.

3. Xuan, Y. and Q. Li, "Heat transfer enhancement of nanofluids," *Int. J. Heat Fluid Flow*, Vol. 21, No. 1, 58–64, 2000.

4. Eastman, J. A., S. U. S. Choi, S. Li, W. Yu, and L. J. Thompson, "Anomalously increased effective thermal conductivities of ethylene glycol-based nanofluids containing copper nanoparticles," *Appl. Phys. Lett.*, Vol. 78, No. 6, 718–720, 2001.

5. Xie, H., J. Wang, T. Xi, Y. Liu, F. Ai, and Q. Wu, "Thermal conductivity enhancement of suspensions containing nanosized alumina particles," *J. Appl. Phys.*, Vol. 91, No. 7, 4568–4572, 2002.

6. Das, S. K., N. Putta, P. Thiesen, and W. Roetzel, "Temperature dependence of thermal conductivity enhancement for nanofluids," *J. Heat Transfer*, Vol. 125, No. 4, 567–574, 2003.

7. Patel, H. E., S. K. Das, T. Sundararagan, A. S. Nair, B. Geoge, and T. Pradeep, "Thermal conductivities of naked and monolayer protected metal nanoparticle based nanofluids: Manifestation of anomalous enhancement and chemical effects," *Appl. Phys. Lett.*, Vol. 83, No. 14, 2931–2933, 2003.

8. Hong, T. K., H. S. Yang, and C. J. Choi, "Study of the enhanced thermal conductivity of Fe nanofluids," *J. Appl. Phys.*, Vol. 97, No. 6, 064311, 2005.

9. Murshed, S. M. S., K. C. Leong, and C. Yang, "Enhanced thermal conductivity of TiO_2 — water based nanofluids," *Int. J. Therm. Sci.*, Vol. 44, No. 4, 367–373, 2005.

10. Hwang, Y., H. S. Park, J. K. Lee, and W. H. Jung, "Thermal conductivity and lubrication characteristics of nanofluids," *Curr. Appl. Phys.*, Vol. 6, Supplement 1, e67–e71, 2006.

11. Li, C. H. and G. P. Peterson, "Experimental investigation of temperature and volume fraction variations on the effective thermal conductivity of nanoparticle suspensions (nanofluids)," *J. Appl. Phys.*, Vol. 99, No. 8, 084314, 2006.

12. Keblinski, P., S. R. Phillpot, S. U. S. Choi, and J. A. Eastman, "Mechanisms of heat flow in suspensions of nano-sized particles (nanofluids)," *Int. J. Heat Mass Transfer*, Vol. 45, No. 4, 855–863, 2002.

13. Koo, J. and C. Kleinstreuer, "A new thermal conductivity model for nanofluids," *J. Nanopart. Res.*, Vol. 6, No. 6, 577–588, 2004.

14. Jang, S. P. and S. U. S. Choi, "Role of Brownian motion in the enhanced thermal conductivity of nanofluids," *Appl. Phys. Lett.*, Vol. 84, No. 21, 4316–4318, 2004.

15. Prasher, R., P. Bhattacharya, and P. E. Phelan, "Thermal conductivity of nanoscale colloidal solutions (nanofluids)," *Phys. Rev. Lett.*, Vol. 94, 025901, 2005.

16. Evans, W., J. Fish, and P. Keblinski, "Role of Brownian motion hydrodynamics on nanofluid thermal conductivity," *Appl. Phys. Lett.*, Vol. 88, No. 9, 093116, 2006.

17. Keblinski, P. and J. Thomin, "Hydrodynamic field around a Brownian particle," *Phys. Rev. E*, Vol. 73, 010502, 2006.

18. Prasher, R., P. E. Phelan, and P. Bhattacharya, "Effect of aggregation kinetics on the thermal conductivity of nanoscale colloidal solutions (nanofluid)," *Nano. Lett.*, Vol. 6, No. 7, 1529–1534, 2006.

19. Buongiorno, J., "Convective transport in nanofluids," *J. Heat Transfer*, Vol. 128, No. 3, 240–250, 2006.

20. Prasher, R., P. Bhattacharya, and P. E. Phelan, "Brownian-motion-based convective-conductive model for the effective thermal conductivity of nanofluids," *J. Heat Transfer*, Vol. 128, No. 6, 588–595, 2006.

21. Prakash, M. and E. P. Giannelis, "Mechanism of heat transport in nanofluids," *J. Comput. Aided Mater. Des.*, Vol. 14, No. 1, 109–117, 2007.

22. Jang, S. P. and S. U. S. Choi, "Effects of various parameters on nanofluid thermal conductivity," *J. Heat Transfer*, Vol. 129, No. 5, 617–623, 2007.

23. Patel, H. E., T. Sundararajan, and S. K. Das, "A cell model approach for thermal conductivity of nanofluids," *J. Nanopart. Res.*, Vol. 10, No. 1, 87–97, 2008.

On the Use Complex Susceptibility Measurements in Investigating the Field Dependence of Resonance and After-effect Function of Nano-particle Colloids

P. C. Fannin

Department of Electronic and Electrical Engineering, Trinity College, Dublin 2, Ireland

Abstract— Measurement of the field and frequency dependent magnetic complex susceptibility, $\chi(\omega, H) = \chi'(\omega, H) - i\chi''(\omega, H)$, of nano-particle colloids, such as magnetic fluids, has proven to be a reliable method for investigating a number of important properties of such fluids including ferromagnetic resonance, relaxation mechanisms and other important parameters. Also, because of the inverse Fourier transform relationship which exists between the after-effect function, $b(t)$, and $\chi''(\omega)$, $b(t)$ may be obtainedgnd used in determining a value of the precessional decay time, τ_0.

Here, examples of results obtained from measurements obtained by means of the transmission line technique in the MHz-GHz range, are presented.

1. INTRODUCTION

Magnetic fluids are stable colloidal systems consisting of single-domain magnetic particles covered by a surfactant in order to prevent particle agglomeration and dispersed in a carrier liquid. The particles have radii ranging from approximately 2–10 nm and when in suspension their magnetic properties can be described by the paramagnetic theory of Langevin. The particles are considered to be in a state of uniform magnetization with a magnetic moment $m = M_s v$, where M_s denotes the saturation magnetization and v is the magnetic volume of the particle.

There are three characteristic times which govern the behaviour of a particle and two of these are associated with the relaxation of the particles magnetic moment, namely the Brownian relaxation time (τ_B) and the Néel relaxation time (τ_N). The third is the decay time associated with precession (τ_0) of the magnetic moment. All three times can be determined by means of complex susceptibility measurement, $\chi(\omega) = \chi'(\omega) - i\chi''(\omega)$, as can the presence of ferromagnetic resonance which occurs at a frequency f_{res}. However the time of particular interest here is τ_0, which can be determined by means of the unique relationship which exists between the after-effect function, $b(t)$, and the imaginary susceptibility component $\chi''(\omega)$.

Here it is demonstrated how $b(t)$, f_{res} and τ_0 are influence by the application of a polarizing field, H.

2. COMPLEX SUSCEPTIBILITY

$\chi(\omega)$, of an assembly of single domain particles can also be described in terms of its parallel, $\chi_{||}(\omega)$, and perpendicular, $\chi_\perp(\omega)$, components, with [1]

$$\chi(\omega) = \frac{1}{3}\left(\chi_{||}(\omega) + 2\chi_\perp(\omega)\right). \tag{1}$$

$\chi_{||}(\omega)$, can be described by the Debye equation, with $\chi_{||}(\omega) = \frac{\chi_{||}(0)}{1 + i\omega\tau_{||}}$, where $\chi_{||}(0)$ is the static parallel susceptibility and $\tau_{||}$ is the parallel relaxation time.

Over the frequency range considered here relaxation due to Brownian rotational diffusion of the particles will, in general, be ignored and thus the $\tau_{||}$ component, will be considered to be dominated by the Néel relaxation mechanism with relaxation time τ_N.

The perpendicular or transverse susceptibility, $\chi_\perp(\omega)$, can have a resonant character, this effect being indicated by a transition in the value of $\chi'(\omega)$ from a $+ve$ to a $-ve$ quantity at an angular frequency, $\omega_{res} = 2\pi f_{res}$.

Under equilibrium conditions, the magnetic moment, m, and the anisotropy field, H_A, of a particle are parallel and any deviation of the magnetic moment from the easy axis direction results in the precession of the magnetic moment about this axis. If the polar angle θ is small, ω_{res}, is given by [1],

$$\omega_{res} = 2\pi f_{res} = \mathrm{g}\gamma H_A \tag{2}$$

$H_A = 2K/M_s$, where K is the anisotropy constant in J/m^3 and γ is the gyromagnetic ratio. If a radio frequency field is applied perpendicular to H_A, the motion of the magnetic moment has a typical resonant character which is commonly described by means of the Landau and Lifshitz equations, namely,

$$\frac{\chi(\omega)}{\chi(0)} = \frac{\left(1 + \alpha^2\right)\omega_0^2 + i\alpha\omega\omega_0}{\left(1 + \alpha^2\right)\omega_0^2 - \omega^2 + 2i\alpha\omega\omega_0} \tag{3}$$

where α is a damping parameter γ and the precessional decay time, $\tau_0 = (\alpha\omega_0)^{-1}$.

The after-effect function, $b(t)$, represents the decay of magnetization after the sudden removal of an external polarizing magnetic field, and $\chi(\omega)$ and $b(t)$ are related by the expression [2],

$$b(t) = 2\mathrm{Re}\left\{F^{-1}\left[\frac{\chi''(\omega)}{\omega}\right]\right\} \tag{4}$$

where F^{-1} denotes the inverse Fourier transform.

Scaife [3] has shown that $b(t)$ for the Landau and Lifshitz equations, has the form,

$$b(t) = b(0)\exp\left(-\frac{t}{\tau_0}\right)\cos\omega_0 t \tag{5}$$

If the area under $b(t)$, $\int_0^\infty b(t)dt = B$ say, then it follows that, $B = b(0)\int_0^\infty dt\cos\omega_0 t\exp(-t/\tau_0)$, resulting in a normalised value of [4],

$$\frac{B}{b(0)} = \frac{(1/\tau_0)}{(1/\tau_0)^2 + \omega_0^2} = \frac{\tau_0}{1 + \omega_0^2\tau_0^2} \tag{6}$$

from whence, by knowing $B/b(0)$ and ω_0, τ_0 can be determined.

3. SUSCEPTIBILITY MEASUREMENT

Measurements reported here were made by means of the short-circuited, coaxial transmission line technique [5, 6] using a Hewlett Packard (HP) 50 Ω coaxial line incorporating a co-axial cell, in conjunction with an HP 8753C network analyser. To obtain polarised measurements the coaxial cell containing the ferrofluid sample, terminated in a standard HP short circuit load, was placed between the pole faces of an electromagnet, the axis of the cell being perpendicular to the biasing field. The biasing field, H, was altered between 0 and $104\,\mathrm{KAm}^{-1}$. Automatic swept measurements of the input impedance of the line containing the sample were measured and from these measurements the complex components, $\chi'(\omega)$ and $\chi''(\omega)$ were determined.

4. RESULTS

Measurements are presented for a 330 Gauss fluid consisting of $\mathrm{Mn}_{0.5}\mathrm{Fe}_{0.5}\mathrm{Fe}_2\mathrm{O}_4$ particles suspended in Isopar M. Measurements were performed over the range 100 MHz to 6 GHz and for 10 values of H. The curves where then fitted [7] up to 10 GHz in order to facilitate a more accurate determination of the inverse Fourier transform of $\chi''(\omega)$.

Figure 1 shows a plot of the $\chi'(\omega)$ and $\chi''(\omega)$ components of the sample and from these plots it can be seen that when $H = 0$, $f_{res} = 1.45\,\mathrm{GHz}$, and when $H = 104\,\mathrm{KAm}^{-1}$, f_{res} rises to 5.7 GHz. The $\chi''(\omega)$ component has a high frequency loss peak that occurs at a frequency of $f_{max} = 0.92\,\mathrm{GHz}$ at $H = 0$. This shifts to 5.7 GHz at $H = 104\,\mathrm{KAm}^{-1}$. Thus the value of f_{max} approaches the value of f_{res} as resonance becomes the dominant process.

A plot of f_{res} against H is shown in Fig. 2 and as $\omega_{res} = 2\pi f_{res} = \gamma(H + H_A)$, the value of H_A is found from the intercept of Fig. 2 and determined as being equal to $37\,\mathrm{KAm}^{-1}$, corresponding to a mean value of K, of $3.5\ 10^3\,\mathrm{J/m}^3$. Form the slope of Fig. 2. γ is found to be $2.5\ 10^5\,\mathrm{s}^{-1}\mathrm{A}^{-1}\mathrm{m}$.

Figure 3 shows the $\chi''(\omega)$ data which was transformed to obtain the $b(t)$ profiles shown in Fig. 4 and one can observe how, over the polarizing field range, $b(t)$ changes from an exponential type decay to an oscillatory one. This transition arises because with increasing H, the parallel relaxation component diminishes its contribution to the overall susceptibility and $b(t)$ becomes similar to that of the Landau Lifshitz form.
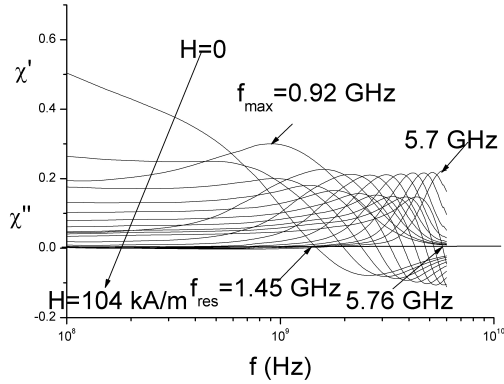
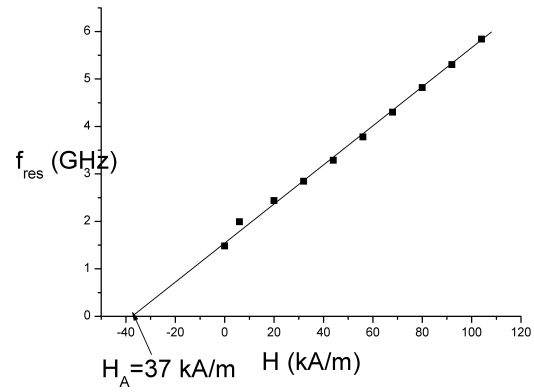Figure 1: Plot of χ' and χ'' against f (Hz).
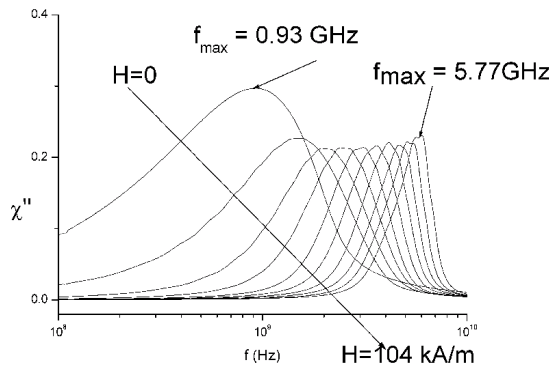

Figure 2: Plot of f_{res} against H.

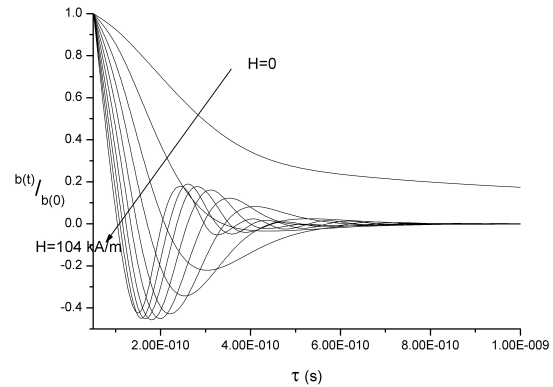Figure 3: Plot of fit to χ'' against f (Hz).

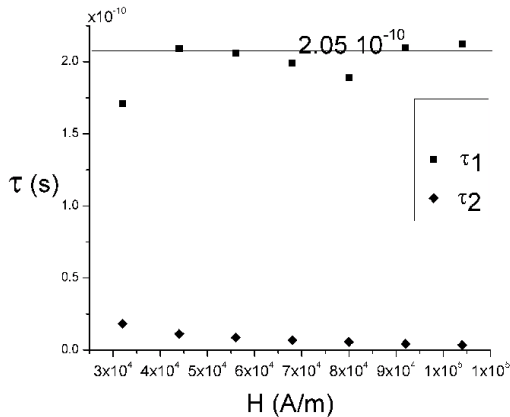Figure 4: Plot of $b(t)$ against t sec.

Figure 5: Plot of τ_1 and τ_2 against H.
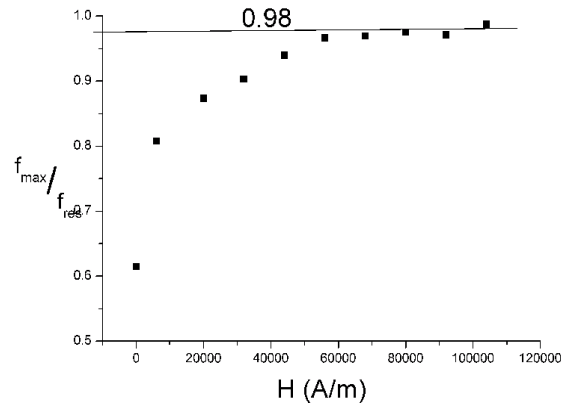

Figure 6: Plot of $f_{\max}/f_{res}$ against H.

Computing the area, B, under the $b(t)$ curves and using the f_{res} values, τ_0 as a function of H, was determined by use of Eq. (6). The results are shown plotted in Fig. 5. As can be observed from Fig. 5, for the analysed sample, Eq. (6) has real solutions only for polarizing fields larger than 60 KA/m. This is due to the fact that the approximation used in computing the solutions of Eq. (6) (i.e., $\omega_0 = 2\pi f_{res}$) is valid only for strong polarizing fields. One also notes that, Eq. (6) is a quadratic equation and has two solutions, τ_1 and τ_2 The correct values of τ_0 can be chosen by simply testing the values of τ_1 and τ_2 with the relation $\alpha = (\tau_0\omega_0)^{-1}$, where α cannot be larger that one. Performing this test for all solutions of τ_1 and τ_2, we determine that the correct solutions of Eq. (6) are those of τ_1 (see Fig. 5) and at large values of H a mean value of $\tau_0 = 2.05 \cdot 10^{-10}$ s is obtained.

In order to determine that the value of α for our sample is indeed less than 1, we have also

determined α by an alternative method, which is described in Ref. [8]. Here it is shown that in the region of strong polarizing fields,

$$\alpha = \sqrt{\frac{1 - (f_{\max}/f_{res})^2}{1 + (f_{\max}/f_{res})^2}} \tag{7}$$

From Fig. 6, one can observe that in polarizing fields larger than $60\,\mathrm{KA/m}$, the ratio $f_{\max}/f_{res}$ tends to have a constant value of 0.98 and using this value in Eq. (7) results in $\alpha = 0.14$. Thus verifying the criteria that α of the sample should be less than 1.

5. CONCLUSIONS

Using the frequency and polarizing field dependencies of the complex magnetic susceptibility, $\chi(\omega, H) = \chi'(\omega, H) - i\chi''(\omega, H)$, and the corresponding after-effect functions, the determination and subsequent observation of the behaviour of the resonant frequency, f_{res}, and the precessional decay time, τ_0, under the strong polarizing fields has been presented. The significance of the area, B, under the after-effect functions has been highlighted and it has been shown that τ_0 can be determined from the expression $\frac{B}{b(0)} = \frac{\tau_0}{1+\omega_{res}^2\tau_0^2}$; this method being based on the analysis of the after effect function of a magnetic fluid at resonance.

The determined value of $\tau_0 = 2.05 \cdot 10^{-10}\,\mathrm{s}$ lies within the generally accepted range for τ_0.

ACKNOWLEDGMENT

REFERENCES

1. Raikher, Y. L and M. I. Shliomis, *Sov. Phys. JETP*, Vol. 40, 526, 1975.
2. Fannin, P. C., "Wideband measurement and analysis techniques for the determination of the frequency-dependent, complex susceptibility," *Adv. in Chem. Phys.*, Vol. 104, 181, 1988.
3. Scaife, B. K. P., "On the theory of the complex frequency-dependent susceptibility of magnetic fluids," *Adv. in Chem. Phys.*, Vol. 109, 1, 1999.
4. Fannin, P. C., C. N. Marin, and C. Couper, *J. Magn. Magn. Mater*, In Press.
5. Roberts, S. and A. R. von Hippel, *J. App. Phys.*, Vol. 17, 610, 1946.
6. Fannin, P. C., T. Relihan, and S. W. Charles, *J. Phys. D: Appl. Phys.*, Vol. 28, 10, 2003, 1995.
7. Fannin, P. C., T. Relihan, and S. W. Charles, *Phys. Rev. B*, Vol. 55, 14423, 1997.
8. Fannin, P. C. and C. N. Marin, *J. Magn. Magn. Mater*, Vol. 299, 425, 2006.

CFD Simulation of Gravitational Sedimentation and Clustering Effects on Heat Transfer of a Nano-ferrofluid

A. Jafari[1], S. M. Mousavi[2, 3], T. Tynjala[1], and P. Sarkomaa[1]

[1]Laboratory of Engineering Thermodynamics, Lappeenranta University of Technology
Lappeenranta, Finland
[2]Biotechnology Group, Chemical Engineering Department, Tarbiat Modares University
Tehran, Iran
[3]Department of Chemical Engineering, Lappeenranta University of Technology
Lappeenranta, Finland

Abstract— In this attempt heat transfer behavior of one type of nanofluids in a disc geometry has been investigated using computational fluid dynamics (CFD) technique. Effects of gravity induced sedimentation and magnetic particle's diameter on thermal treatment of the fluid have been studied. In order to compare the effect of single phase approximation and mixture model on prediction of nanofluid' behavior, both methods were applied in simulations. It was obtained that in the presence of gravity because of natural convection heat transfer will enhance. In addition, gravity causes separation of phases and also sedimentation in the nanofluid. It was also confirmed that at value of the ratio of thermal energy to dipole-dipole contact energy more than unity, thermal forces are able to disrupt agglomerates.

1. INTRODUCTION

Nanofluids are produced by adding only a small amount of nanoparticles or nanotubes (up to 10%) into the base fluid [1]. The presence of these small particles has been shown to increase the heat transfer of the base fluid. As Keblinski et al. [2] explored there are different factors influencing the heat transport capability of nanofluids such as: Brownian motion of nanoparticles; molecular-level layering of the liquid at the nanoparticle surface; nature of heat transport in nanoparticles; and the effects of nanoparticles clustering.

Due to the high surface energy of nanoparticles, it is easy for nanoparticles to coagulate and difficult to disperse in water. Therefore, controlling the coagulation of nanoparticles in the nanofluid has become the primary issue for the initial research of nanofluid. It is also important to investigate dispersion and stability of nanofluids in order to exploit their potential benefits and applications [3]. It has been established that the mean droplets coalescence time is proportional to droplet's diameter, i.e., $t_c \sim d^n$ where n varies in the range of 0.6–3.15 [4]. The colloidal instability is one of the biggest problems that make the commercial use of nanofluids difficult. Aggregated nanoparticles cause sedimentation which lead to non homogenous dispersion of particles.

Sedimentation phenomena affect the distribution of the particle concentration in the flow. As a result of the sedimentation of particles, a layer of the sedimented particles will form on the wall of the channel. The form of this layer is nonsymmetrical in horizontal channels [5]. In the midst of a sedimenting suspension, regions of particle density higher than average will constantly be formed and destroyed [6].

Nano-ferrofluids are one type of nanofluids which are suspension of magnetic nanoparticles in a carrier liquid such as water or kerosene [7]. In the presence of an external magnetic field, a ferrofluid becomes magnetized as the particles align with the magnetic field [8]. So they have a wide range of potential application in biomedicine and technology. In ferrofluids transfer of particles in the presence of temperature gradients and magnetic fields [9], and also settling of particles and their aggregates [10] lead to heterogeneity of the fluid. As stability of such fluids is an important property that insures the investigator of a well-defined material suitable for fluid applications, the main objective of this work was to study effect of gravitational sedimentation and aggregation on heat transfer of a kerosene based ferrofluid using CFD technique. In addition application of both single and two phase methods was studied.

2. MATHEMATICAL FORMULA AND NUMERICAL METHOD

In the literature two approaches, single phase model and mixture method, have been applied to investigate the heat transfer characteristics of nanofluids. The first approach assumes that

the continuum assumption is still valid for fluids with suspended nano size particles. The other approach uses a two-phase model for description of both the fluid and the solid phases. In this study effect of these two models have been compared with each other. For more details about formula used in this study refer to [11, 12].

As Fig. 1 illustrates, a disc with a height and diameter of 3.5 mm and 75 mm, respectively was used in simulations. A grid independency check has been performed, and finally a grid with 6×10^5 tetrahedral elements, and 10^5 nodes was chosen for performing simulations. Simulations were done using a commercial software, Fluent, and a user defined function (UDF) was added to apply a uniform external magnetic field parallel to the temperature gradient. Constant temperature boundary conditions were applied for both bottom and top of the disc, and sidewall was insulated.
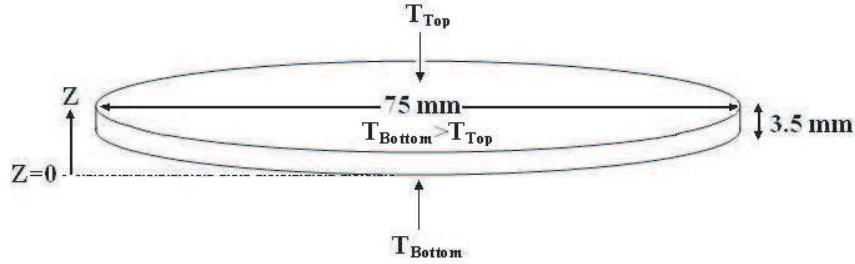

Figure 1: Schematic of the geometry used in this study.

The solver specifications for the discretization of the domain involve the presto for pressure and second-order upwind for momentum and energy in both models. In addition first-order upwind was used for volume fraction in mixture model. The under-relaxation factors, which are significant parameters affecting the convergence of the numerical scheme, were set to 0.5 for the pressure, 0.7 for the momentum, and 0.5 for the volume fraction. Using mentioned values for the under-relaxation factors a reasonable rate of convergence was achieved.

3. RESULTS AND DISCUSSION

A kerosene-based magnetic fluid with magnetization 48 kA/m, particle magnetic moment 2.5×10^{-19} Am^2, vacuum permeability $4\pi \times 10^{-7}$ H/m and thermal expansion coefficient $0.00086 \, 1/K$ was used in this study. Other properties are listed in Table 1.

Table 1: Properties of the studied ferrofluid (**c** and **p** denote continuous and dispersed phase, respectively).

Model					
Single		**Mixture**			
Property	**Value**	**Property (c)**	**Value (c)**	**Property (p)**	**Value (p)**
Thermal conductivity	0.22 W/mK	Thermal conductivity	0.149 W/mK	Thermal conductivity	1 W/mK
Dynamic viscosity	0.008 kg/ms	Dynamic viscosity	0.0024 kg/ms	Dynamic viscosity	0.03 kg/ms
Heat capacity at constant pressure	3259 J/kgK	Heat capacity at constant pressure	2090 J/kgK	Heat capacity at constant pressure	4000 J/kgK
Density	1250 kg/m^3	Density	$1248 - 1.56 \times T$ kg/m^3	Density	5400 kg/m^3

Magnetic Rayleigh number, Ra_m, is a good criterion to characterize convection. Ra_m is defined as follow:

$$Ra_m = \frac{\mu_0 \beta_m \mathbf{M}_0 \Delta T l^2 \Delta \mathbf{H}}{\rho \upsilon k} \qquad (1)$$

where μ_0, β_m, $\mathbf{M}_0$, ΔT, l, $\Delta \mathbf{H}$, ρ, υ, and k represents magnetic permeability in vacuum, pyromagnetic coefficient, magnetization vector at initial condition, the temperature difference over the gap,

a characteristic length equal to the gap width, the magnetic field difference over the gap, density, kinematic viscosity, and thermal diffusivity, respectively.

Figure 2 represents effect of gravity on heat transfer of the ferrofluid using both single and two phase methods. As this figure shows for both methods in the presence of gravity because of natural convection the heat transfer will enhance. Simulations for the mixture model were performed at particle diameter 10 nm and solid volume fraction 0.1. There are few experimental data to validate the single phase approximation and two phase method. As the particle–liquid interaction and the movement between the particle and liquids play important roles in affecting the convective heat transfer performance of nanofluids, it seems that mixture method can better predict the behavior of such fluids. But some researchers [13–15] deal with the nanofluids as single-phase fluid because it is much simpler and computationally more efficient.
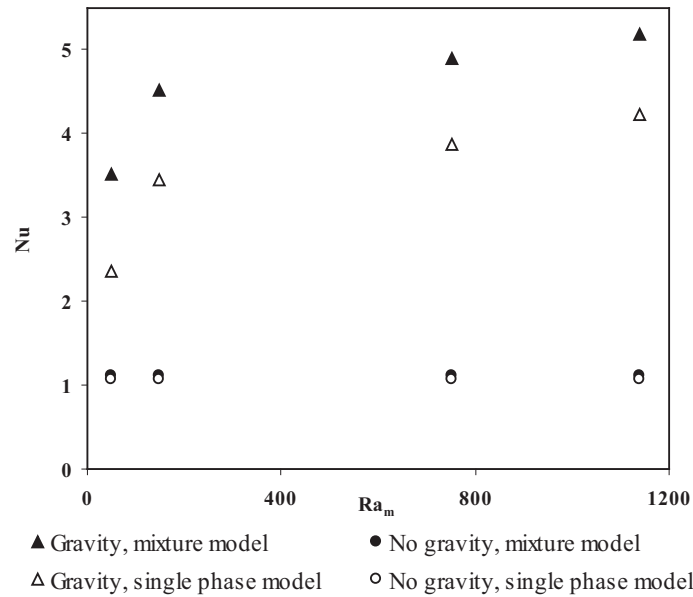

Figure 2: Effect of gravity on Nusselt number using mixture and single phase methods.

In order to better illustrate the effect of gravity on heat transfer of the ferrofluid, temperature contours in the presence and absence of gravity has been presented in Fig. 3. As this figure shows in the presence of gravity, the system is more non-uniform and sedimentation can occur.
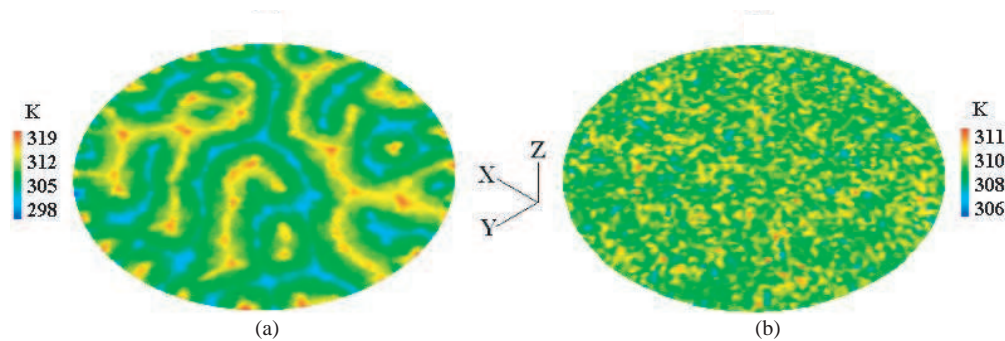

Figure 3: Temperature Contours at $z = 0.00175$ m in the (a) presence and (b) absence of gravity.

Gravitation induces phase separation; therefore it causes particle sedimentation within the nanofluid. In other words, there is an interaction between gravity-and suspension-driven sedimentation. Sedimentation can damage the stability of the system, and then lead to oscillatory convection. This phenomenon will stop as phase equilibrium is approached. For ferrofluids the following ratio can describe the relative influence of gravity to magnetism [7];

$$\frac{Gravitational\ energy}{Magnetic\ energy} = \frac{\Delta\rho\mathbf{g}l}{\mu_0\mathbf{MH}} \tag{2}$$

if the above ratio is sufficiently smaller than 1, therefore the gravity is less of a threat to sedimentation.

The presence of gravity sedimentation of particles causes the nanofluid to be more heterogeneous. Fig. 4 shows separation of phases because of gravity using the mixture model at $\Delta T = 38\,\mathrm{K}$, and $\mathbf{H} = 20\,\mathrm{kA/m}$. Not only the nanoparticles move around, but also significantly large bodies of fluid around nanoparticles interact with each other, which leads to a strong micro convection.
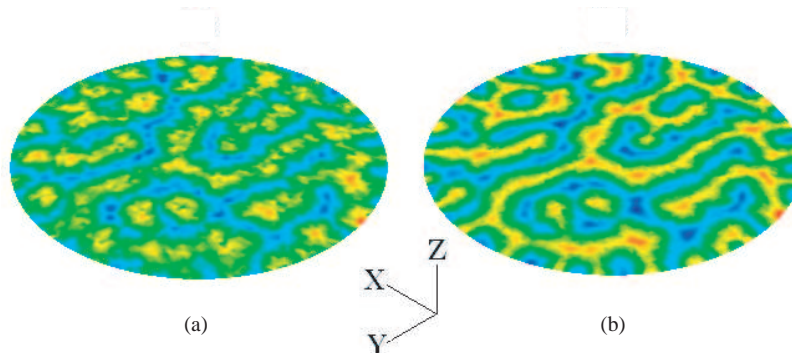

Figure 4: Solid volume fraction of dispersed phase (magnetic particles) on a z-plane after 100 s. (a) $z = 0$ and (b) $z = 0.00175\,\mathrm{m}$. Red and blue colors refer to the maximum and minimum values.

In order to investigate the influence of aggregation on heat transfer, particle's diameter 50 nm and solid volume fraction 0.1 was used in simulations, and results compared with the case contained particle's diameter 10 nm. The ratio of thermal energy to dipole-dipole contact energy can govern the stability of the system. To escape agglomeration, this ratio must be greater than unity [7], i.e.,

$$\frac{Thermal\ energy}{Dipole - dipole\ contact\ energy} = \frac{12k_BT}{\mu_0\mathbf{M}^2V} > 1 \tag{3}$$

where k_B, and V represent Boltzmann's constant and volume for a spherical particle, respectively.

As in our simulations the above ratio was greater than 1, the thermal agitation was able to disrupt agglomerates, so the amount of heat transfer in both cases was the same. In the previous work [16] it was shown that in the presence of magnetic field greater than 20 kA/m particle's diameter can affect the heat transfer, because dipole-dipole contact energy is significant and the above ratio is smaller than 1.

4. CONCLUSION

In this study the effect of gravity on heat transfer of kerosene based nano-ferrofluid has been investigated using CFD tools. Significance of single phase and mixture models on prediction of nanofluid's behavior also was studied. Results show that gravity induces sedimentation and therefore causes the system to be more unstable. It was obtained that for both mathematical models in the presence of gravity because of natural convection heat transfer will enhance. It was confirmed to escape agglomeration; thermal energy to dipole-dipole contact energy ratio must be greater than unity.

REFERENCES

1. Choi, S. U. S., "Enhancing thermal conductivity of fluids with nanoparticles," *Developments and Applications of Non-Newtonian Flows*, Siginer, D. A. and Wang, H. P. (Eds.), FED-Vol. 231/MD-Vol. 66, 99–105, The American Society of Mechanical Engineers, New York, 1995.
2. Keblinksi, P., S. R. Phillpot, S. U. S. Choi, and J. A. Eastman, "Mechanisms of heat flow in suspensions of nano-sized particles (nanofluids)," *International Journal of Heat and Mass Transfer*, Vol. 45, 855–863, 2002.
3. Li, X., D. Zhu, and X. Wang, "Evaluation on dispersion behavior of the aqueous copper nano-suspensions," *Journal of Colloid and Interface Science*, Vol. 310, 456–463, 2007.
4. Thornton, J. D., *Science and Practice of Liquid-Liquid Extraction*, Vol. 1, 255–337, Clarendon Press, Oxford, 1992.
5. Sarimeseli, A. and G. Kelbaliyev, "Sedimentation of solid particles in turbulent flow in horizontal channels," *Powder Technology*, Vol. 140, 79–85, 2004.

6. Abade, G. C. and F. R. Cunha, "Computer simulation of particle aggregates during sedimentation," *Comput. Methods Appl. Mech. Engrg.*, Vol. 196, 4597–4612, 2007.

7. Rosensweig, R. E., *Ferrohydrodynamics*, 2nd ed., Dover Publications Inc., New York, 1997.

8. Borglin, S. E., G. J. Moridis, and C. M. Oldenburg, "Experimental studies of the flow of ferrofluid in Porous media," *Transport in Porous Media*, Vol. 41, No. 1, 61–80, 2000.

9. Blums, E., M. M. Maiorov, and A. Cebers, *Magnetic Fluids*, Zinatne, Riga, 1989.

10. Bozhko, A. A. and G. F. Putin, "Heat transfer and flow patterns in ferrofluid convection," *Magnetohydrodynamics*, Vol. 39, 147–169, 2003.

11. Jafari, A., T. Tynälä, S. M. Mousavi, and P. Sarkomaa, "CFD simulation and evaluation of controllable parameters effect on thermomagnetic convection in ferrofluids using Taguchi technique," *Computers and Fluids*, Vol. 37, 1344–1353, 2008.

12. Tynjala, T., "Theoretical and numerical study of thermomagnetic convection in magnetic fluids," PhD Thesis, Lappeenranta University of Technology Press, Finland, 2005.

13. Xuan, Y. and W. Roetzel, "Conceptions for heat transfer correlation of nanofluids," *International Journal of Heat and Mass Transfer*, Vol. 43, 3701–3707, 2000.

14. Khanafer, K., K. Vafai, and M. Lightstone, "Bouyancy-driven heat transfer enhancement in a twodimensional enclosure utilizing nanofluids," *International Journal of Heat and Mass Transfer*, Vol. 46, 3639–3653, 2003.

15. Maiga, S. E. B., C. T. Nguyen, N. Galanis, and G. Roy, "Heat transfer behaviours of nanofluids in a uniformly heated tube," *Superlattices and Microstructures*, Vol. 35, 543–557, 2004.

16. Jafari, A., T. Tynjälä, S. M. Mousavi, and P. Sarkomaa, "Simulation of heat transfer in a ferrofluid using computational fluid dynamics technique," *International Journal of Heat and Fluid Flow*, Vol. 29, 1197–1202, 2008.

Electronic Properties of Quantum Wells Structures with Gaussian Potential Profiles

S. Jelev-Vlaev, **A. Enciso-Muñoz**, and **D. A. Contreras-Solorio**

Unidad Académica de Física, Universidad Autónoma de Zacatecas
Calzada Solidaridad Esquina con Paseo La Bufa s/n, C. P. 98060, Zacatecas, ZAC., México

Abstract— The first semi-empirical tight-binding calculation of the electronic properties of some quantum wells structures with Gaussian confining potential is presented. The obtained results demonstrate that this methodology can be useful for realistic calculations of the electronic structure in this kind of systems.

1. INTRODUCTION

The design of many semiconductor devices containing quantum wells structures is based on numerical calculations of their electronic properties [1]. The electronic spectrum of these heterostructures depends on the potential profile in the growth direction. A lot of confinement potentials have been exploited in the applications. The Gaussian profile causes specific features in the electronic structure of quantum wells and superlattices which have been used in some opto-electronic devices [2–9]. Only simple one-dimensional models based on effective mass approximation have been applied for these systems [2]. More realistic calculations can be done using the methodology of this work. We take into account the real crystal and atomic structure of Gaussian shaped quantum wells, barriers and superlattices based on $GaAs$ and $AlGaAs$ materials. The numerical calculations reveal the specific features in the electronic spectrum of these systems.
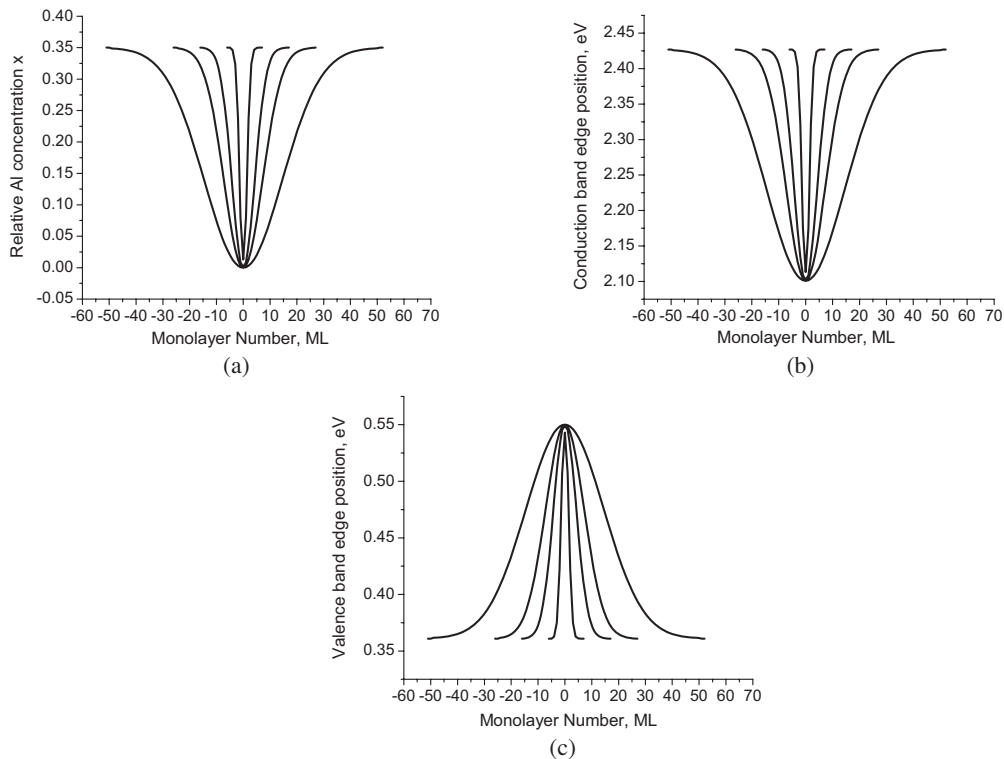

Figure 1: Concentration and potential profiles of Gaussian shaped quantum wells (GQWs) of different widths. The total well width in MLs: 10; 30; 50; 100. (a) Relative Al concentration profiles $x(z)$. (b) Electron potential profiles $Ec[x(z)]$. (c) Hole potential profiles $Ev[x(z)]$. The zero of the energy is fixed at the top of the $AlAs$ valence band.

2. MODEL AND METHOD

A single Gaussian quantum well (GQW) and barrier (GQB), as well as finite and infinite superlattices consisting of GQW and GQB of different widths are considered. The sp₃s* spin dependent semi-empirical tight-binding model taking into account nearest-neigbors is used and the surface Green function matching method (SGFM) within the framework of the virtual crystal approximation is applied to treat these large and inhomogeneous planar heterostructures [10–12]. The numerical calculations are fast and precise due to the application of two efficient algorithms for inversion of large block tri-diagonal matrices [10, 12]. The first algorithm [10] is exact and the second [12] is iterative based on the decimation technique. We consider $Al[x(z)]Ga[1 - x(z)]As$ materials where the Al concentration $x(z)$ depends on the growth coordinate z and the functional dependence is of Gaussian type. Due to the linear dependence of the conduction and valence band borders on the Al concentration x, the confinement potentials for electrons and holes obey Gaussian profiles too [11] (See Fig. 1).
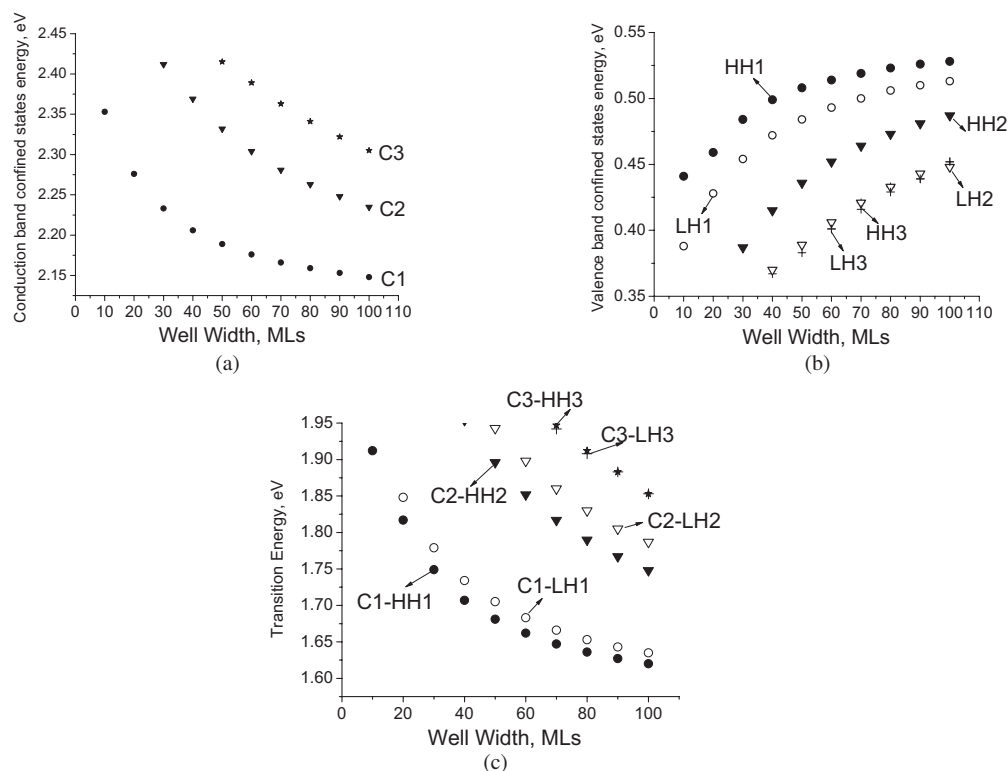

Figure 2: Energies of the electronic confined states and the transitions between them in Gaussian shaped quantum wells (GQWs) of different widths. First (ground) state and two excited states are considered. (a) Conduction band states C1, C2, C3. (b) Valence band states HH1, LH1; HH2, LH2; HH3, LH3. (c) Transitions C1-HH1, C1-LH1; C2-HH2, C2-LH2; C3-HH3, C3-LH3. The zero of the energy is fixed at the top of the $AlAs$ valence band.
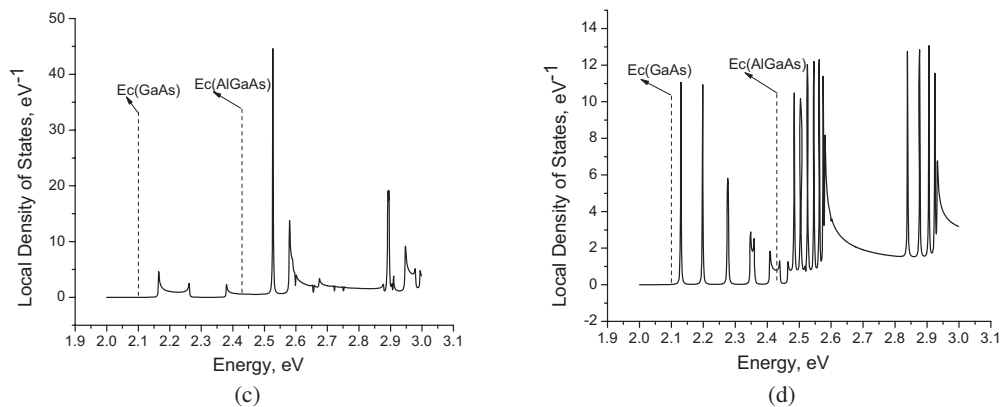
(c)

(d)
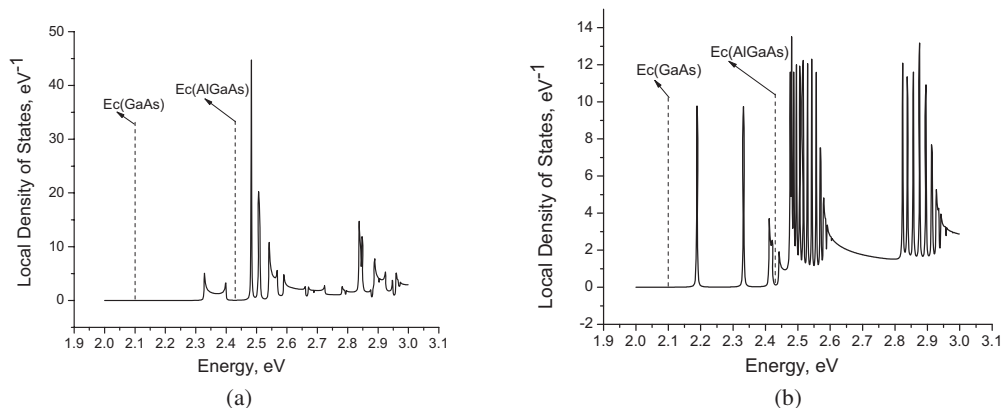
Figure 3: Conduction mini-band electronic structure of *infinite* Gaussian shaped superlattices. The *Al* concentration in $Al(x)Ga(1-x)As$ is x=0.35. (a) The superlattice period: $2Al(x)Ga(1-x)As$ /10GQW/ $2Al(x)Ga(1-x)As$.; (b) The superlattice period: $Al(x)Ga(1-x)As$ /50GQW/ $2Al(x)Ga(1-x)As$.; (c) The superlattice period: $2GaAs$ /10GQB/ $GaAs$. (d) The superlattice period: $2GaAs$ /50GQB/ $2GaAs$. The zero of the energy is fixed at the top of the *AlAs* valence band.

3. RESULTS AND DISCUSSION

The energies of the confined electron and hole states and the energies of the main optical interband transitions are shown in Fig. 2 for GQW of widths from 10 to 100 MLs.

The conduction mini-band formation of *infinite* superlattices with different periods is presented in Fig. 3.
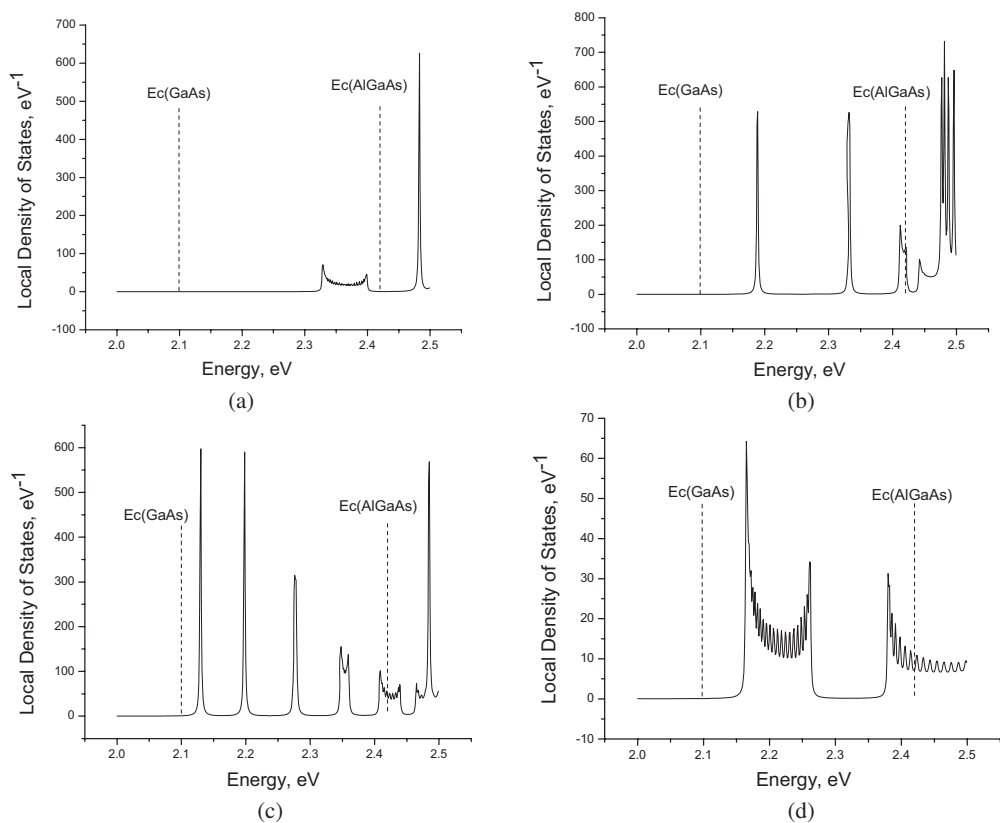

(a)

(b)

(c)

(d)

Figure 4: Conduction mini-band electronic structure of *finite* Gaussian shaped superlattices. The *Al* concentration in $Al(x)Ga(1-x)As$ is x=0.35. (a) 50 periods: $2Al(x)Ga(1-x)As$ /10GQW/ $2Al(x)Ga(1-x)As$; (b) 20 periods: $2Al(x)Ga(1-x)As$ /50GQW / $2Al(x)Ga(1-x)As$; (c) 50 periods: $2GaAs$ /10GQB /$2GaAs$; (d) 20 periods: $2GaAs$ /50GQB/ $2GaAs$. The zero of the energy is fixed at the top of the *AlAs* valence band.

The dashed vertical lines show the energy positions of the conduction band edges for *GaAs* (2.101 eV) and *AlGaAs* (2.429 eV).

The conduction mini-band formation of *finite* superlattices with different periods is presented in Fig. 4. The dashed vertical lines show the energy positions of the conduction band edges for *GaAs* (2.101 eV) and *AlGaAs* (2.429 eV). The finite number of periods causes the differences between Fig. 4 and Fig. 3, but the main features of the superlattice mini-bands are similar.

The distances between the GQW (GQB) in the superlattices determine the degree of the energy splitting of the interacting electronic levels. The mini-band widths and positions agree with the electronic structure of the single GQW (GQB) and the general properties of the superlattices. The parameters of the superlattice (period, number of periods, interwell distance) act as tuning parameters in the design of Gaussian shaped superlattices. The intensive maxima and pseudogaps observed in the electronic spectra agree with the results obtained in [1–3].

4. CONCLUSIONS

The methodology used in this work permits a realistic treatment of quantum wells structures with Gaussian potential profile. The calculations conducted in the present study reproduce well the known specific features of these heterostructures and show that this approach can be used for theoretical band-gap engineering of systems with this kind of confining potential.

REFERENCES

1. Singh, J., *Physics of Semiconductors and Their Heterostructures*, McGraw-Hill, Inc., New York, 1993.
2. Gómez, I., F. Domínguez-Ádame, E. Diez, and V. Bellani, "Electron transport across a Gaussian superlattice," *J. Appl. Phys.*, Vol. 85, No. 7, 3916–3921, 1999.
3. Asakura, K., M. Suzuki, H. Sanada, and N. Nagai, "Band structure estimation for Gaussian superlattices by applying image parameters in circuit theory," *2000 IEEE International Symposium*, Vol. 4, No. 3, 533–536, 2000.
4. Banfi, F., V. Bellani, I. Gomez, E. Diez, and F. Dominguez-Ádame, "Interface roughness effects in Gaussian superlattices," *Semicond. Sci. & Technology*, Vol. 16, No. 5, 304–309, 2001.
5. Lara, G. A. and P. Orellana, "Resonant tunneling through Gaussian superlattices," *Rev. Mex. Fis.*, Vol. 48, No. 3, 40–42, 2002.
6. Sanada, H. and K. Watanabe, "A study on electron-wave filters using barrier height modulated multiple barrier structures," *Electronics and Communications in Japan (Part II: Electronics)*, Vol. 86, No. 9, 11–19, 2003.
7. Gomez, I., E. Diez, F. Dominguez-Ádame, and V. Bellani, "Inusual behaviour of the conductance in Gaussian superlattices," *Physica E: Low-dimensional Systems and Nanostructures*, Vol. 23, No. 1–2, 65–69, 2004.
8. Villegas, D., F. del León-Pérez, and R. Pérez-Alvarez, "Gaussian superlattice for phonons," *Microelectronics J.*, Vol. 36, No. 3–6, 411–412, 2005.
9. Asakura, K., H. Sanada, O. Ogurisu, and M. Suzuki, "Practical band estimation for periodic superlattices by using semi-infinite periodic model," *IEEE Int. Symposium on Communications and Information Technologies*, 932–935, 2007.
10. Vlaev, S., V. R. Velasco, and F. García-Moliner, "Electronic states in graded-composition heterostructures," *Phys. Rev. B*, Vol. 49, No. 16, 11222–11229, 1994.
11. Vlaev, S. and D. A. Contreras-Solorio, "Electronic states in diffused quantum wells," *J. Appl. Phys.*, Vol. 82, No. 8, 3853–3856, 1997.
12. Vlaev, S. J., I. Rodríguez-Vargas, L. M. Gaggero-Sager, and V. R. Velasco, "An alternative way of calculating a superlattice Green function for discrete media," *Surface Sci.*, Vol. 554, No. 2–3, 245–252, 2004.

Total Density of States in Rectangular Quantum Wells

S. Jelev-Vlaev[1], Romeo de Coss[2],
A. Del Río de Santiago[1], and J. C. Martínez-Orozco[1]

[1]Unidad Académica de Física, Universidad Autónoma de Zacatecas
Calzada Solidaridad Esquina con Paseo La Bufa s/n, C. P. 98060, Zacatecas, ZAC, México
[2]Departamento de Física Aplicada, CINVESTAV-Mérida
Mérida 97310, Yucatán, México

Abstract— The total density of states in rectangular $AlGaAs/GaAs$ quantum wells are calculated numerically summing in the two-dimensional Brillouin zone. The quasi-two-dimensional energy bands are observed for typical wells and the integrated spectral strengths are presented. The conditions for the formation of quantum wells for all electron K-vectors are discussed.

1. INTRODUCTION

The applications of the quantum wells are based mainly on the optical transitions around the high symmetry points in the two-dimensional Brillouin zone [1]. By this reason the electronic structure of the quantum wells has been studied extensively only in these points [2]. But all macroscopic properties of any condensed matter system depend on the integrated (total) density of states [3]. The experimental methods photoemission and inverse photoemission measure directly the total density of states of the occupied and non-occupied electronic states respectively in atoms, molecules, crystals and nanostructures [4]. There is few experimental and theoretical information about the total density of states in quantum wells [1, 2]. In the present work we have conducted numerical calculations of this quantity in typical rectangular quantum wells. The method of the special points was applied summing in 10, 36, 136, 528 and 2080 points of the two-dimensional Brillouin zone [5]. We have studied the specific properties of the electronic spectrum in these quasi-two-dimensional systems within the framework of the semi-empirical tight-binding model and the Green function formalism [6].

2. MODEL AND METHOD

The semi-empirical tight-binding model was applied within the framework of the Surface Green Function Matching (SGFM) method [6]. Nearest neighbor interactions and sp^3s^* atomic base including spin were taken into account. The (001) crystal growth direction was considered and the integration over the two-dimensional Brillouin zone was made in 10, 36, 136, 528 and 2080 special points. The bulk material $Al(x)Ga(1-x)As$ was treated in the virtual crystal approximation and the Al barrier concentration was $x = 0.35$. The well width has been varied between 10 and 50 monolayers. The small imaginary energy part was taken to be $0.01\,\text{eV}$ or $0.001\,\text{eV}$ depending
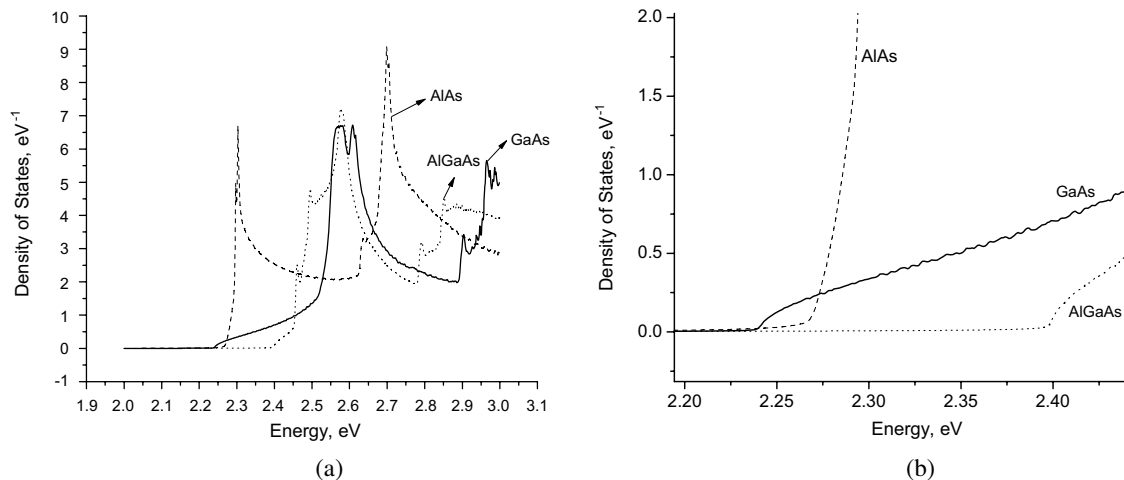

Figure 1: Total density of states for bulk $Al(x)Ga(1-x)As$, $GaAs$ and $AlAs$ in the conduction band.

on the special points number. To determine the confinement conditions we calculated the total (integrated) density of states for the homogeneous bulk materials *GaAs* and *AlGaAs*. A quantum well for all electron K-vectors appears with a barrier height of 150 meV. (See Figs. 1(a), 1(b)).

We can see in the Fig. 1 that the *AlGaAs* conduction band starts at energy 150 meV higher than the *GaAs* conduction band. This means that the electrons will be confined for all energies in the energy interval between the band edges and for all K-vectors, if a finite *GaAs* slab is sandwiched between *AlGaAs* barriers.

3. RESULTS AND DISCUSSION

In Fig. 2(a) the total density of states in the energy interval of the confinement is presented for a quantum well of 15 monolayers.
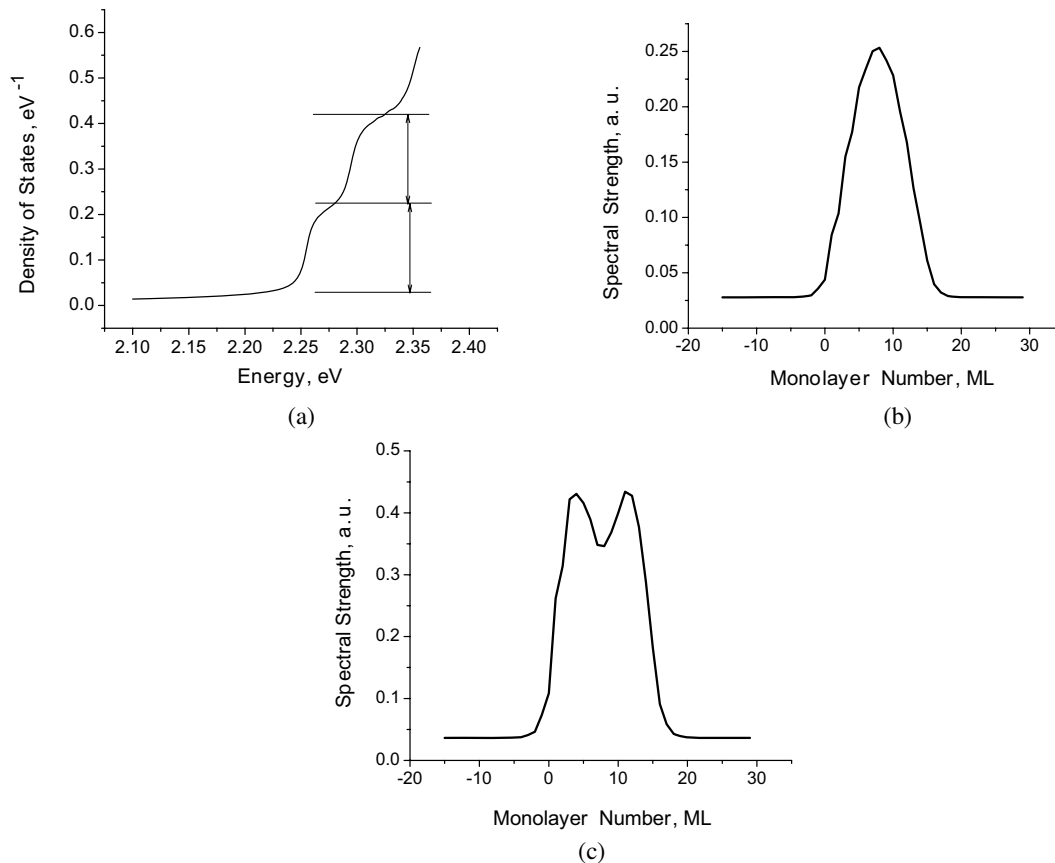

Figure 2: Integrated density of states and spatial distributions in the conduction band of rectangular $Al(x)Ga(1 - x)As/GaAs$, $x = 0.35$ quantum well of 15 monolayers. (a) Integrated (total) density of states. (b) Integrated spectral strengths of the first sub-band. (c) Integrated spectral strengths of the second sub-band. The zero of the energy is fixed at the top of the *AlAs* valence band in Gamma point.

The integration is conducted summing in 2080 special points of the Brillouin zone. There are two sub-bands of the same height with step-like behavior. It indicates that for all K-points there are no more than two bound states. The spatial distributions of the first and second sub-band are shown in Figs. 2(b), (c). The spectral strengths are localized in the well region. The features in Figs. 2(a), (b), (c) correspond to the basic properties of the rectangular quantum wells of infinite barriers [2]. There are contributions from the bulk density of states and the curves in Fig. 2 are not completely as in the ideal case. Fig. 3 presents similar results for the density of states in wells of 11 and 21 monolayers when only one bound states exists for all K-points.

The zero of the energy is fixed at the top of the *AlAs* valence band in Gamma point.
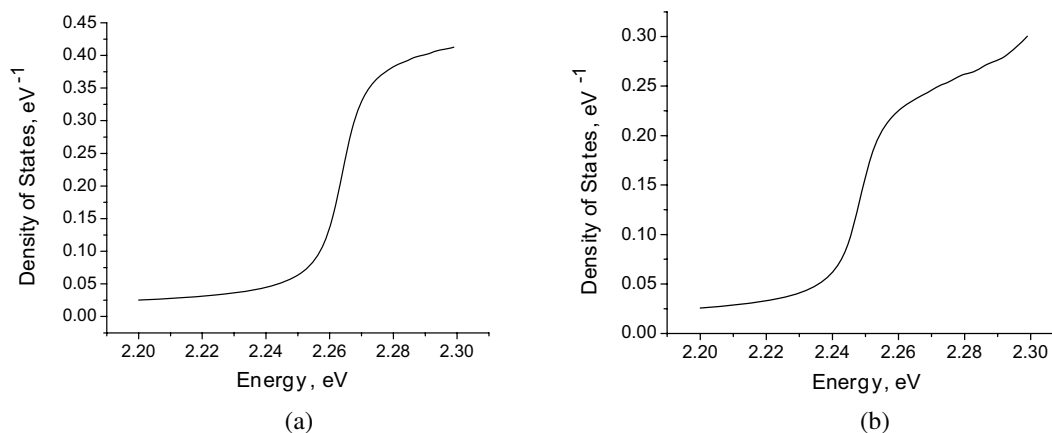
Figure 3: Total density of states for a single rectangular $Al(x)Ga(1-x)As/GaAs$, $x = 0.35$ quantum well integrating in 2080 special points over the 2DBZ. The well width is (a) 11 and (b) 21 monolayers and the projection of the density of states is done in the middle of the well.

4. CONCLUSIONS

We have studied the problem of the electron confinement in rectangular quantum wells formed from the INTEGRATED density of states of the bulk materials *GaAs* and *AlGaAs*. We have found energetic intervals where the density of states has two-dimensional properties. The integrated spectral strengths for these energies are localized in the well region.

REFERENCES

1. Singh, J., *Physics of Semiconductors and Their Heterostructures*, McGraw-Hill, Inc., New York, 1993.
2. Harrison, P., *Quantum Wells, Wires and Dots*, John Wiley & Sons Ltd., 2000.
3. Harrison, W. A., *Elementary Electronic Structure*, World Scientific, Singapore, 1999.
4. Hughes, H. P. and H. I. Starnberg, *Electron Spectroscopies Applied to Low-dimensional Materials*, Kluwer Academic Publishers, 2000.
5. Cunningham, S. L., "Special points in the two-dimensional Brillouin zone," *Phys. Rev. B*, Vol. 10, No. 12, 4988–4994, 1974.
6. Vlaev, S., V. R. Velasco, and F. García-Moliner, "Electronic states in graded-composition heterostructures," *Phys. Rev. B*, Vol. 49, No. 16, 11222–11229, 1994.

Scattering Parameters Measuring Technology Research on Two-port Surface Mounted Device

Hui Huang[1,3], **Ke Wang**[2], **Xin Meng Liu**[3], **and Xin Lv**[1]

[1]Beijing Institute of Technology, Beijing 100081, China
[2]China Electronic Standardization Institute, Beijing 100007, China
[3]National Institute of Metrology, Beijing 100013, China

Abstract— This paper derives a novel principle how to design the fixture and the standard kit from the physical structure of surface mounted device (SMD). The characteristic coefficients of the standard are described. A perfect de-embedding algorithm is implemented. Then an example of two port surface acoustic wave (SAW) filter scattering parameters (S-parameters) measurement is used to show how to design the fixture and the calibration kit by applying the de-embedding algorithm.

1. INTRODUCTION

Surface mounted technology is widely used in modern electronic and information industry, because this technology can be met with the high frequency, high integration and low power. Accurate S-parameters measurement of the surface mounted device (SMD) must be solved. To obtain the accurate results, researcher uses the vector network analyzer (VNA) calibrated with the coaxial calibration kit to measure the overall of the SMD and fixture, then compute the real and accuracy measurement results of the SMD by applying the de-embedding algorithm. The fixture must have well stability and repeatedly. The characteristic coefficients of the standard kit are known.

2. SURFACE MOUNTED DEVICE MEASUREMENT PRINCIPLE

2.1. General Description

It is easy to measure the S-parameters of the coaxial device using the VNA calibrated by the coaxial calibration kit, and the measurement accuracy is quite high. However, this method is invalid to the SMD. We use the calibrated VNA to measure the overall of the SMD and the corresponding fixture, compute the real and accuracy measurement results of the SMD by applying de-embedding algorithm. The measurement system of the SMD is shown in Figure 1.
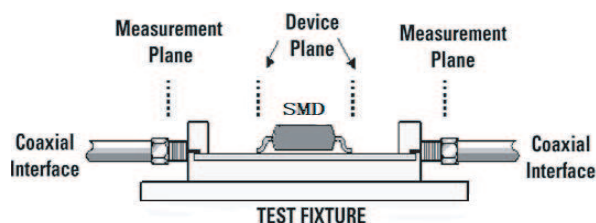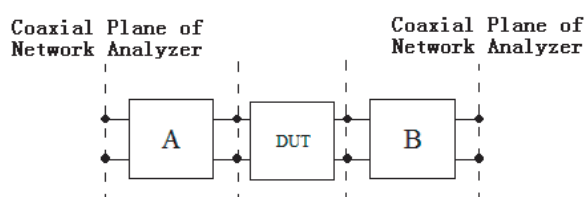

Figure 1: Schematic of measurement system of the SMD.

Figure 2: Schematic of cascade of the test fixture and DUT.

2.2. The S-parameters of SMD De-embedding Algorithm

In actual measurement procedure, we regard the test fixture as the two symmetry parts. The two symmetry parts are defined the left (A) and right (B) fixture, the device under test (DUT) is cascade between the two symmetry parts. The port of the fixture connecting the VNA is coaxial, such as N type, 3.5 mm, 2.4 mm. The other port is corresponding with the DUT, such as stripline, microstrip line, coplanar waveguide. Then the overall fixture and the DUT is cascade between the two coaxial ports of the VNA. The cascade of the fixture and the DUT is shown in Figure 2.

The measurement result of the calibrated VNA is overall of the fixture A, DUT, and fixture B. The measured result of this cascade is most easily calculated by using the cascade matrix definition (transmission-parameters). The flow graph is shown in Figure 3.
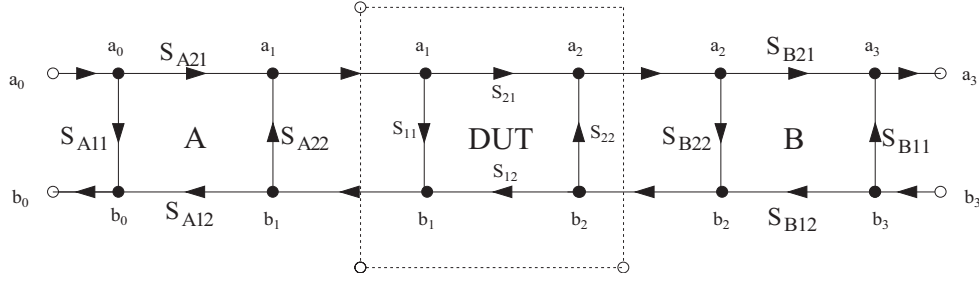
Figure 3: Flow graph of cascade of the test fixture and DUT.

Let [A], [T], [B] are cascade matrices of the S-parameters.

$$T_M = T_A T T_B = \frac{1}{(S_{A21}S_{B21})} \begin{bmatrix} -\Delta_A & S_{A11} \\ -S_{A22} & 1 \end{bmatrix} T \begin{bmatrix} -\Delta_B & S_{B11} \\ -S_{B22} & 1 \end{bmatrix} = \frac{1}{(S_{A21}S_{B21})} ATB \quad (1)$$

where, $\Delta_A = S_{A11}S_{A22} - S_{A21}S_{A12}$, $\Delta_B = S_{B11}S_{B22} - S_{B21}S_{B12}$, $\Delta_M = S_{11M}S_{22M} - S_{12M}S_{21M}$.

T is the parameters of the device under test, T_A and T_B are the parameters of the fixture on the input and output of the device under test. T_M is the overall measured result matrix.

There is another mathematical formulation for the cascade matrix. This T-parameter matrix (T) can be partitioned into the four sub matrixes T_1, T_2, T_3, T_4.

$$\begin{bmatrix} b_0 \\ b_3 \\ a_0 \\ a_3 \end{bmatrix} = \left[\frac{T_1 | T_2}{T_3 | T_4} \right] \begin{bmatrix} b_1 \\ b_2 \\ a_1 \\ a_2 \end{bmatrix} \quad (2)$$

where, $T_1 = \begin{bmatrix} -\Delta_A & 0 \\ 0 & -k\Delta_B \end{bmatrix}$, $T_2 = \begin{bmatrix} S_{A11} & 0 \\ 0 & kS_{B22} \end{bmatrix}$, $T_3 = \begin{bmatrix} -S_{A22} & 0 \\ 0 & -kS_{B22} \end{bmatrix}$, $T_4 = \begin{bmatrix} 1 & 0 \\ 0 & k \end{bmatrix}$,

$k = \frac{S_{A21}}{S_{B21}}$.

Using this approach the measured S-parameters formulation is a 'bilinear matrix equation'. It looks much the same as the one-port bilinear transformation described earlier. The equation can be easily 'inverted' to solve for the actual S-parameters. And most important the relationship can be input in linear form. The matrix equation for the two-port expand to yield 4 equations, with 4 measured S-parameters, 4 actual S-parameters, and 8 S-parameters of the fixture. Note that these 4 equations are linear with regards to the 8 S-parameters of the fixture, we get Equation (3).

$$T_1 S + T_2 - S_M T_3 S - S_M T_4 = 0 \quad (3)$$

Equation (3) can be rewritten as following,

$$S_{A11} + S_{D11}S_{MD11}S_{A22} - S_{D11}\Delta_A + S_{D21}S_{MD12}(kS_{B11}) = S_{MD11} \quad (4)$$

$$S_{D12}S_{MD11}S_{A22} - S_{D12}\Delta_A + S_{D22}S_{MD12}(kS_{B11}) - S_{MD12}k = 0 \quad (5)$$

$$S_{D11}S_{MD21}S_{A22} + S_{D21}S_{MD22}(kS_{B11}) - S_{D21}(k\Delta_B) = S_{MD21} \quad (6)$$

$$S_{D12}S_{MD21}S_{A22} + k(S_{B22}) + S_{D22}S_{MD22}(kS_{B11}) - S_{D22}(k\Delta_B) - S_{MD22}k = 0 \quad (7)$$

The first step is the matched line ($S_{T11} = 0$, $S_{T22} = 0$) connection, The second step is matched load ($S_{M11} = 0$, $S_{M22} = 0$), the last step is the same reflection coefficient termination connecting the fixture. Then $S_{A11}S_{A22}\Delta_A S_{B11}S_{B22}\Delta_B S_{A12}S_{B12}S_{A21}S_{B21}$ are 8 S-parameters of the fixture, solving for the 8 S-parameters of the fixture A and B,

$$S_{A11} = S_{MM11}, \quad S_{A22} = \frac{(S_{MS11} - S_{MM11})(S_{MM22} - S_{MT22})}{\left[(S_{MS11} - S_{MT11})(S_{MM22} - S_{MT22}) - S_{MT12}^2\right]S_{S11}},$$

$$\Delta_A = \frac{S_{MT12}^2 + S_{MT11}(S_{MM22} - S_{MT22})}{S_{MM22} - S_{MT22}}S_{A22}, \quad kS_{B11} = \frac{S_{MT11} - S_{MM11}}{S_{MT12}S_{T12}}, \quad S_{B22} = S_{MM22},$$

$$k\Delta_B = \frac{S_{MT22}(S_{MT11} - S_{MM11}) - S_{MT12}^2}{S_{T12}S_{MT12}}, \quad k = \frac{(S_{MM11} - S_{MS11})S_{MT12}S_{T12}}{\left[(S_{MS11} - S_{MT11})(S_{MM22} - S_{MT22}) - S_{MT12}^2\right]S_{S11}},$$

$$S_{A12}S_{B12} = \frac{S_{A11}S_{A22} - \Delta_A}{k}, \quad S_{A21}S_{B21} = kS_{B11}S_{B22} - k\Delta_B.$$

Now solving for the 4 S-parameters of DUT using inverted equation $T = T_A^{-1}T_M T_B^{-1}$ and solved 8 S-parameters,

$$DUT_{11} = \frac{S_{B11}(S_{A11}S_{MD22} - \Delta_{MD}) + (S_{MD11} - S_{A11})\Delta_B}{\Delta}, \quad DUT_{12} = \frac{-S_{A21}S_{B21}S_{MD12}}{\Delta},$$

$$DUT_{22} = \frac{S_{A22}(S_{B22}S_{MD11} - \Delta_{MD}) + (S_{MD22} - S_{B22})\Delta_A}{\Delta}, \quad DUT_{21} = \frac{-S_{A12}S_{B12}S_{MD21}}{\Delta},$$

$$\Delta_{MD} = S_{MD11}S_{MD22} - S_{MD12}S_{MD21}, \quad \Delta = S_{B11}(S_{MD22}\Delta_A - S_{A22}\Delta_{MD}) + (S_{A22}S_{MD11} - \Delta_A)\Delta_B.$$

2.3. The Structure and the Characteristic of the DUT

The DUT is the SBP040A surface acoustic wave (SAW) filter. The structure and characteristic are show in Figure 4. The footprint 11 is input port and the footprint 5 is output port. The others footprint are ground.
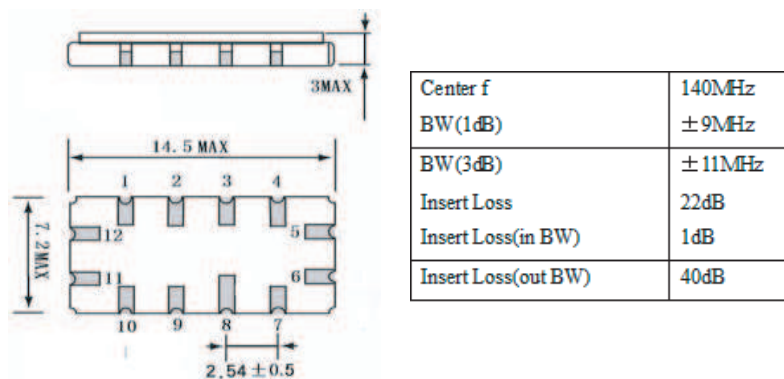

Center f	140MHz
BW(1dB)	±9MHz
BW(3dB)	±11MHz
Insert Loss	22dB
Insert Loss(in BW)	1dB
Insert Loss(out BW)	40dB

Figure 4: Schematic of the structure and the characteristic of the SAW filter SBP040A.

2.4. Designed Fixture and the Standard

The port of the designed fixture is 3.5 mm coaxial, the other port is the ground-signal-ground (GSG) transmission line structure. The gap of the ground electrode and the signal electrode is equal to the DUTs'. The whole structure of the fixture is shown in Figure 5.

The standard substrate is Al_2O_3 ($\varepsilon_r = 9.9$) and its thickness is 1 mm. The impedance of all designed standard is $50\,\Omega$, including thru standard, short standard, and load standard. The gold-filled thickness is 0.007 mm. The structure of the standard kit is shown in Figures 5 and 6.
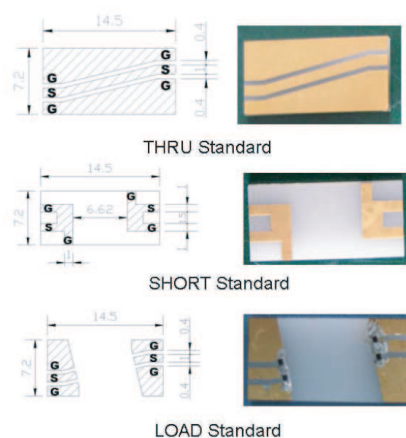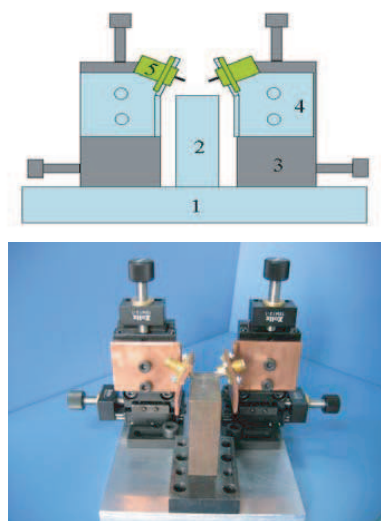

Figure 5: Schematic of the structure of the fixture.

THRU Standard

SHORT Standard

LOAD Standard

Figure 6: Schematic of the structure of the standard.

2.5. Measurement Result of the SAW Filter

The measurement result of the DUT is according with specifications of the SAW filter. The measurement system and the measurement result of the SAW filter is shown in Figure 7.
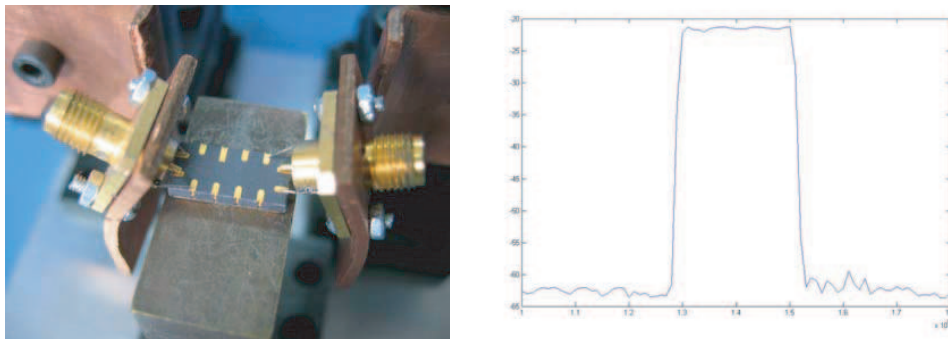
Figure 7: Schematic of measurement system and the measurement result.

3. CONCLUSIONS

We found the measurement system for the SAW filter, including VNA, designed fixture and the standard kit. Then we use the designed the fixture and the standard kit corresponding with DUT, and run the programme of the de-embedding algorithm implementing with C++ and Matlab. At last, we obtain the well measurement results.

REFERENCES

1. Zhu, N. H., "New algorithms of the TSM and TOM methods for calibrating microwave test fixtures," *Microwave and Optical Technology Letters*, Vol. 34, No. 1, July 5, 2002.
2. Wong, K., "The 'unknown thru' calibration advantage," *63th ARFTG Conference*, 2004.
3. Rytting, D., "TRL calibration," *Hewlett Packard.*

A Ka Band LTCC-based Small Encapsulated Transceiver Module

Ye Yuan, **Yubo Cui**, **Shengchang Zhang**, and **Kai Zhang**
The 10th Institute of Chinese Electronic Technology Corporation
P. O. Box 94-13, Chengdu 610036, China

Abstract— With the development of the millimeter-wave system technology, system on chip (SOC) and system on package (SOP) millimeter wave system become an important direction. Low-temperature co-fired ceramics (LTCC) technology has been widely used in the design because of its relatively high dielectric constant, the technics stability and the three-dimention (3D) system on package in the millimeter-wave transceiver design.

In this paper, we describes a small Ka-band transceiver using LTCC technology and its application in a small channel extension. All part of the millimeter-wave circuit are built on Ferro A6 tape in a standard multi-layer LTCC technology, and be hermetic encapsulated with tungsten copper and kovar. This transceiver includes the internal components of a receiving channel and a transimitting channel, which are switched by a MMIC switch chip, and share a fourth subharmonic image rejection mixer which achieve the downconvert frequency image rejection and upconvert sideband constrain at the same time. The circuit in this transceiver is high-density integrated, system integration package (SOP) of the millimeter-wave MMIC chips and other passive circuit. In this paper it is described in detail on the principle of transceiver, LTCC waveguide-microstrip transition, the fourth subharmonic image rejection mixer, RF bandpass filters, LO bandstop filter and other passive circuits and the hermetic packaging technology of the small transceiver module.

Transceiver has the entire volume of $32\,\text{mm} \times 45\,\text{mm} \times 6\,\text{mm}$. Working in the 400 MHz bandwidth in the receiver gain measured greater than 35 dB, noise figure of less than 5 dB, image frequency compression of greater than 30 dB, output power greater than 21 dBm. The result provides a good high-volume production consistency of circuit, low-cost, easy assembly, with a high value of the LTCC encapsulated transceiver module.

1. INTRODUCTION

With the development of radar, communications systems the millimeter wave band circuits and systems are becoming the inevitable trend. The requirement of the millimeter wave T/R module becomes higher in many engineering applications, such as need for small-scale, high-reliable gas-tight package, the production with high consistency, and so on. LTCC technology have a huge advantage in these areas, we can achieve system on chip, system on package, integrate all kinds of passive circuit in a small space, reach the goal of miniaturization, as well as a reduction of all types of assembly and testing process to improve the reliability of the system and productive. In this paper, on the Ferro's FA6 substrate, a Ka-band millimeter-wave module of the hermetic package is presented.

2. PRINCIPLE

Figure 1 shows the LTCC millimeter wave T/R module diagram. It includes receiving channel and transmitting channel. The receiving channel includes LNA and amplifier, two amplifiers and a power amplifier are in the transmitting channel. The two channels share a fourth sub-harmonic image rejection mixer. Two SPDT MMIC are used to switch the two channels. The fourth sub-harmonic image rejection mixer can realize the image rejection in downconvert and side band constrain in upconvert.

2.1. Passive Circuits

2.1.1. Microstrip-to-waveguide Transition

With the advantage of high performance and small size, probe is widely used in the traditional Microstrip-to-waveguide transition in the millimeter-wave circuits. However, LTCC ceramic substrate is a fragile material, so we need to improve the traditional transition structure. And eventually we choosed the structure with no back-shot shown in Figure 2. It is not only to solve the problem of assembling the fragile LTCC probe, but also realize the hermetic packaging of T/R module. The simulation result of the probe is shown in Figure 3.
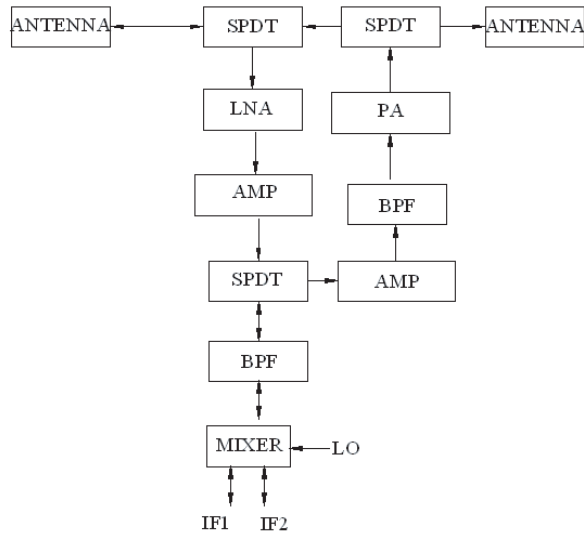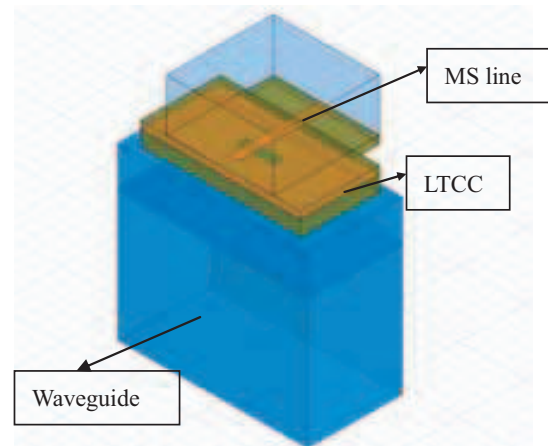
Figure 1: Principle of millimeter wave T/R module.
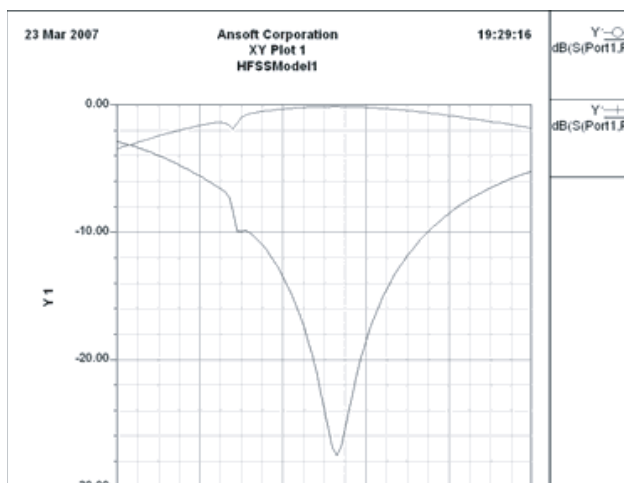

Figure 2: Microstrip-to-waveguide transition.
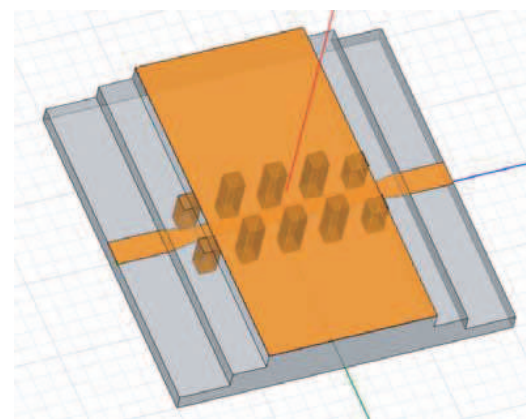

Figure 3: Simulation result.

Figure 4: Microstrip-to-stripline transition.

2.1.2. Microstrip-to-stripline Transition

Because of the high-frequency, space radiation, high integration density and high gain, in order to improve the isolation of RF signal, we used a large number of stripline. But MMIC cirduits must be bonded in the microstrip line, so we have to design the microstrip-to-stripline transition. Figure 4 shows the structure of microstrip-to-stripline transition, the simulation result is shown in Figure 5.

2.1.3. Fourth Subharmonic Image Rejection Mixer

In order to achieve the goal of small size of the transceiver, we choosed the one convertion project, which has the directe advantage of cutting the L-band signal channel and a channel of LO source of L-band frequency synthesizer, so the fourth subharmonic image rejection mixer is necessary. The advantage of fourth subharmonic image rejection mixer is below. Firstly the realization of image rejection in the downconvert mode, secondly it can achieve sideband clutter constraining in the upconvert mode. At the same time the fourth subharmonic mode can achieve the suppression of the fourth harmonic and noise of LO source, and reduce the difficulty and cost of the frequency synthesizer. Figure 6 shows the fourth subharmonic image rejection mixer diagram, a 1/4 wavelength transmission line achieve the RF 90 degrees phase-shifted, lumped inductance and capacitance realizes the IF 90 degrees phase-shifting. Figure 7 shows the simulation resultof the convertion and image rejection of the fourth subharmonic image rejection mixer. Simulation convertion loss is less than 10 dB, image rejection more than 35 dB.
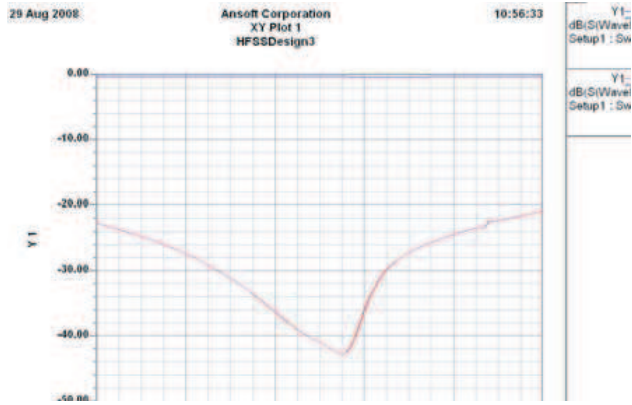
Figure 5: Simulation result in Ansoft HFSS.
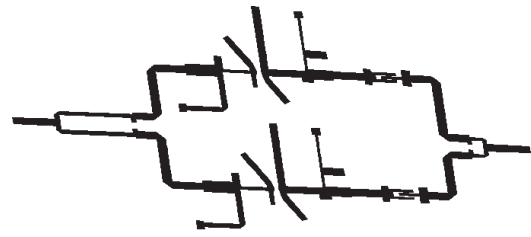
Figure 6: Fourth subharmonic image rejection mixer.
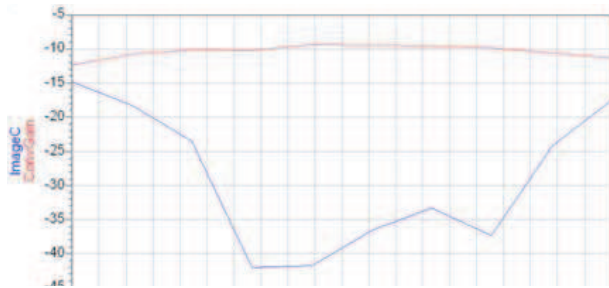
Figure 7: Simulation result of mixer.
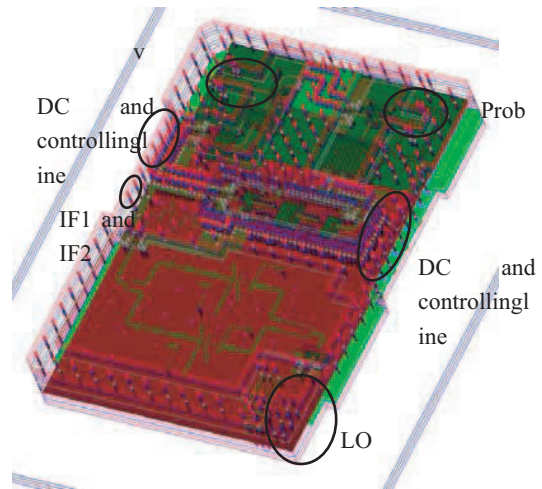
Figure 8: 3D view of LTCC module.

2.2. LTCC Module Design

There are twelve layers in the substrate of LTCC circuit with a total thickness of 1.2 mm. Figure 7 shows the three-dimensional map of LTCC circuits, which has an area of 30 mm × 35 mm. RF goes in the six layer, the top five layers are used as the top ground of the stripline, the bottom three layers are used as the bottom ground of stripline line and the last four layer are used as DC line
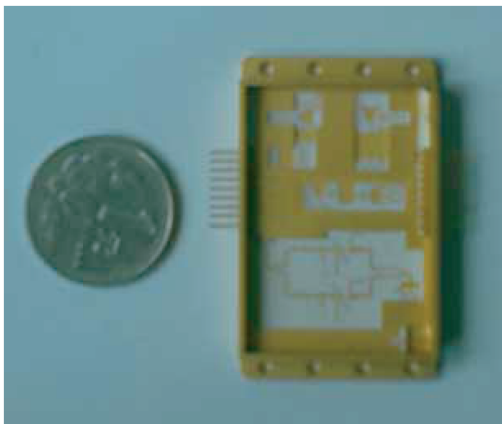
Figure 9: Photograph of LTCC module.
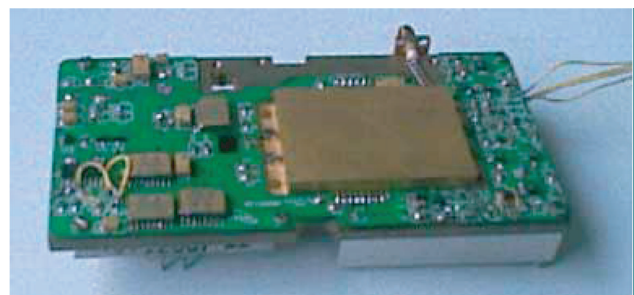
Figure 10: Photograph of transceiver.

and controlling line. Figure 8 shows the 3D view of LTCC module. LTCC substrate is welded in the W-Cu carrier and Kovar box wall around the substrate. These three materials have the similar expansion coefficient, so that it can be packaged and worked in a wide temperature band. The adapters are used to connect low-frequency signal and DC line using, and microwave adapters are used to connect high frequency signal. Photograph of LTCC module is shown in Figure 9.

3. MEASUREMENT

Figure 10 shows the photograph of transceiver, including the frequency synthesizer module, DC-DC power modules, LTCC millimeter-wave circuit module. The module control line and DC power line are fabricated on the PCB. Frequency Synthesizer in LTCC module connected on a Microwave substrate. The transceiver has the entire volume of $120\,\text{mm} \times 60\,\text{mm} \times 15\,\text{mm}$, which was measured in the $400\,\text{MHz}$ band at Ka band. Figure 11 shows the measurement of noise figure and gain and image rejection, and the measurement of transmitting power is shown in Figure 12. The measurement shows that the transceiver has the gain greater than $35\,\text{dB}$, noise figure less than $5\,\text{dB}$, image rejection greater than $30\,\text{dB}$, transmitting power greater than $21\,\text{dBm}$.
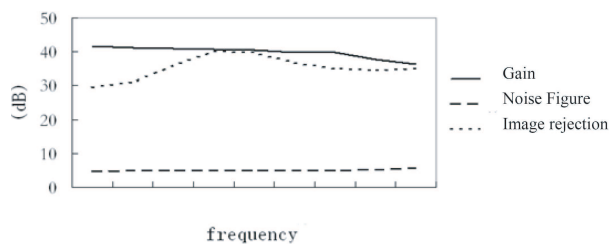
Figure 11: Measured result of transceiver.   Figure 12: Measured result of transceiver.
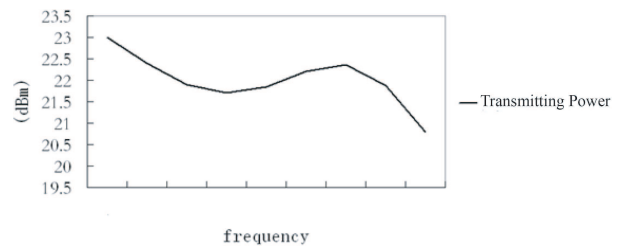
4. CONCLUSION

In this paper, a miniaturized millimeter-wave transceiver based on LTCC technology is presentated. Modules with high-density integrated circuit realized the hermetic package, based on which the transceiver is miniaturized, has the good solution of high-volume production, lower costs, etc, and it has a high value in engineering applications.

REFERENCES

1. Lee, C.-H., A. Sutono, S. Han, et al., "A compact LTCC-based Ku-band transmitter module," *IEEE Transactions on Advanced Packaging*, Vol. 25, No. 3, Aug. 2002.
2. Mancuso, Y., P. Gremillet, and P. Lacomme, "T/R-modules technological and technical trends for phased array antennas," *IEEE Transactions*, Vol. 614, 2006.
3. Heyen, J., T. von Kerssenbrock, A. Chernyakov, et al., "Novel LTCC/BGA modules for highly integrated millimeter-wave transceivers," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 51, No. 12, 2589–2596, Dec. 2003.
4. Lin, Y.-S., C.-C. Liu, K.-M. Li, et al., "Design of an LTCC Tri-band transceiver module for GPRS mobile applications," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 52, No. 12, Dec. 2004.
5. Davis, M. F., A. Sutono, K. Lim, J. Laskar, et al., "Multi-layer fully organic-based system on package (SOP) technology of RF application," *Proceeding of EPEP*, 103–106, 2000.
6. Lim, K., A. Obatoyinbo, A. Sutono, S. Chakraborty, et al., "A highly integrated transceiver module for 5.8 GHz OFDM communication system using multi-layer packaging technology," *IEEE MTT-S Int. Microwave Symp. Dig.*, 1739–1742, 2001.

HFSS$^{\text{TM}}$ Modelling Anomalies with THz Metal-Pipe Rectangular Waveguide Structures at Room Temperature

Yun Zhou and Stepan Lucyszyn

Imperial College London, UK

Abstract— Air-filled metal-pipe rectangular waveguides (MPRWGs) represent one of the most important forms of guided-wave structure for terahertz applications. Well-known commercial electromagnetic modelling software packages currently employ over-simplified intrinsic frequency dispersion models for the bulk conductivity of normal metals used in terahertz structures at room temperature. This paper has compared various conductivity modelling strategies for normal metals at room temperature and characterized rectangular waveguides and associated cavity resonators between 0.9 and 12 THz. An expression for the geometrical factor of a rectangular cavity resonator has been derived for the general case of a metal characterized with $\mu_r \neq 1$ and $\omega\tau > 0$. In addition, a method for determining the corresponding lossless frequency of oscillation has been given for the first time for such models. Using these techniques, a quantitative analysis for the application of different models used to describe the intrinsic frequency dispersion nature of bulk conductivity at room temperature has been undertaken. When compared to the use of the accurate relaxation-effect model, it has been found that HFSS$^{\text{TM}}$ (Versions 10 and 11) gives a default error in the attenuation constant for MPRWGs of 108% at 12 THz and 41% errors in both Q-factor and overall frequency detuning with a 7.3 THz cavity resonator. With the former, measured transmission losses will be significantly lower than those predicted using the current version of HFSS$^{\text{TM}}$, which may lead to an underestimate of THz losses attributed to extrinsic effects. With the latter error, in overall frequency detuning, the measured positions of return loss zeros, within a multi-pole filter, will not be accurately predicted by the current version of HFSS$^{\text{TM}}$. This paper has highlighted a significant source of errors with the electromagnetic modeling of terahertz structures, operating at room temperatures, which can be rectified by adopting the classical relaxation-effect model to describe the frequency dispersive behavior of normal metals.

1. INTRODUCTION

The past decade has seen rapid developments in the exploitation of the lower terahertz frequency band [1–5]. Applications include radiometric imaging and remote sensing; spectroscopy and radio astronomy; as well as high-speed communications. Air-filled metal-pipe rectangular waveguides (MPRWGs) represent one of the most important forms of guided-wave structure for applications at (sub-)millimetre-wave frequencies. With the development of low-cost microfabrication technologies, high tolerance structures are becoming more affordable [6, 7]. This is helping to open up the lower terahertz frequency band to new, and possibly ubiquitous, applications in the not too distant future.

In addition to the advances being made for the manufacture of terahertz MPRWG structures [1–5], AB MILLIMETRE claim that their 8 GHz to 1 THz vector network analyzer (MVNA-8-350) can be pushed to perform measurements up to *ca.* 2 THz [8]. However, some well-known commercial electromagnetic modelling software packages currently employ over-simplified frequency dispersion models for the conductivity of metals used to predict the performance of terahertz structures at room temperature. For example, Ansoft's High Frequency Structure Simulator (HFSS$^{\text{TM}}$) is considered by some to represent a benchmark standard in electromagnetic modelling software, even though it can give anomalous results under certain conditions (e.g., electrically thin-walled MPRWGs [9]).

We investigate modelling anomalies found with HFSS$^{\text{TM}}$ for the simulation of terahertz air-filled MPRWGs and associated cavity resonators. To this end, recommended standards for MPRWGs have been adopted here for operation between 0.9 and 12 THz, based on ISO 497:1973 *Preferred Metric Sizes* [10].

2. FREQUENCY DIPSERISON IN NORMAL METALS AT ROOM TEMPERATURE

For any normal metal at room temperature, the generic equations for intrinsic surface impedance Zs, propagation constant γ_s, skin depth δ_s and complex skin depth δ_c are given by:

$$Z_S \equiv R_S + jX_S = \sqrt{\frac{j\omega\mu_0\mu_r}{\sigma + j\omega\varepsilon_o}} \cong \sqrt{\frac{j\omega\mu_0\mu_r}{\sigma}} \quad \text{with} \quad \omega < 10^{15} \text{ rad/s}$$

$$\gamma_S \equiv \alpha_S + j\beta_S = \frac{j\omega\mu_0\mu_r}{Z_S} = \sqrt{j\omega\mu_o\mu_r\sigma} \quad \text{and} \quad \delta_S = \frac{1}{\Re\{\gamma_S\}} = \frac{1}{\alpha_S} \quad \text{and} \quad \delta_c \equiv \delta_c' - j\delta_c'' = \frac{1}{\gamma_S}$$

$$(1)$$

where $R_s = \Re\{Z_s\}$ is the surface resistance; $X_s = \Im\{Z_s\}$ is the surface reactance; $j = \sqrt{-1}$ is the complex operator; $\omega = 2\pi f$ is the angular frequency; f is the frequency of the driving electromagnetic field; μ_o is the permeability of free space; and μ_r is the relative permeability; σ is the intrinsic bulk conductivity of the metal; ε_o = permittivity of free space; $\alpha_S = \Re\{\gamma_S\}$ is the attenuation constant; and $\beta_S = \Im\{\gamma_S\}$ is the phase constant.

2.1. Intrinsic Frequency Dispersion Models for Normal Metals at Room Temperature

The "classical relaxation-effect model" (variables identified by the suffix R) is well understood and recognized as an analytical model that accurately describes the frequency dispersive nature of normal metals at room temperature [11, 12]. With the classical relaxation-effect (or Drude) model, intrinsic bulk conductivity is given by:

$$\sigma_R \equiv \sigma_R' - j\sigma_R'' = \frac{\sigma_o}{(1 + j\omega\tau)} \Rightarrow \begin{cases} \sigma_R' = \frac{\sigma_o}{1+(\omega\tau)^2} & \text{Simple Relaxation-Effect Model} \\ \sigma_o & \text{Classical Skin-Effect Model} \end{cases} \quad (2)$$

where, σ_o is the intrinsic bulk conductivity at DC; and τ is the phenomenological scattering relaxation time for the free electrons (i.e., mean time between collisions).

The "simple relaxation-effect model" (variables identified by the suffix R') takes only the real part of the conductivity from the classical relaxation-effect model. This model tried to remove the need for a complex conductivity, while still acknowledging that it has both a frequency and temperature dependency.

The "classical skin-effect model" (variables identified by the suffix o) removes the frequency dependency, while still keeping a temperature dependency. This model is traditionally used by RF, microwave and even millimeter-wave engineers, by default. With both the simple relaxation-effect and classical skin-effect models, equations for surface impedance, propagation constant and skin depth are greatly simplified; albeit overly-simplified at terahertz frequencies. Commercial electromagnetic modelling packages, such as HFSS$^{\text{TM}}$ (Versions 10 and 11), assume the classical skin-effect model by default, while giving the option to input frequency-dependent real values only for conductivity [13]. It should be noted here that HFSS$^{\text{TM}}$ (Versions 10 and 11) allows complex values for conductivity to be entered, but the imaginary term does not appear to be used in calculations.

The "effective relaxation-effect model" represents a logical way around the problem of not being able to incorporate the imaginary part of conductivity. It can be shown that the effective intrinsic bulk permittivity of the metal can be represented as:

$$\varepsilon_{eff} = \varepsilon_o\varepsilon_{reff} \equiv \varepsilon_o(\varepsilon_{reff}' - j\varepsilon_{reff}'') = \sigma_{eff}/j\omega = \left[\sigma' + j(\omega\varepsilon_o - \sigma'')\right]/j\omega \quad (3)$$

where ε_{reff} is the effective relative intrinsic bulk permittivity of the metal (also know as the dielectric function), and the effective dielectric constant of the metal $\varepsilon_{reff}' = \Re\{\varepsilon_{reff}\}$ is given by:

$$\varepsilon_{reff}' = 1 - \sigma''/\omega\varepsilon_o \quad (4)$$

Therefore, with the effective relaxation-effect model, only the real values of intrinsic bulk conductivity and effective relative permittivity are needed to represent the classical relaxation-effect model. Unfortunately, it will be shown that with HFSS$^{\text{TM}}$ (Versions 10 and 11) this solution does not appear to work either, as the real part of effective relative intrinsic bulk permittivity also appears to be ignored for normal metals.

2.2. Surface Impedance and Skin Depth Calculations

Using (1)–(2), the surface impedance and skin depth can be easily defined and calculated for a normal metal at room temperature:

Classical relaxation-effect model: $\sigma(\omega) \to \sigma_R(\omega)$

$$Z_{SR} \equiv R_{SR} + jX_{SR} = \sqrt{\frac{j\omega\mu_0\mu_r}{\sigma_R}} = Z_{So}\sqrt{1+ju};$$

$$\delta_{SR} = \frac{1}{\alpha_{SR}} = \Re\left\{\frac{j\omega\mu_o\mu_r}{Z_{SR}}\right\}^{-1} = \delta_{So}\sqrt{1+u^2}\frac{R_{SR}}{R_{So}} \quad \text{where} \quad u = (\omega\tau)$$

$$\delta_{cR} = \frac{1}{\gamma_{SR}} = \sqrt{\frac{1+ju}{j\omega\mu_o\mu_r\sigma_o}} = \delta_{So}\frac{\sqrt{1+ju}}{(1+j)} = \frac{\delta_{So}}{j2}\sqrt{1+ju}\,(1+j)$$

$$= \frac{\delta_{So}}{2}\left(\sqrt{\sqrt{1+u^2}+u} - j\sqrt{\sqrt{1+u^2}-u}\right)$$

(5)

where $R_{SR} = R_{So}\sqrt{\sqrt{1+(\omega\tau)^2}-\omega\tau}$ is the surface resistance and $X_{SR} = R_{So}\sqrt{\sqrt{1+(\omega\tau)^2}+\omega\tau}$ is the surface reactance.

Simple relaxation-effect model: $\sigma(\omega) \to \sigma_{R'}(\omega)$

$$Z_{SR'} \equiv R_{SR'} + jX_{SR'} = \sqrt{\frac{j\omega\mu_o\mu_r}{\sigma'_R}} = Z_{So}\sqrt{1+u^2};$$

$$\delta_{SR'} = \frac{1}{\alpha_{SR'}} = \frac{1}{\sigma'_R R_{SR'}} = \sqrt{\frac{2}{\omega\mu_o\mu_r\sigma'_R}} = \delta_{So}\sqrt{1+u^2}$$

(6)

$$\delta_{cR'} = \frac{1}{\gamma_{SR'}} = \sqrt{\frac{1+u^2}{j\omega\mu_o\mu_r\sigma_o}} = \delta_{So}\frac{\sqrt{1+u^2}}{(1+j)} = \frac{\delta_{So}}{2}\sqrt{1+u^2}(1-j) = \frac{\delta_{SR'}}{2}(1-j)$$

where $R_{SR'} = R_{So}\sqrt{1+(\omega\tau)^2}$ is the surface resistance and $X_{SR'} \equiv R_{SR'}$ is the surface reactance.

Classical skin-effect model: $\sigma(\omega) \to \sigma_o$

$$Z_{So} \equiv R_{So} + jX_{So} = \sqrt{\frac{j\omega\mu_o\mu_r}{\sigma_o}} = R_{So}(1+j);$$

$$\delta_{So} = \frac{1}{\alpha_{So}} = \frac{1}{\sigma_o R_{So}} = \sqrt{\frac{2}{\omega\mu_o\mu_r\sigma_o}}$$

(7)

$$\delta_{co} = \frac{1}{\gamma_{So}} = \sqrt{\frac{1}{j\omega\mu_o\mu_r\sigma_o}} = \frac{\delta_{So}}{(1+j)} = \frac{\delta_{So}}{2}(1-j)$$

where $R_{So} = \sqrt{\frac{\omega\mu_o\mu_r}{2\sigma_o}}$ is the surface resistance and $X_{So} \equiv R_{So}$ is the surface reactance.

For gold at room temperature, the electrical parameters may differ between published sources; depend on the exact temperature and method of deposition [11, 13, 14]. The values adopted here for bulk DC conductivity, relative permeability and relaxation time are $\sigma_o = 4.517 \times 10^7$ S/m, $\mu_r = 0.99996$ and $\tau = 27.135 fs$, respectively. In order to be consistent, these values have been used in all methods of modelling (i.e., in both calculations and HFSS$^{\text{TM}}$ simulations). Fig. 1 shows the calculated values from 0.9 to 12 THz for gold at room temperature for surface resistance, surface reactance, skin depth and the imaginary part of complex skin depth.

It can be seen from Fig. 1(a) that all three models give diverging values for surface resistance as frequency increases. Indeed, contrary to normal expectations, the rise in value for the relaxation-effect model is much less than with the other models; resulting in lower levels of predicted attenuation and higher unloaded Q-factors with the relaxation-effect model for MPRWG structures. From Fig. 1(b), it can be seen that surface reactance is very similar for both classical and simple relaxation-effect models, and these have a much greater increase with frequency when compared to the classical skin-effect model; resulting in higher levels of predicted frequency detuning error with the classical skin-effect model for cavity resonators.

With reference to Fig. 1(c), both classical relaxation- and skin-effect models give similar reductions in skin depth $\delta_S(\omega)$ with frequency. However, with the simple relaxation-effect model, there is a turning point at the relaxation frequency $f_\tau = 5.865$ THz; above which the skin depth begins
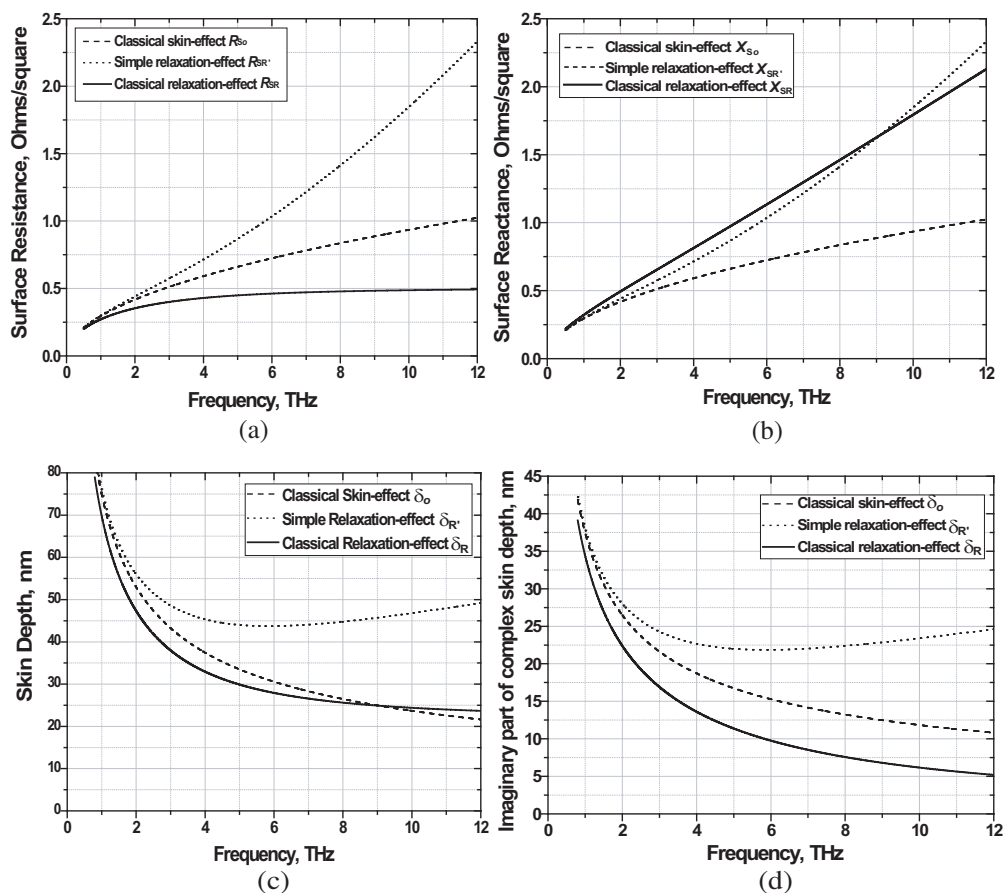
Figure 1: Calculated values for gold at room temperature: (a) surface resistance; (b) surface reactance; (c) skin depth; and (d) imaginary part of complex skin depth.

to increase with frequency — contrary to conventional expectations. It will be seen in Section 4 that, for the general case of a metal with $\omega\tau > 0$, the unloaded Q-factor is inversely proportional to the imaginary part of the complex skin depth $\Im\{\delta_c(\omega)\}$; this has been plotted in Fig. 1(d). It can be seen that while the simple relaxation- and classical skin-effect models have frequency responses that scale those of skin depth, the classical relaxation-effect model has its imaginary part of complex skin depth that decreases much more than with the other models. This will result in higher levels of unloaded Q-factor and significant errors with the simple relaxation- and classical skin-effect models for cavity resonators.

3. THZ METAL-PIPE RECTANGULAR WAVEGUIDE MODELLING

3.1. Proposed ISO 497:1973-based Standards for THz Frequencies

In Fig. 2, a uniform air-filled MPRWG is defined within the Cartesian coordinate system xyz. Transmission is along the z direction, and over a distance d, with internal cross-sectional dimensions a and b.
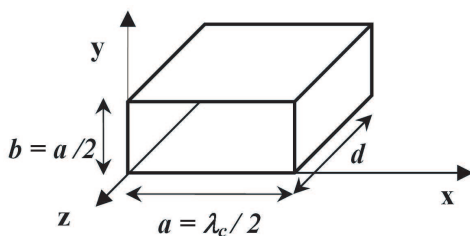
Figure 2: Internal spatial variable definitions for a uniform air-filled MPRWG.

The structure in Fig. 2 has the ideal (i.e., lossless) dominant-mode guided wavelength given by the following textbook expression [15]:

$$\lambda_{g\_ideal} = \frac{\lambda_o}{\sqrt{1 - \left(\frac{\lambda_o}{\lambda_c}\right)^2}} = \frac{\lambda_o}{\sqrt{1 - \left(\frac{f_c}{f}\right)^2}} \tag{8}$$

where, λ_o is the free space wavelength; $\lambda_c = 2a$ is the ideal cut-off wavelength; a is the internal width dimension of the MPRWG; $f_c = c/2a$ is the ideal cut-off frequency for the dominant TE_{10} mode; and c is the speed of light in free space.

For many decades, MPRWGs have been standardized by different designations (e.g., IEC-R, EIA-WR and WG). Currently, commercialization of such waveguides exists up to 500 GHz (i.e., WR-2.2) by Millitech$^\circledR$ [15] and up to 1.157 THz (i.e., WR-0.51) by Virginia Diodes Inc. [16]. However, there are still no global standards for frequency bands at terahertz frequencies. Cross-sectional dimension variations between well-established standards create serious problems, which are exacerbated by inadequate mechanical tolerances, as frequencies approach the terahertz spectrum. In order to provide compatible (interchangeable) hardware at terahertz frequencies, the width dimension can be generated using the ISO 497:1973 *Preferred Metric Sizes* standard [10]; as recently proposed for operational frequencies between 0.9 and 2.9 THz [17, 18]. The ISO 497:1973 standard is already adopted widely by industry for other applications, is infinitely extendable and finer choices can accommodate smaller dimensions. Based on this standard, as shown in Table 1, we have extended the four lower terahertz frequency bands [17, 18] up to 12 THz.

Table 1: Proposed air-filled MPRWG and cavity resonator definitions and specifications.

ISO 497 Preferred Metric Size		Proposed Frequency Band Designation (μm)		Internal Dimensions $a \times b$ (μm²)	TE_{10}-mode Ideal Cutoff Frequency f_C (THz)	Lower-Band Frequency Factor f_L/f_c	Mid-Band Frequency $f_O = 1.55 f_C$ (THz)	Upper-Band Frequency Factor f_U/f_c	Useful Frequency Range $f_L \to f_U$ (THz)	TE_{101}-mode Ideal Cavity Resonance Frequency $f_{101\_ideal} = \sqrt{1.5}f_c$ (THz)
B	i	JPL	Our							
20'	6	200		200×100	0.75	1.20	1.162	1.93	0.90→1.45	0.919
20'	4	160		160×80	0.94	1.23	1.452	1.92	1.15→1.80	1.148
20'	2	125		125×62.5	1.20	1.21	1.860	1.92	1.45→2.30	1.470
20'	0	100		100×50	1.50	1.20	2.325	1.93	1.80→2.90	1.837
40	35		75	75×32.5	2.00	1.20	3.100	1.90	2.40→4.00	2.449
20'	14		50	50×25	3.00	1.20	4.650	1.90	3.60→6.00	3.674
20'	10		32	32×16	4.68	1.20	7.260	1.90	5.62→8.90	5.732
20'	8		25	25×12.5	6.00	1.20	9.300	1.90	7.20→12.00	7.348

3.2. Calculation of Attenuation Constant

The propagation constant for a lossy metal-pipe rectangular waveguide can be accurately calculated using the *variational* method [19] for the TE_{mo} mode [7], as follows:

$$\gamma_{mo}^2 = \Gamma_d^2 - j\frac{2Z_s}{\omega\mu_o\mu_r b}\left[\left(\frac{\Gamma_d}{k_c}\frac{m\pi}{a}\right)^2 - k_c^2\left(1 + \frac{2b}{a}\right)\right]; \quad \Gamma_d^2 = k_c^2 - k_{od}^2$$

$$k_c = \omega_c\sqrt{\mu_o\mu_r\varepsilon_o\varepsilon_r} \to \omega_c\sqrt{\mu_o\varepsilon_o} = \frac{\omega_c}{c} \quad \text{in free space} \tag{9}$$

$$k_{od} = \omega\sqrt{\mu_o\mu_r\varepsilon_o\varepsilon_r(1 - j\tan\delta)} \to k_o = \omega\sqrt{\mu_o\varepsilon_o} = \frac{\omega}{c} \quad \text{in free space}$$

where ε_r is the dielectric constant and $\tan\delta$ is the loss tangent for the dielectric filler.

The attenuation constant for this guided-wave structure can be obtained directly from the real part of the propagation constant, i.e., $\alpha = \Re\{\gamma_{mo}\}$. For the dominant TE_{10} mode, the following expressions for attenuation constant can be calculated, based on the intrinsic frequency dispersion models for normal metals at room temperature.

<u>Classical relaxation-effect model</u>

$$\alpha_R = \Re\{\gamma_{10R}\}; \quad \gamma_{10R} = f\{Z_{SR}\} \tag{10a}$$

<u>Simple relaxation-effect model</u>

$$\alpha_{R'} = \Re\{\gamma_{10R'}\}; \quad \gamma_{10R'} = f\{Z_{SR'}\} \tag{10b}$$

<u>Classical skin-effect model</u>

$$\alpha_o = \Re\{\gamma_{10o}\}; \quad \gamma_{10o} = f\{Z_{So}\} \tag{10c}$$

Using (10) with the simple Power Loss approximation method, as used in [6], for calculating attenuation constant in terms of the real part of the surface impedance only, it is easy to calculate the percentage error in attenuation constant for the simple relaxation-effect model $E_{\alpha_{R'}}$ and classical skin-effect model E_{α_o}, relative to the classical relaxation-effect model, as follows:

$$
\begin{aligned}
E_{\alpha_{R'}} &= \left(\frac{\alpha_{R'} - \alpha_R}{\alpha_R}\right) \cdot 100\% \cong \left[\sqrt{\frac{1 + (\omega\tau)^2}{\sqrt{1 + (\omega\tau)^2} - \omega\tau}} - 1\right] \cdot 100\% \\[2mm]
E_{\alpha_o} &= \left(\frac{\alpha_o - \alpha_R}{\alpha_R}\right) \cdot 100\% \cong \left[\sqrt{\sqrt{1 + (\omega\tau)^2} + \omega\tau} - 1\right] \cdot 100\%
\end{aligned}
\tag{11}
$$

For simplicity, it will be assumed throughout that the MPRWGs will have a height dimension of $b = a/2$. The attenuation constants and resulting errors have been plotted against frequency, and are shown in Fig. 3. To a first degree of approximation, it can be seen that the error increases linearly with frequency for the classical skin-effect model; with a 108% error at 12 THz. The error obtained with the simple relaxation-effect model is 373% at 12 THz.
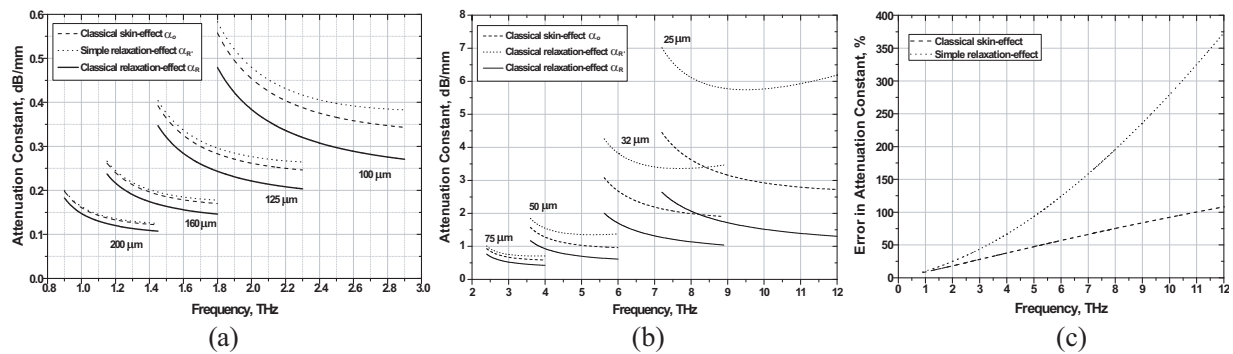
Figure 3: Calculated attenuation constants for the dominant TE_{10} mode: (a) JPL bands; (b) our bands; and (c) resulting errors in attenuation constants.

3.3. Electromagnetic Simulation of Attenuation Constant

$HFSS^{TM}$ is often used to design terahertz guided-wave structures [1–5]. Therefore, to test its suitability at these frequencies, a number of simple, uniform THz air-filled MPRWGs were simulated. The $HFSS^{TM}$ results are shown in Fig. 4(a), while the calculated values are superimposed, for comparison, in Fig. 4(b). It can be clearly seen in Fig. 4 that all the relaxation-effect model values entered into $HFSS^{TM}$ give identical results to the calculated simple relaxation-effect model, while $HFSS^{TM}$ (by default) gives results that conform to the calculated classical skin-effect model. This clearly shows that $HFSS^{TM}$ (Versions 10 and 11) cannot give accurate results and that very large errors will result, corresponding to those given in Fig. 3(c).
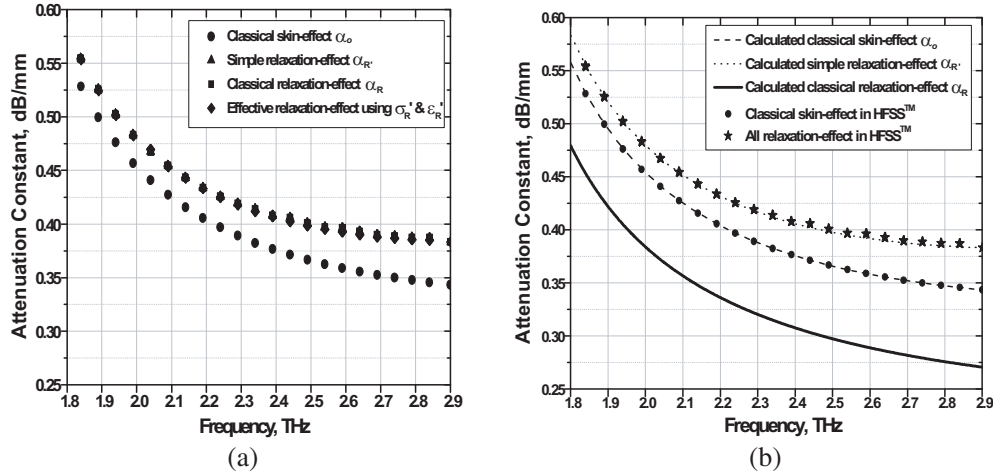
Figure 4: HFSS^{TM} simulated attenuation constant for the dominant TE$_{10}$ mode in the 100 μm JPL band: (a) HFSS^{TM} only; and (b) HFSS^{TM} and calculated values.

4. THZ CAVITY RESONATOR MODELLING

4.1. Derivation of Basic Equations

With reference to Fig. 2, the internal dimensions of a cavity resonator are a, b and d. The corresponding textbook expression for the resonant frequencies for the TE$_{mnl}$ modes in an ideal (i.e., lossless) cavity is give by:

$$f_{mnl\_ideal} = \frac{c}{2\pi}\sqrt{\left(\frac{m\pi}{a}\right)^2 + \left(\frac{n\pi}{b}\right)^2 + \left(\frac{l\pi}{d}\right)^2} \tag{12}$$

For simplicity, it will be assumed throughout that the cavity will have spatial dimensions of $b = a/2$ and $d = \sqrt{2}a$. Moreover, only the dominant TE$_{101}$ mode will be considered:

$$\therefore f_{101\_ideal} = \sqrt{1.5}f_c = \sqrt{\frac{3}{8}}\left(\frac{c}{a}\right) \equiv \frac{\omega_I}{2\pi} \tag{13}$$

In HFSS^{TM}, the eigenmode solver determines the complex resonant frequency of a structure $\tilde{\omega}_o$ [20]:

$$\tilde{\omega}_o \equiv \omega_o' + j\omega_o'' = \omega_o\sqrt{1 - \left(\frac{1}{2Q_u(\omega_o)}\right)^2} + j\frac{\omega_o}{2Q_u(\omega_o)}$$

$$\omega_o = |\tilde{\omega}_o| = \sqrt{(\omega_o')^2 + (\omega_o'')^2} \quad \text{and} \quad Q_u(\omega_o) = \frac{\omega_o}{2\omega_o''} \tag{14}$$

With a non-zero surface reactance, the equivalent surface inductance effectively reduces the lossless frequency of oscillation ω_o from the ideal case ω_I by a small perturbation, $\Delta\omega_I = (\omega_o - \omega_I)$. Moreover, with most applications, the unloaded Q-factor $Q_u(\omega_o)$ is large enough so that the corresponding frequency detuning caused by ohmic losses (e.g., surface resistance) is negligible, such that $\omega_o' \cong \omega_o$. It has been shown by Slater [20] that, to the first order of small quantities:

$$\tilde{\omega}_o = \omega_o + j\frac{\omega_I}{2Q_u(\omega_o)} \quad \text{where} \quad \omega_I = \pi\sqrt{\frac{3}{2}}\left(\frac{c}{a}\right) \tag{15}$$

$$\text{when} \quad \sigma_o \to \infty: \quad \omega_o \to \omega_I \quad \text{and} \quad Q_U(\omega_I) \to \infty: \quad \Delta\tilde{\omega}_o = (\omega_o - \omega_I) + j\frac{\omega_I}{2Q_u(\omega_o)} \tag{16}$$

Now, the cavity perturbation formula can be used to relate the change in complex resonant frequency to the corresponding change in surface impedance, as follows [20]:

$$\Delta Z_S = -j2\Gamma \cdot \Delta\tilde{\omega}_o \tag{17}$$

It can be shown that, for the general case, the geometrical factor Γ is given by:

$$\Gamma = \mu_0 \left\{ \frac{abd\left(a^2 + d^2\right)}{2\left[2b\left(a^3 + d^3\right) + ad\left(a^2 + d^2\right)\right]} \right\} = \mu_0 \left(\frac{3a}{2(\sqrt{2} + 10)} \right) \tag{18}$$

But, when $\sigma_o \to \infty$: $\quad Z_S(\omega_I) \to 0 \quad \Delta Z_S = Z_S(\omega_o)$:

$$\therefore Z_S\left(\omega_o\right) = \omega_I \Gamma \left(\frac{1}{Q_u\left(\omega_o\right)} + j\frac{2\left(\omega_I - \omega_o\right)}{\omega_I} \right)$$

$$\text{where} \quad R_S\left(\omega_o\right) = \frac{\omega_I \Gamma}{Q_u\left(\omega_o\right)} \quad \text{and} \quad X_S\left(\omega_o\right) = 2\Gamma\left(\omega_I - \omega_o\right) \tag{19}$$

We have found that for the general case of a metal with $\mu_r \neq 1$ and $\omega\tau > 0$, in order to achieve a self-consistent solution, the unloaded Q-factor can be expressed as follows:

$$Q_u\left(\omega_o\right) = \frac{\lambda_I(\omega_I)}{8\mu_r \Im\left\{\delta_c\left(\omega_o\right)\right\}} \left(\frac{\omega_I}{\omega_o} \right) \left\{ \frac{2b\left(a^2 + d^2\right)^{\frac{3}{2}}}{\left[2b\left(a^3 + d^3\right) + ad\left(a^2 + d^2\right)\right]} \right\}$$

$$= \frac{1}{\mu_r \Im\left\{\delta_c(\omega_o)\right\}} \left(\frac{\omega_I}{\omega_o} \right) \left(\frac{3\sqrt{2}a}{4\left(5\sqrt{2} + 1\right)} \right)$$

$$\text{where} \quad \delta_c(\omega_o) \equiv \delta_c'(\omega_o) - j\delta_c''(\omega_o) = \frac{1}{\gamma_S(\omega_o)} \tag{20}$$

Now, the frequency of oscillation ω_o' is defined in (14). The overall frequency detuning, which takes both perturbation and frequency detuning due to ohmic losses into account, is given by $\Delta\omega_o' = (\omega_o' - \omega_I)$. The corresponding errors in overall frequency detuning, relative to the classical relaxation-effect model, are given by:

$$E_\omega = \left| \frac{\Delta\omega_o' - \Delta\omega_{oR}'}{\Delta\omega_{oR}'} \right| \times 100\% = \left| \frac{\omega_o' - \omega_{oR}'}{\omega_{oR}' - \omega_I} \right| \times 100\%$$

$$\text{where} \quad \Delta\omega_{oR}' = \left(\omega_{oR}' - \omega_I\right) \quad \text{and} \quad \Delta\omega_o' \to \Delta\omega_{oo}' = \left(\omega_{oo}' - \omega_I\right) \quad \text{or} \quad \Delta\omega_{oR'}' = \left(\omega_{oR'}' - \omega_I\right) \tag{21}$$

At the frequency of oscillation:

$$\therefore Q_u\left(\omega_o'\right) = \frac{\omega_I \Gamma}{R_S\left(\omega_o'\right)} = \frac{1}{\mu_r \Im\left\{\delta_c\left(\omega_o'\right)\right\}} \left(\frac{\omega_I}{\omega_o'} \right) \left(\frac{3\sqrt{2}a}{4\left(5\sqrt{2} + 1\right)} \right)$$

$$\text{where} \quad \delta_c(\omega_o') \equiv \delta_c'(\omega_o') - j\delta_c''(\omega_o') = \frac{1}{\gamma_S(\omega_o')} \tag{22}$$

The corresponding errors in unloaded Q-factor, relative to the classical relaxation-effect model, are given by:

$$E_Q = \left| \frac{Q_U(\omega_o') - Q_{UR}(\omega_{oR}')}{Q_{UR}(\omega_{oR}')} \right| \times 100\% = \left| \left(\frac{R_{SR}(\omega_{oR}')}{R_S(\omega_o')} \right) - 1 \right| \times 100\%$$

$$\text{where} \quad Q_U(\omega_o') \to Q_{Uo}(\omega_{oo}') \quad \text{or} \quad Q_{UR'}(\omega_{oR'}') \quad \text{and} \quad R_S(\omega_o') \to R_{So}(\omega_{oo}') \quad \text{or} \quad R_{SR'}(\omega_{oR'}') \tag{23}$$

4.2. Derivation of Lossless Frequency of Oscillation

In order to determine the level of unloaded Q-factor and overall frequency detuning, both at the frequency of oscillation, it is first necessary to determine the lossless frequency of oscillation and then the unloaded Q-factor at this frequency, so that the frequency of oscillation can be found using (14). To this end, by equalizing equations for surface reactance, the lossless frequency of oscillation can be found by solving the roots of the characteristic equation given for each of the intrinsic frequency dispersion model.

The classical relaxation-effect model: $\omega_o \to \omega_{oR}$

$$X_{SR}(\omega_{oR}) = R_{So}(\omega_{oR})\sqrt{\sqrt{1 + (\omega_{oR}\tau)^2} + \omega_{oR}\tau} \equiv 2\Gamma\left(\omega_I - \omega_{oR}\right)$$

$$\therefore \frac{\sqrt{\omega_{oR}\left(\sqrt{1 + (\omega_{oR}\tau)^2} + \omega_{oR}\tau\right)}}{\left(\omega_I - \omega_{oR}\right)} - K \equiv 0 \quad \text{where} \quad K = \sqrt{\frac{8\mu_o\sigma_o}{\mu_r}} \cdot \left(\frac{3a}{2(\sqrt{2} + 10)} \right) \tag{24}$$

For the simple relaxation-effect model: $\omega_o \rightarrow \omega_{oR'}$

$$X_{SR'}(\omega_{oR'}) = R_{So}(\omega_{oR'})\sqrt{1 + (\omega_{oR'}\tau)^2} \equiv 2\Gamma\left(\omega_I - \omega_{oR'}\right)$$

$$\therefore \cdot \frac{\sqrt{\omega_{oR'}\left(1 + (\omega_{oR'}\tau)^2\right)}}{(\omega_I - \omega_{oR'})} - K \equiv 0 \tag{25}$$

For the classical skin-effect model: $\omega_o \rightarrow \omega_{oo}$

$$X_{So}(\omega_{oo}) = \sqrt{\frac{\omega_{oo}\mu_o\mu_r}{2\sigma_o}} \equiv 2\Gamma\left(\omega_I - \omega_{oo}\right)$$

$$\therefore \frac{\sqrt{\omega_{oo}}}{(\omega_I - \omega_{oo})} - K \equiv 0$$

$$\omega_{oo} = (W + \omega_I) - \sqrt{(W + \omega_I)^2 - \omega_I^2} \quad \text{where} \quad W = \frac{1}{2K^2} \tag{26}$$

4.3. Calculation and Electromagnetic Simulation of Unloaded Q-factor and Frequency Detuning

The calculated and HFSS$^{\text{TM}}$ simulated values for unloaded Q-factor and overall frequency detuning are plotted in Fig. 5(a), for rectangular waveguide cavity resonators operating in the dominant TE$_{101}$ mode. It can be seen that the results from HFSS$^{\text{TM}}$ correspond exactly to those calculated using the classical skin-effect model. As previously explained in Section 2.2, the classical relaxation-effect model will predict higher unloaded Q-factor and overall frequency detuning, when compared to the other models.

The resulting errors in unloaded Q-factor and overall frequency detuning, relative to the classical relaxation-effect model, show an almost identical frequency response for both with the classical skin-effect model. Here, a 41% error is calculated for a 7.3 THz cavity resonator. This level of error in overall frequency detuning can have a dramatic impact on the positions of return loss zeros, within a multi-pole filter. A much lower error is found in the overall frequency detuning with the simple relaxation-effect model; with a worst-case value of 12% for a 3.7 THz cavity resonator. However, a 63% error in the unloaded Q-factor has been calculated with the simple relaxation-effect model for a 7.3 THz cavity resonator.
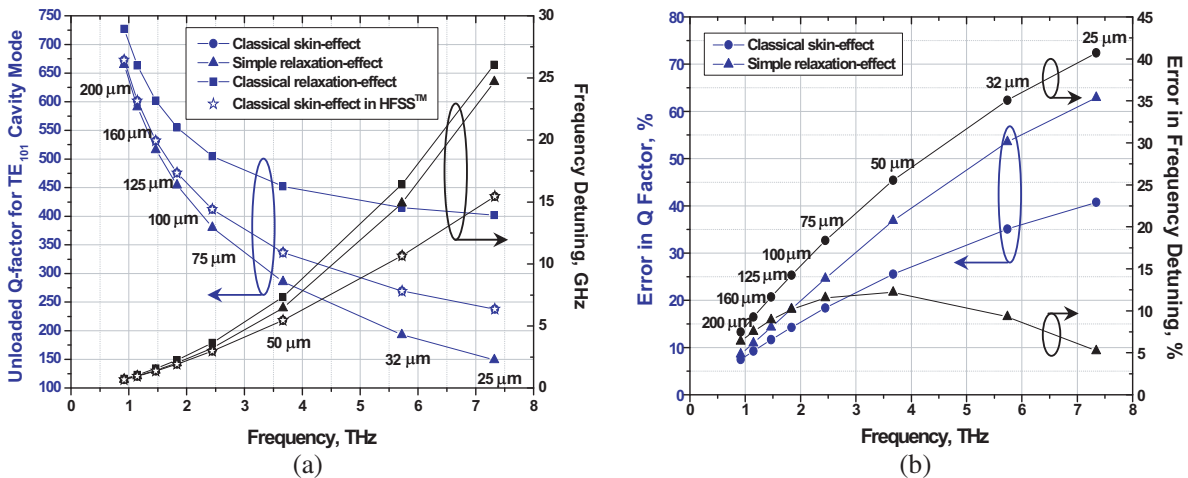
Figure 5: (a) Unloaded Q-factor and overall frequency detuning for TE$_{101}$ cavity mode, at the resonant frequencies; and (b) resulting errors in Q factors and frequency detuning.

5. CONCLUSIONS

This paper has compared various conductivity modelling strategies for normal metals at room temperature and characterized rectangular waveguides and associated cavity resonators between 0.9 and 12 THz. It has been found that the current versions of HFSS$^{\text{TM}}$ (Versions 10 and 11) cannot

accurately predict the performance of structures operating at terahertz frequencies. For example, with simple uniform MPRWGs, measured transmission losses can be significantly lower, which may lead to an underestimate of THz losses attributed to extrinsic effects; such as poor mechanical tolerances and surface roughness. Also, from the errors found in the frequency detuning of simple rectangular cavity resonators, the measured positions of return loss zeros, within a multi-pole filter, will not be accurately predicted.

It should be noted that while this paper has focused on the current version of HFSS^{TM}, the same problem can be found with other commercial electromagnetic modelling software packages that use the overly-simplified classical skin-effect model by default. This paper has highlighted a significant source of errors with the electromagnetic modeling of terahertz structures, operating at room temperatures, which can be rectified by adopting the classical relaxation-effect model to describe the frequency dispersive behavior of normal metals.

ACKNOWLEDGMENT

REFERENCES

1. Kazemi, H., S. T. Wootton, N. J. Cronin, S. R. Davies, R. E. Miles, R. D. Pollard, J. M. Chamberlain, D. P. Steenson and J. W. Bowen, "Active micromachined integrated terahertz circuits," *International Journal of Infrared and Millimeter Waves*, Vol. 20, No. 5, 967–974, Mar. 1999.

2. Maiwald, F., S. Martin, J. Bruston, A. Maestrini, T. Crawford and P. H. Siegel, "2.7 THz waveguide tripler using monolithic membrane diodes," *IEEE MTT-S International Microwave Symposium Digest*, Vol. 3, 1637–1640, May 2001.

3. Pavolotsky, A., D. Meledin, C. Risacher, M. Pantaleev, and V. Belitsky, "Micromachining approach in fabricating of THz waveguide components," *Microelectronics J.*, Vol. 36, 683–686, Jun. 2005.

4. Bowen, J. W., S. Hadjiloucas, B. M. Towlson, L. S. Karayzas, S. T. G. Wootton, N. J. Cronin, S. R. Davies, C. E. McIntosh, J. M. Chamberlain, R. E. Miles, and R. D. Pollard, "Micromachined waveguide antennas for 1.6 THz," *IET Electronics Letters*, Vol. 42, No. 15, 842–843, Jul. 2006.

5. Nordquist, C. D., M. C. Wanke, A. M. Rowen, C. L. Arrington, M. Lee, and A. D. Grine, "Design, fabrication, and characterization of metal micromachined rectangular waveguides at 3 THz," *Antennas and Propagation Society International Symposium (AP-S 2008)*, 1–4, Jul. 2008.

6. Lucyszyn, S., Q. H. Wang, and I. D. Robertson, "0.1 THz rectangular waveguide on GaAs semi-insulating substrate," *IEE Electronics Letters*, Vol. 31, No. 9, 721–722, Apr. 1995.

7. Lucyszyn, S., D. Budimir, Q. H. Wang, and I. D. Robertson, "Design of compact monolithic dielectric-filled metal-pipe rectangular waveguides for millimetre-wave applications," *IEE Proceedings — Microwaves, Antennas and Propagation*, Vol. 143, No. 5, 451–453, Oct. 1996.

8. http://www.abmillimetre.com/Products.htm.

9. Choi, J. Y. and S. Lucyszyn, "HFSS^{TM} modelling anomalies with electrically thin-walled metal-pipe rectangular waveguide simulations," *10th IEEE High Frequency Postgraduate Student Colloquium (10th HF-PgC) Digest*, 95–98, ISBN: 0-7803-9500-X, Leeds, Sep. 2005.

10. ISO 497:1973, "Guide to the choice of series of preferred numbers and of series containing more rounded values of preferred numbers," 1973. http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=4548

11. Lucyszyn, S., "Investigation of anomalous room temperature conduction losses in normal metals at terahertz frequencies," *IEE Proceedings — Microwaves, Antennas and Propagation*, Vol. 151, No. 4, 321–329, Aug. 2004.

12. Lucyszyn, S., "Investigation of Wang's model for room temperature conduction losses in normal metals at terahertz frequencies," *IEEE Trans. on Microwave Theory Tech.*, Vol. 53, 1398–1403, Apr. 2005.

13. HFSS^{TM} help files for versions 10 & 11.

14. Lide, D. R. (Editor-in-chief), *CRC Handbook of Chemistry and Physics*, CRC Press, 2003.

15. Millitech® Inc., "Rectangular waveguide specifications and MIL-specification cross reference". http://www.millitech.com/pdfs/recspec.pdf

16. Virginia Diodes Inc., "Waveguide band designations". http://www.vadiodes.com/V-DI/pdf/VDI%20Waveguide%20Chart.pdf

17. Ward, J. S., "New standards for submillimeter wavelengths," *17th International Symposium on Space Terahertz Technology*, Paris, May 2006.

18. Ward, J. S., "Terahertz waveguide standards," *IEEE MTT-S International Microwave Symposium Workshop*, WFD 06-2, San Francisco, USA, 2006.

19. Collin, R. E., *Field Theory of Guide Waves*, 182–195, McGraw-Hill, 1960.

20. Slater, J. C., "Microwave electronics," *Rev. of Modern Physics*, Vol. 18, No. 4, 441–512, Oct. 1946.

High-speed I/O Buffer Modeling for Signal-integrity-based Design of VLSI Interconnects

Yi Cao and Qi-Jun Zhang

Department of Electronics, Carleton University, Ottawa, ON, K1S 5B6, Canada

Abstract— Digital I/O buffers play an important role for the signal integrity (SI) simulation and timing analysis of high-speed VLSI interconnect networks, which often require the consideration of electromagnetic (EM) effects. In this paper, we give an overview of the recent advances in efficient macromodeling of nonlinear digital I/O buffers, including equivalent-circuit-based and neural-network-based approaches. The detailed equivalent circuit models are accurate but computationally slow. On the other hand, the simplified equivalent circuit models are fast but only provide limited accuracy. The neural-network-based models are good alternatives to those equivalent-circuit-based models, maintaining a good overall performance in terms of accuracy and speed. We demonstrate the neural-network-based approaches through an example of modeling a commercial high-speed integrated circuit (IC) device and its application to the SI simulation of high-speed interconnect networks.

1. INTRODUCTION

With the ever-increasing speed and complexity in today's high-speed VLSI system, electromagnetic (EM) effects have to be taken into account to achieve first-pass success of the design. In such a case, minimizing signal integrity (SI) effects such as signal delays, crosstalks, and ground-bounce noises has become a necessary yet challenging task [1]. Efficient and accurate representation of the SI effects is a key step towards ensuring the reliable signal propagations in the high-speed signal path, including transmission lines and nonlinear digital I/O buffers. The nonlinear I/O buffer, i.e., nonlinear drivers and receivers, are complex nonlinear circuits that are used to drive and load the interconnect structures. Digital I/O buffers play an important role for the SI simulation and timing analysis of large-scale digital system. Recently, behavioral modeling of nonlinear I/O buffers has become an important topic [2–8]. Behavioral models are simplified representations of the essential nonlinear behaviors of the detailed transistor-level descriptions of the buffer. When used in the system level SI analysis, these behavioral models can lead to faster CPU time and less memory consumption, compared to the detailed transistor-level models.

In this paper, we provide an overview of the approaches for modeling nonlinear I/O buffers. Firstly, the equivalent-circuit-based modeling approaches will be introduced. A widely used approach in this category is the I/O buffer information specification (IBIS) method, which is supported by many commercial computer-aided design (CAD) tools. However, because it mainly captures the static behavior of the buffer, the model accuracy may be limited when the buffer exhibits strong dynamic effects. The second type of methods to be reviewed is based on the use of artificial neural networks (ANN). We will review several recent approaches to ANN-based modeling of nonlinear buffers, which make use of different forms of neural network structures, such as radio basis function (RBF) neural networks [3–5], recurrent neural network [6] and state-space dynamic neural networks (SSDNN) [7, 8]. A numerical example is included to demonstrate the accuracy and robustness of the ANN-based approaches for nonlinear buffer modeling. Compared to the IBIS model, the ANN-based model is shown to produce more accurate results when the model is simulated together with different transmission lines.

2. EQUIVALENT-CIRCUIT-BASED MODEL

High-speed I/O buffers can be most accurately represented by SPICE transistor-level models [9], which are often directly extracted from the detailed circuit design of the buffer. The detailed transistor-level model can accurately account for the second-order effects (e.g., ground bounce noises) and is suitable for small-scale simulations of interconnects with a few buffers. However, because transistor-level models contain detailed process and design information, they tend to be slow and are not well suited for the simulation of the complex board consisting of a large number of nets and drivers. Another drawback of the transistor-level model is that it reveals the sensitive information of the design and manufacturing process. Due to this reason, it is often not easy to obtain the transistor-level models from the manufactures.

To improve model efficiency, simplified equivalent-circuit-based approach has been introduced to I/O buffer modeling. The most widely applied model of this type is the IBIS model [2], which was originally developed in late 1990s by the semiconductor industry to provide input-output behavioral information of integrated circuits (IC), without having to revealing the proprietary process and design information. IBIS model adopts a standard parsed file format, i.e., a table-based format, to store only the essential behavioral information about input-output characteristics of ICs. Due to this standard format, IBIS model file does not contain any confidential information about the original circuit to be modeled. Another advantage is that the IBIS model tends to run much faster than the detailed transistor-level model since it is a behavioral-level representation of the original device. Furthermore, this standard is widely supported by most electrical design automation (EDA) vendors and semiconductor manufacturers, leading to a large set of model libraries of commercial devices.

IBIS models can be generated either by measurement, which requires well-controlled measurement devices, or through multiple SPICE simulations to obtain the necessary voltage/current (V/I) tables and voltage transition tables [2]. Once all the necessary V/I data and other parameters are obtained, one can create the IBIS ASCII file using the file format defined in the IBIS standard. Next, the model file is checked by using the "Golden Parser" to detect any possible syntax errors. If the model passes this check, the model is then incorporated into an EDA tool and validated for accuracy by using the measured or simulated data from the actual device. If the model has good correlation with the actual device, the model is ready to be used for the system-level SI simulations.

3. ANN-BASED MODEL

In recent years, ANN-based techniques have emerged as an attractive alternative to the equivalent-circuit-based approach, such as IBIS model. ANNs are information-processing systems that are inspired by the abilities of the human brains to learn from observation and generalize by abstraction [10]. ANNs can be trained by an optimization process called training to learn the input-output behaviors of the nonlinear I/O buffer. The trained ANN models can then be used in SI-based analysis and design for quick and accurate estimation of the performance of the high-speed digital systems.

In [3], a parametric model based on the theory from input-output system identifications has been introduced. This method is a useful complement of the conventional IBIS models, taking into account more physical effects of the I/O drivers and maintaining higher-level of accuracy and efficiency. In this method, the nonlinear relationship between the currents and voltages of the driver output port is represented by a piece-wise linear formulation composed of two submodels that describe the port nonlinear behavior of the driver when the drive inputs are at HIGH and LOW logic states, respectively. Each submodel is represented by a weighted sum of Gaussian RBF functions. The vector of inputs to the RBF functions is called the repressor vector that collects the port voltage and current at the past several time instants and the port voltage at the present time. Each Gaussian basis function is defined by its position in the regressor space (centers) and by its spreading in terms of a scale parameter. The internal weight parameters of the RBF functions, i.e., centers and scale parameters, are estimated by standard methods such that the mode responses can fit those of the reference model for the original circuit under consideration. In the overall piece-wise model, two time- varying weight coefficients that act as the switches for submodels are utilized to combine the submodel responses depending on the transitions of the logic state at the input port. The overall model can be generated from the input-output transient waveforms at the device ports following a systematic procedure. A detailed description of the modeling estimation process can be found in [3]. This technique has been used for parametric modeling of single-ended CMOS buffers [3] as well as the low voltage differential signaling buffers [4]. An extended formulation of this approach for modeling the effects of device temperatures and non-ideal power supplies can be found in [5].

As a related technique to the piece-wise formulation with RBF functions, recurrent neural networks (RNN) are exploited in [6], for modeling the highly nonlinear driver circuits. RNNs are a generic type of discrete-time neural network that can be used to effectively model the nonlinear patterns with memory or feedback. In this case, the nonlinear submodels in the piece-wise linear formulation are represented by RNN functions which contains 3-layer multilayer perceptron (MLP) neural networks [10]. The training of the RNN is carried out by using back propagation through time (BPTT) algorithm to estimate the internal weight parameters of the RNN. By taking into account the feedback effects of RNN outputs during training, the BPTT algorithm can lead to more

robust RNN models for representing the highly nonlinear buffer behavior. In [6], it is demonstrated that the RNN model maintains the similar accuracy as the detailed transistor-level model, but being 6–7 times faster in terms of computational time.

Recently, a direct neural-network-based technique for modeling nonlinear transient input-output behavior is presented [7, 8]. This technique combines the state-space concept and the continuous RNN format to establish a generalization time-domain formulation, called SSDNN. Let $\boldsymbol{u} \in \Re^M$ be the input signals of a nonlinear circuit, e.g., input voltages and currents, and $\boldsymbol{y} \in \Re^K$ be the output signals of a nonlinear circuit, e.g., output voltages and currents where M and K are the numbers of circuit inputs and outputs respectively. The SSDNN format can be described as

$$\begin{cases} \dot{\boldsymbol{x}}(t) = -\boldsymbol{x}(t) + \tau \boldsymbol{g}_{\text{ANN}}(\boldsymbol{u}(t),\, \boldsymbol{x}(t),\, \boldsymbol{w}) \\ \boldsymbol{y}(t) = \boldsymbol{C}\boldsymbol{x}(t) \end{cases} \tag{1}$$

where $\boldsymbol{x} = [x_1,\, \ldots,\, x_N]^T \in \Re^N$ represents the internal state variables and N denotes the dimension of the state-space, or order of the model. The $\boldsymbol{g}_{\text{ANN}} = [g_{\text{ANN\_1}},\, \ldots,\, g_{\text{ANN\_N}}]^T$ represents a feed-forward Multilayer Perceptron (MLP) neural network [10], which has $M + N$ input neurons and N output neurons, with weight parameters $\boldsymbol{w}$ and a constant scaling parameter τ. $\boldsymbol{C} = [c_{ij}] \in \Re^{K \times N}$ is the output matrix, which maps the state space into the output space. The SSDNN model can accurately represent the behavior of a nonlinear circuit, such as a nonlinear I/O driver, only after the model has properly learned from the training data about the original circuit through an optimization process called training. The training data for SSDNN are in the form of transient waveforms, which are generated from the detailed nonlinear circuit. Let L denote total number of transient waveforms used for training. Let $\boldsymbol{u}_d^i(t)$ and $\boldsymbol{y}_d^i(t)$ represent ith input and output transient waveforms sampled in the time-interval $[0,\, T_1]$. Let $\boldsymbol{y}^i(t)$ represent the SSDNN prediction of the ith input waveform $\boldsymbol{u}_d^i(t)$. The goal of training process is to adjust parameters in $\boldsymbol{w}$ and $\boldsymbol{C}$ such that the objective function (dynamic training error) is minimized. Efficient training (optimization) of the model requires the sensitivity information of the model to be computed accurately and quickly. To facilitate fast sensitivity computation, the concept of the adjoint dynamic neural network [11] is expanded to cover the case of SSDNN. Specifically, an adjoint SSDNN system is formulated as

$$\dot{\hat{\boldsymbol{x}}} = \hat{\boldsymbol{x}} - \tau \frac{\partial \boldsymbol{g}_{\text{ANN}}}{\partial \boldsymbol{x}}^T \hat{\boldsymbol{x}} + \boldsymbol{C}^T(\boldsymbol{y} - \boldsymbol{y}_d) \tag{2}$$

where $\hat{\boldsymbol{x}} \in \Re^N$ is the vector of the adjoint state variables and initial condition of system is given by $\hat{\boldsymbol{x}}(T_1) = \boldsymbol{0}$.

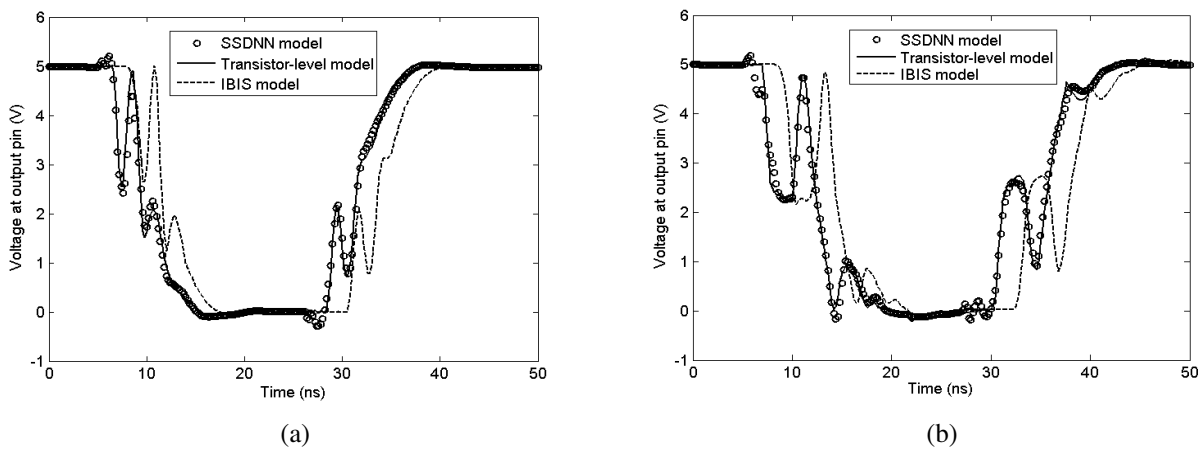(a)   (b)

Figure 1: Sample comparisons of the output voltage responses using detailed transistor-level model (–), IBIS model (- -), and the SSDNN model (o) for example D. The comparisons shown here are the cases for (a) $T_r = 1.25\,\text{ns}$ and $d = 1.25\,\text{cm}$, and (b) $T_r = 1.75\,\text{ns}$ and $d = 2.75\,\text{cm}$. Good agreement is achieved between the responses of the SSDNN models and those of detailed transistor-level models even though such test waveforms are never used in training.

4. NUMERICAL EXAMPLE

In this example, we present an example for the transient modeling of a commercial IC package using SSDNN techniques [8]. For this IC buffer, both the detailed transistor-level model and the IBIS model are available, where the former model is slow and accurate, and the latter model is faster but less accurate. In this example, the SSDNN model is used to represent the nonlinear dynamic behavior between the input pin and output pin of the IC package. The training waveforms are obtained from HSPICE simulations by varying the rise time of the input pulse T_r [0.5 ns–2.0 ns, step 0.5 ns] and the length of the transmission line load d [1 cm–3 cm, step 0.5 cm]. After training, the SSDNN model accuracy is confirmed by an independent set of test waveforms that are never applied during the training. Subsequently, the trained SSDNN model is used in commercial software HSPICE to perform the transient simulation of the buffer for SI analysis. Figure 1 shows the comparison of the voltages at the output pin using three types of models, i.e., detailed transistor-level model, IBIS model and the SSDNN model. It is demonstrated that the SSDNN model provide more accurate results than the IBIS model for predicting the SI effects of the high-speed interconnect with nonlinear buffer terminations.

REFERENCES

1. Achar, R. and M. S. Nakhla, "Simulation of high-speed interconnects," *Proc. IEEE*, Vol. 89, No. 5, 693–728, 2001.
2. *I/O Buffer Information Specification (IBIS) Ver. 4.1*, Electronic Industries Alliance, Jan. 2003. Available: http://eda.org/pub/ibis/ver4.1/
3. Stievano, I. S., I. A. Maio, and F. G. Canavero, "Parametric macromodels of digital I/O ports," *IEEE Trans. Adv. Packag.*, Vol. 25, No. 2, 255–264, 2002.
4. Stievano, I. S., I. A. Maio, and F. G. Canavero, "Parametric macromodels of differential drivers and receivers," *IEEE Trans. Adv. Packag.*, Vol. 28, No. 2, 189–196, 2005.
5. Stievano, I. S., I. A. Maio, and F. G. Canavero, "Mpilog macromodeling via parametric identification of logic gates," *IEEE Trans. Adv. Packag.*, Vol. 27, No. 1, 15–23, 2004.
6. Mutnury, B., M. Swaminathan, and J. P. Libous, "Macromodeling of non-linear digital I/O drivers," *IEEE Trans. Adv. Packag.*, Vol. 29, No. 1, 102–113, 2006.
7. Cao, Y., R. T. Ding, and Q. J. Zhang, "A new nonlinear transient modeling technique for high-speed integrated circuit applications based on state-space dynamic neural network," *IEEE MTT-S Int. Microwave Symp. Dig.*, 1553–1556, Fort Worth, TX, June 2004.
8. Cao, Y., R. T. Ding, and Q. J. Zhang, "State-space dynamic neural network technique for high-speed IC applications: Modeling and stability analysis," *IEEE Trans. Microwave Theory Tech.*, Vol. 54, No. 6, 2398–2409, 2006.
9. HSPICE ver. 2007.03, Synopsys, Inc., Mountain View, CA, 2007.
10. Zhang, Q. J. and K. C. Gupta, *Neural Networks for RF and Microwave Design*, Artech House, MA, 2000.
11. Cao, Y., J. J. Xu, V. K. Devabhaktuni, R. T. Ding, and Q. J. Zhang, "An adjoint dynamic neural network technique for exact sensitivities in nonlinear transient modeling and high-speed interconnect design," *IEEE MTT-S Int. Microwave Symp. Dig.*, 163–168, Philadelphia, PA, June 2003.

Near Field Imaging of Synthetic Aperture Radiometer

C. Zhang, **J. Wu**, **H. Liu**, and **W. Y. Sun**

Center for Space Science and Applied Research, Chinese Academy of Sciences
Beijing 100190, China

Abstract— The near field applications of synthetic aperture interferometric radiometer (SAIR) are receiving great interest in the recent years. Because the traditional far field Fourier imaging theory can not stand any longer in near field condition, developing the associate imaging theory will be an important objective of SAIR technique. This work is devoted to explore an effective near field imaging method for SAIR systems with traditional far field plane antenna array. The discrete numerical inversion method with focus plane approximation is developed. The Tikhonov regularization method is introduced to deal with the ill condition and measurement errors.

1. INTRODUCTION

Synthetic aperture interferometic radiometer (SAIR) is a promising technique for passive remote sensing. It can achieve high resolution without the problem of deploying large scanning antenna for traditional real aperture radiometer. Several airborne SAIR systems have been developed. ESTAR [1], 2D-STAR [2] and HUT-2D [3] are the representative ones. Space borne application projects of SAIR system such as ESA's SMOS/MIRAS [4] and NASA's GeoSTAR [5] are also being implemented. All these applications of aircraft and spacecraft instruments are operated in far field of the antenna array. With the development and maturity of integrate digital correlator technology, there has been growing interest in applying SAIR in short range imaging applications such as security detection of concealed weapons or other contrabands, all weather reconnaissance and surveillance, and ground penetrating imaging for landmine detection or archeology. As a cost-effective technique, SAIR will be widely used in near field applications. The associate near field imaging theory also should be well explored. Developing near field imaging technique is also helpful for SAIR system on-ground calibration and characterization inside range limited anechoic chamber or at open air experimental field.

2. BASIC THEORY OF NEAR FIELD INTERFEROMETRY

SAIR is a passive interferometry technique to measure the objects' natural radiation, which is composed by a set of correlation interferometer pairs. The electromagnetic signals received by each pair are cross correlated to get the so-called visibility function. In condition of far field measurement, i.e., the viewing scene is far away from the antenna array, the curvature of the incoming wave fronts can be neglected, so the incoming wave performs as plane wave (see Figure 1(a)). The cross correlation results of the antenna pairs with different antenna pair spacing (baselines) can be expressed as

$$V_{ij}(u,v) = \iint \frac{T_B(l,m) \cdot F_{ni}(l,m)F_{nj}^*(l,m)}{\sqrt{\Omega_i \Omega_j} \cdot \sqrt{1 - l^2 - m^2}} \cdot e^{-j2\pi(u_{ij}l + v_{ij}m)} dl \, dm \qquad (1)$$
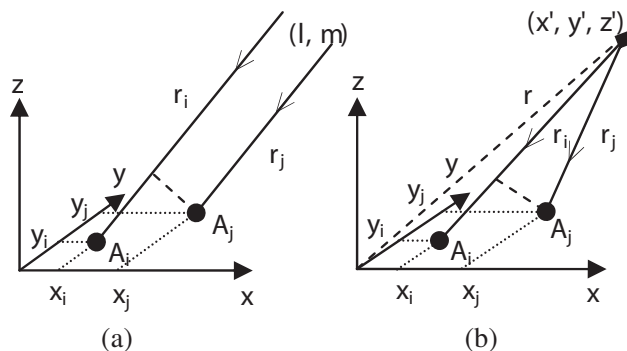

Figure 1: Geometry sketch of the far field (a) and near field (b) interferometric measurement.

where Ω_i and Ω_j are the solid angles of the antennae, $T_B(l,m)$ is the brightness temperature (with unit of Kelvin), F_n is the normalized antenna voltage pattern; $(u_{ij}, v_{ij}) = (x_i - x_j, y_i - y_j)/\lambda$ is the antennae separation measured in wavelengths; (l,m) is the direction cosines of the incidence angle. Thus the far field visibilities are exactly correspond to the samples of the spatial frequency of the incoming brightness temperature. Then the original brightness temperature image can be reconstructed by inverse Fourier transform.

For near field applications, the typical distance from targets to antenna array is limited in a short range. Comparing to the antenna array size, the viewing distance is not large enough to neglect the incident wave front curvature (see Figure 1(b)). A general separation criterion between the near field and field for SAIR system is [6]

$$R_{\text{far−field}} \geq 20 \frac{D_{syn}^2}{\lambda} \tag{2}$$

where D_{syn} is the synthetic aperture size, which approximately is twice as large as the largest baseline. When the targets are in the near field range of the synthetic aperture, the far field paraxial approximation that adopted in Equation (1) will not stand, and the relationship between the cross correlations and incident brightness temperature will not be Fourier transform anymore. The exact visibilities for near field can be expressed as

$$V_{i,j}^{NF} = \iint\limits_{\Omega_s} \frac{T_B(\theta, \varphi) F_{ni}(\theta, \varphi) F_{nj}^*(\theta, \varphi)}{\sqrt{\Omega_i \Omega_j}} e^{-jk(r_i - r_j)} \cdot \frac{r^2}{r_i r_j} d\Omega \tag{3}$$

where, r_i and r_j are the distances from the two associate antenna elements to the image pixel. The distances are changing with the image pixel location and antenna baselines, so theoretically it is impossible to analytically inverse Equation (3) like what the inverse Fourier transform applied for far field case. other appropriate techniques should be developed to deal with the near field imaging problem.

Now some solutions have been proposed for SAIR near field imaging problem. It can be classified into two types. One is based on hardware transformation as what described in [6] and [7]. It physically rearranges the antenna elements from planar array to circular or spherical array. The other one is based on software modification, as proposed in [8] and [9], that still keep the traditional planar array but theoretically add a correction phase term on the measured near field visibility to get the equivalent far field visibility. Both the two kind approaches can directly use traditional Fourier inversion algorithm, but have obvious disadvantages in practical case. The spherical array only fit for still scene with targets fixed on a given place, while the corrected Fourier method only fit for point source. It still remains large errors in off-boresight areas for complex extended source especially large 3D targets.
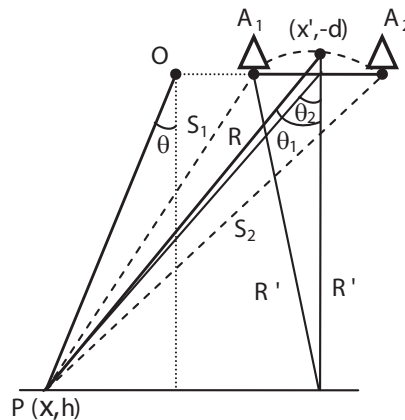

Figure 2: Geometry sketch of the near field interferometric measurement by one antenna pair.

3. NEAR FIELD DISTORTION

The distortion relationship between the far field and near field visibility should be well explored, that will be helpful for investigating the near field imaging mechanism. For the radiation source in

near field, the incident angles are different according to different antenna baselines. Take 1D linear array for example, the array center 'O' is defined as the coordinates' origin to scale the incident brightness distribution. Figure 2 shows out the geometry sketch of a baseline pair receiving the radiation wave from a random point source. The distance from antenna A_1 to point source P is

$$
\begin{aligned}
S_1 &= \sqrt{(x_1 - x)^2 + h^2} = \sqrt{[(x_1 - x') - (x - x')]^2 + h^2} \\
&= \sqrt{(x - x')^2 - 2(x_1 - x')(x - x') + R'^2} = R\sqrt{1 - \frac{2(x_1 - x')(x - x')}{R^2}}
\end{aligned} \tag{4}
$$

where $(x', -d)$ is the arc center between the two antenna elements (the arc radius is R'). The same result can be drawn for S_2. Then after taking the second-order Taylor approximation, these two distances can be expressed as

$$
S_1 = R\left[1 - \frac{(x_1 - x')(x - x')}{R^2} + \frac{(x_1 - x')^2(x - x')^2}{2R^4} + \sigma_1\right] \tag{5}
$$

$$
S_2 = R\left[1 - \frac{(x_2 - x')(x - x')}{R^2} + \frac{(x_2 - x')^2(x - x')^2}{2R^4} + \sigma_2\right] \tag{6}
$$

Noticing $|x_1 - x'| = |x_2 - x'|$, the even order terms of Equations (5) and (6) can be eliminated when calculating the wave path difference

$$
\Delta S = S_1 - S_2 = (x_2 - x_1)\frac{(x - x')}{R} + (\sigma_1 - \sigma_2) \approx u \cdot \sin\theta_1 \tag{7}
$$

where u is the baseline length. Considering the relations formula

$$
\sin(\theta_2 - \theta_1) = \frac{d\sin\theta_1}{h/\cos\theta_2}, \quad \tan\theta_2 = \tan\theta + \frac{x'}{h} \tag{8}
$$

after neglecting the amplitude errors due to the free space propagation, the near field visibility can be defined as

$$
V(u) = \int_{-\pi/2}^{\pi/2} T_B(\theta)F_i(\theta)\frac{F_j^*(\theta)}{\sqrt{\Omega_1\Omega_2}}\exp\left[jku\frac{\sin(\theta + \varphi)}{\zeta}\right]d\theta \tag{9}
$$

where

$$
\varphi \approx \cos^2\theta \cdot \frac{x'}{h} - \cos^3\theta \cdot \sin\theta \cdot \left(\frac{x'}{h}\right)^2, \quad \zeta = \sqrt{\left(1 + \frac{d}{h}\right)^2\cos^2(\theta + \varphi) + \sin^2(\theta + \varphi)} \tag{10}
$$

Comparing to far field visibility, Equation (9) has two more parameters φ and ζ, which can be deemed as the distortion factors of near field visibilities. These factors are determined by the antenna position and viewing distance. They can not be simply compensated by far field to near field transformation.

4. NUMERICAL INVERSION METHOD

We are intent to develop a more effective SAIR near field imaging technique for traditional planar antenna array. Taking into account the finite resolution of the reconstruction image, the original extended radiation source can be represented as a collection of independent point sources. The visibility integral can be discretized into a linear system of matrix equations

$$
V_{M\times1} = D_{M\times N} \cdot T_{N\times1} \tag{11}
$$

where V_M is the M measured near field visibilities. The elements of the column vector T_N represent the samples of the original brightness function. D is the coefficient matrix, defined by Equation (12), which characterizes the system configuration and objects spatial distribution.

$$
D_{m,n} = \frac{F_{1m}(\theta_n)F_{2m}^*(\theta_n)}{\sqrt{\Omega_{1m}\Omega_{2m}}}\exp\left[-jk2\pi\frac{\Delta s(m, n)}{\lambda_0}\right]\Delta\theta \tag{12}
$$

In order to minimize the discretization error in Equation (11), the discretization interval $\Delta\theta$ should be small. As a general criterion it should be $N \geq 3M$.

The problem in date processing is to reconstruct the unknown T_N by means of the observation V_M. But as a critical step, precisely getting every elements of the coefficient matrim D is a difficult work because the exact target's spatial distribution must be known to calculate the Δs. As mentioned in [9] focusing the near field visibilities to the pixel at bore-sight could only produces a smooth defocusing error for the pixels off-boresight. Following this idea and extending it in the numerical inversion, we adopt a new method that focuses the near field visibilities on a plane where the targets are equivalently located. Essentially this also means to approximate the objects' spatial distribution by a plane. Combined with the known antenna position, the matrix D will be easily computed.

Straightforward computing solution T' from Equation (11) is usually unpractical due to the error in the left hand side V and the ill condition of matrix D. Regularization methods can be used to suppress the amplified error components in V. The Tikhonov regularization has a wide application [11], which is defined as

$$\min_{T} \left(\parallel DT - V \parallel_2^2 + \lambda \parallel LT \parallel_2^2 \right) \tag{13}$$

where L is a regularization operator chosen to obtain a solution with desirable properties, such as the standard operator $L = I$ (the identity matrix) for a small norm. Thus Equation (13) can also be expressed as

$$\left(D^*D + \lambda I \right) T = D^*V \tag{14}$$

The regularization parameter λ is a positive scalar parameter that viewed as a posteriori turning parameter. L-curve can be used to compute the optimal λ at the L-shaped corner.

5. NUMERICAL SIMULATIONS

Numerical simulations are performed to demonstrate the algorithm feasibility. An 8-elements X band one dimensional SAIR system, shown in Figure 3, is used to simulate the near field measurement, where $\Delta u = 0.735\lambda\,(\lambda = 3.2\,\mathrm{cm})$. In point source imaging simulation, the point source is placed in the near field with the distance of $2\,\mathrm{m}$ to the antenna array. The reconstruction results by different methods are shown in Figure 4(a). It shows out that direct applying far field (FF) Fourier method in the near field case will produce heavy distortions that the main-lobe is shifted and decreased while the side-lobe is disorderly increased. The near field (NF) numerical reconstruction has a good coincidence with the ideal point response. The imaging simulations of an extended line source are also implemented. This extended source has different brightness temperature and range distance in different parts. The equivalent focus plane is set $2\,\mathrm{m}$ away from the antenna array. The near field Fourier method with bore-sight focus correction used in [9] is also applied. The reconstruction results are shown in Figure 4(b), it indicates that the plane focus based near field numerical method has the best performance. It performs close to the ideal far field imaging results with reasonable Gibb's wrinkles.
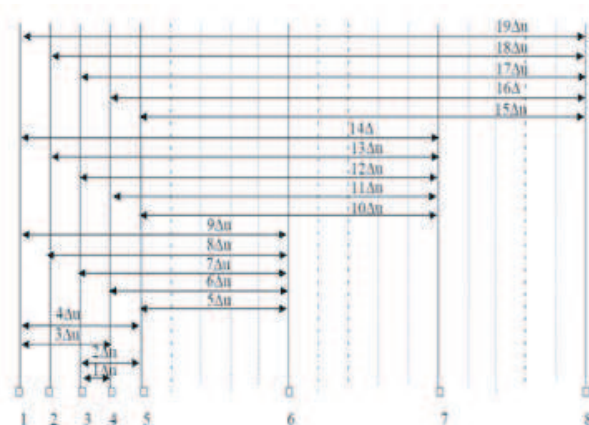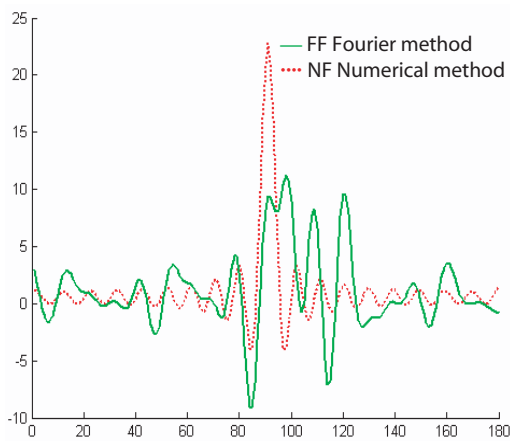

Figure 3: Arrangement of the SAIR array.
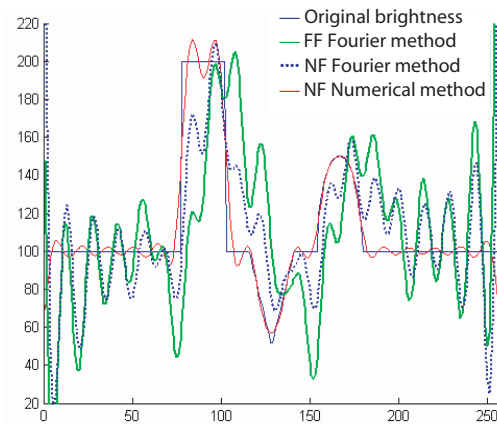
Figure 4: Simulations of point source.

Figure 5: Simulations of extended source

6. CONCLUSIONS

The near field imaging problem of SAIR is analyzed and mathematically modeled in this paper. A new imaging method for SAIR near field application is developed. It essentially bases on focus plane approximation and numerical inversion. This method can be applied to planar array, so the traditional far field SAIR system can be directly implemented in near field applications with only taking some theoretically modification in the imaging algorithm. Note that, in order to overcome the ill condition of the near field problem, the Tikhonov regularization technique has been employed. Numerical simulations demonstrate the advantage and feasibility of this method.

ACKNOWLEDGMENT

REFERENCES

1. Le Vine, D. M., A. J. Griffis, C. T. Swift, et al., "ESTAR: A synthetic aperture microwave radiometer for remote sensing applications," *Proceedings of the IEEE*, Vol. 82, No. 12, 1787–1801, 1994.

2. Le Vine, D. M., T. J. Jackson, and M. Haken, "Initial images of the synthetic aperture radiometer 2D-STAR," *IEEE Trans. Geosci. Remote Sens.*, Vol. 45, No. 11, 3623–3632, 2007.

3. Rautiainen, K., R. Butora, T. Auer, et al., "Development of airborne aperture synthetic radiometer (HUT-2D)," *IEEE Proceedings of IGARSS*, No. 2, 1232–1234, 2003.

4. Silvestrin, P., M. Berger, Y. Kerr, and J. Font, "ESA's second earth explorer opportunity mission: The soil moisture and ocean salinity mission-SMOS," *IEEE Geoscience and Remote Sensing Newsletter*, No. 118, 11–14, 2001.

5. Lambrigtsen, B., "GeoSTAR: Developing a new payload for GOES satellites," *Proceedings of the 2006 IEEE/AIAA Aerospace Conference*, Big Sky, MT, USA, March 2006.

6. Laursen, B. and N. Skou, "Synthetic aperture radiometry evaluated by a two channel demonstration model," *IEEE Trans. Geosci. Remote Sens.*, Vol. 36, No. 3, 822-832, 1998.

7. Walsh, K. P., B. Schulkin, et al., "Terahertz near field interferometric and synthetic aperture imaging," *Proceeding of SPIE*, Vol. 5411, 1–9, 2004.

8. Duffo, N., I. Corbella, F. Torres, A. Camps, and M. Vall-Llossera, "Advantages and drawback of near field characterization of large aperture synthesis radiometers," *MicroRad*, Rome, 2004.

9. Tanner, A., B. Lambrigsten, T. Gaier, and F. Torres, "Near field characterization of the GeoSTAR demonstrator," *Proceedings of IEEE International Geoscience and Remote Sensing Symposium*, 2529–2532, Denver, Co, USA, July 2006.

10. Peichel, M., H. Suess, and M. Suess, "Microwave imaging of the brigthness temperature distribution of pextended areas in the near and far field using two-dimensional aperture synthesis with high spatial resolution," *Radio Science*, Vol. 33, No. 3, 781–801, 1998.

11. Tikhonov, A. and V. Y. Arseninn, *Solution of Ill-Posed Problems*, John Wiley & Sons, New York, 1977.

A Method to Estimated Winter Wheat Yield with the MERIS Data

Xin Du[1, 2], Bingfang Wu[1], Qiangzi Li[1], Jihua Meng[1], and Kun Jia[1, 2]

[1]Institute of Remote Sensing Applications, Chinese Academy of Sciences, China
[2]Graduate University of Chinese Academy of Sciences, China

Abstract— The MERIS data will provide us time-serial parameters, with which we can get the information that describe the growth process of winter wheat, so inputted these information into a biomass model with the real time meteorological data, and then we could get the biomass of winter wheat. For calculating the yield of winter wheat, the harvest index (HI) must be estimated. There were very rarely people to estimated HI of crop with the remote sensing data. Based on the own character of winter wheat, we found that the correlation between HI and the ratio of $NDVI_{pre}$ (the average value of NDVI form emergence to anthesis) and $NDVI_{post}$ (the average value of NDVI from anthesis to maturity) is quite well. Therefore, with the NDVI calculated from remote sensing data, we can estimate the HI of winter wheat for every pixel. Finally, we can use the follow expression to calculate the yield of winter wheat, $YIELD = BIOMASS \times HI$, and $HI = f(NDVI)$. With this method to estimate the yield of winter wheat, the temporal and spacial information of MERIS data has been used adequately. And the precision of the yield of winter wheat estimated with this method is better than that with a fixed HI. Because of the advantage and objectivity of the remote sensing data, this method is very useful and adequate for estimating the yield of winter wheat in daily operation.

1. INTRODUCTION

It is very important to estimate crop yield accurately, and now a lot of studies have been carried out, but most of them were statistical and experiential, which only crop yield value could usually be obtained, and seldom got the spatial distribution information of crop yield [2, 4]. With the remote sensing data, it is timely and external to get plenty of information of the grand, so it is preponderant to estimate crop yield promptly [5]. Nowadays, it is popular topics in remote sensing applications to estimate crop yield at regional or global scale.

Remote sensing is a powerful and efficient tool for estimating crop yield at regional or global scale. A lot of methods using remote sensing data, such as AVHRR, MODIS [1], have been developed for estimating crop yield. And this paper is going to improve the methods with the MERIS data.

2. STUDY AREA AND DATA DESCRIPTION

2.1. Study Area

Because of its key position in food security, winter wheat yield estimated has global importance. China is the biggest wheat producing countries in the world, and its wheat production accounts for nearly 20% of that of the whole world. Shandong province is one of the biggest wheat producing provinces in China, the wheat production of Shandong province accounts for 20% of China's total winter wheat production. Yucheng city locates at west of Shandong province, and it is an agriculture city, most people there rely on crop production. Due the above reasons, around Yucheng city, there has been chosen as the monitoring area of the study.

According to the traditional tillage practice, there is the intensive dual-cropping system based on winter wheat and autumn crops, such as maize, soybean and sorghum et al. Different from the variety of crops in autumn, winter wheat is the dominant summer crop in the area, and it is sown in early October, harvested in early or mid June next year.

Annual precipitation in the area is about 600 mm, and it is concentrated between July and September during the summer monsoon, which is more than 50%. As the limited and variable precipitation, the winter wheat must be irrigated to guarantee the productivity in spring.

2.2. RS Data

The RS data used in the study is MERIS (MEdium Resolution Imaging Specrometer Instrument). MERIS is one of the sensors on ENVIronmental SATellite (ENVISAT) which was launched by the European Space Agency (ESA) on March 2002. Although MERIS was primarily dedicated to ocean color, its bands configuration broadened its application to vegetation monitoring [6].

The MERIS data can provide time-serial terrestrial parameters at s several-day frequency, with which we can monitor the whole growing process of winter wheat.

The normalized different vegetation index (NDVI) derived from red band and near infrared band is a directly remote sensing indicator, and it was calculated using red and near infrared (NIR) reflectance as follows

$$\text{NDVI} = (\text{NIR} - \text{RED})/(\text{NIR} + \text{RED}) \tag{1}$$

where RED and NIR is canopy reflectance at red and near infrared, respectively. As for MERIS, band 13 (855–875 nm) and band 8 (677.5–685 nm) are assigned as near infrared and red bands.

In the frame of the Dragon 2 Program, MERIS data was collected since February to July 2008 covering Shandong province. After these data was pre-processed, NDVI was computed form each scene of MERIS data.

2.3. Basic Data

The most important basic data in the study is the $1:100,000$ land-use/cover digital data and land use map, which was derived from Landsat Thematic Mapper (TM) images. And the land use/cover digital data was used to identify arable fields.

2.4. Observation Data

One $5\,\text{Km} \times 5\,\text{Km}$ experimental area are selected in Yucheng (116 34 09E-30 49 52N, Shandong province), and there are 25 observing points (in 5 $300\,\text{m} \times 300\,\text{m}$ plots) distributed in the experimental area. In every key phenophases of winter wheat such as heading, flowering and harvesting date, the key crop parameters including leaf area index (LAI), height, dry weights et al. are observed and recorded. And these data was used not only to calibrate the model, but also to validate the estimating result.

3. METHODS

3.1. Estimation of above Ground Biomass

The total above ground biomass of winter wheat accumulated over the growing season was calculated as

$$\text{BIOMASS}_{\text{AGB}} = \sum_{i=1}^{n} \varepsilon_i \text{FAPAR}_i \text{PAR}_i \tag{2}$$

where ε_i is the radiation use efficiency (RUE, gMJ^{-1}), which is variable on the basis of winter wheat developmental stage, and PAR$_i$ is the total incident photosynthetically active radiation (MJ) at day i, which is a fraction (0.48 in this study) of incoming solar radiation; FAPAR the fraction of absorbed PAR, which could be estimated form the simpleratio (SR) or NDVI by linear functions. Here FAPAR was calculated as a linear function of SR [9]:

$$\text{FPAR} = \frac{(\text{SR} - \text{SR}_{\text{min}})(\text{FAPAR}_{\text{max}} - \text{FAPAR}_{\text{min}})}{\text{SR}_{\text{max}} - \text{SR}_{\text{min}}} + \text{FAPAR}_{\text{min}} \tag{3}$$

where SR = NIR/RED, and SR$_{\text{min}}$ and SR$_{\text{max}}$ correspond to the second and 98th percentile of SR for the entire cropland regions, and FAPAR$_{\text{min}}$ and FAPAR$_{\text{max}}$ are defined as 0.01 and 0.95, respectively. SR$_{\text{min}}$ and SR$_{\text{max}}$ are computed in every 10 days.

3.2. Estimation of Harvest Index

In the paper, NDVI during the grain filling stage was hypothesized to be indirectly related to the yield of winter wheat accumulation. Therefore, the distribution of the above ground biomass can be simulated with the available winter wheat NDVI values to modify the final harvest index.

The harvest index (HI) of winter wheat was estimated as

$$\text{HI} = \text{HI}_{\text{max}} - \text{HI}_{\text{range}} \left(1 - \frac{\text{NDVI}_{\text{post}}}{\text{NDVI}_{\text{pre}}} \right) \tag{4}$$

where HI$_{\text{max}}$ is the optimal value of HI (here chosen equal to 0.5), HI$_{\text{range}}$ is the possible range of HI (here chosen equal to 0.2), NDVI$_{\text{pre}}$ is the average value of NDVI form emergence to anthesis, and NDVI$_{\text{post}}$ is the average value of NDVI from anthesis to maturity.

3.3. Estimation of Winter Wheat Yield

Using the computed BIOMASS$_{\text{AGB}}$ and HI, winter wheat yield (WY) was simply estimated as

$$\text{WY} = \text{BIOMASS}_{\text{AGB}} * \text{HI} \tag{5}$$

4. RESULTS

There are the estimated results of winter wheat around Yucheng city, Shandong province in 2008 in Fig. 1.
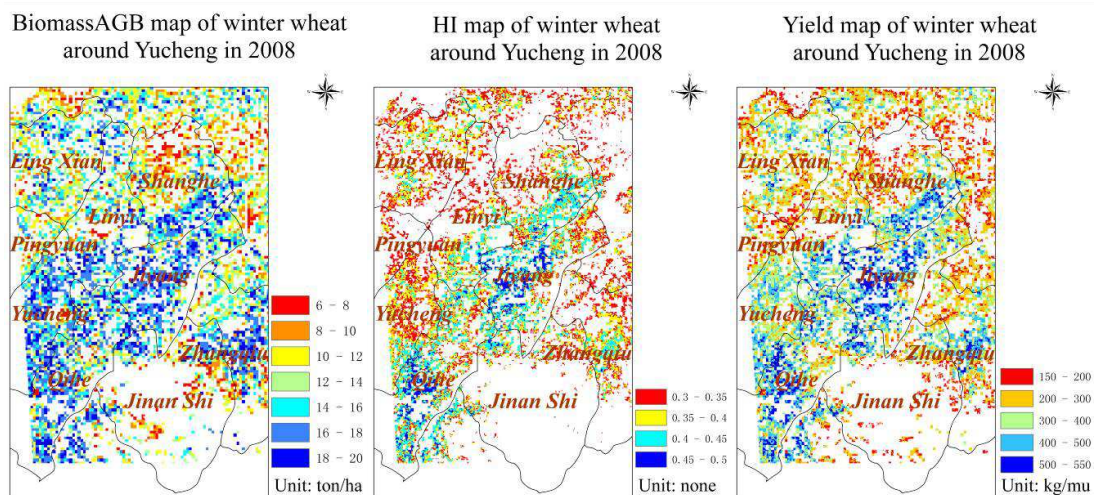

Figure 1: Biomass, HI, yield map of winter wheat around Yucheng in 2008.

In this area, the total above ground biomass of winter wheat is generally higher in the southwest than that in the northeast. Compared with the above ground biomass around Huimin county, that around Yucheng city was about 8∼10 ton/ha higher. As for the HI of winter wheat, it is generally 0.1 higher in the southwest than that in the northeast.

The final yield of winter wheat is much more accurate than that by using a fixed HI ($r = 0.75$ with HI map versus $r = 0.6$ with HI = 0.45).

5. CONCLUSION

We have developed a method for estimating the yield of winter wheat in the Plain of China from time series MERIS data. There are four steps to estimate the yield of winter wheat: 1) pre-processing of time series MERIS data; 2) estimation of the total above ground biomass of winter wheat; 3) estimation of HI of winter wheat; 4) estimation of yield of winter wheat.

The advancement of this method to estimate yield of winter wheat with remote sensing data is the estimated HI map, the result was more accurate and external compared with a fixed HI. And with this method, we can utilize the remote sensing data much better in the spatial and temporal dimensions. This method must have potential in application to other grain crops, such as corn, rice et al.

It was proved that the MERIS data also has an immense potential in time series monitoring of crops. The application of the yield of winter wheat estimating in the other region and other grain crop yield estimating with MERIS data are going to be addressed in our future study.

REFERENCES

1. Moriondo, M., F. Maselli, and M. Bindi, "A simple model of regional wheat yield based on NDVI data," *Europ. J. Agronomy*, Vol. 26, 266–274, 2007.
2. MacDonald, R. B. and F. G. Hall, "Global crop forecasting [J]," *Science*, Vol. 208, 670–679, 1980.

3. Meng, J., B. Wu, Q. Li, and X. Du, "Monitoring crop phenology with MERIS data — A case study of winter wheat in North China Plain," received by *Progress In Electromagnetics Research Symposium*, Beijing, China, March 2009.

4. Meng, Q., Q. Li, and B. Wu, "Operational method for crop yield prediction," *Journal of Remote Sensing*, Vol. 8, No. 6, 602–610, 2004.

5. Wu, B., "China crop watch system with remote sensing," *Journal of Remote Sensing*, Vol. 8, No. 6, 481–194, 2004.

6. Dente, L., G. Satalino, F. Mattia, and M. Rinaldi, "Assimilation of leaf area index derived from ASAR and MERIS data into CERES — Wheat model to map wheat yield," *Remote Sens. Environ.*, Vol. 112, No. 4, 1395–1407, 2008.

7. Field, C. B., J. T. Randerson, and C. M. Malmstrom, "Global net primary production: Combining ecology and remote sensing," *Remote Sens. Environ.*, Vol. 51, 74–88, 1995.

8. Potter, C. S., J. T. Randerson, C. B. Field, P. A. Matson, P. M. V. Vitousek, H. A. Mooney, and A. K. Steven, "Terrestrial ecosystem production: A process model based on global satellite and surface data," *Glob. Biogeochem. Cycles*, Vol. 7, 8110–841, 1993.

9. Sellers, P. J., D. A. Randall, G. J. Collatz, J. A. Berry, C. B. Field, et al., "A revised land surface parameterization (SiB2) for atmospheric GCMs," *Model Formulation J. Climate*, Part 1, Vol. 9, 676–705, 1996.

VLF Sferics Propagating in the Earth-ionosphere Waveguide

A. B. Bhattacharya[1], **S. Joardar**[2], **R. Bhattacharya**[3], **S. Sarkar**[1], **and S. Das**[1, 4]

[1]Department of Physics, University of Kalyani, Kalyani 741235, West Bengal, India
[2]Giant Meterwave Radio Telescope, Tata Institute of Fundamental Research
Khodad, Narayagaon, Pune 410504, Maharashtra, India
[3]Department of Environmental Science, University of Kalyani
Kalyani 741235, West Bengal, India
[4]Department of Physics, Serampore College
Serampore, Hooghly 712201, West Bengal, India

Abstract— The electromagnetic waves originating in lighting discharges at frequencies in the VLF band are reflected by the ground as well as by the conducting layer of the ionosphere and thus they are guided efficiently around the earth. As in this earth-ionosphere waveguide sferics propagate with low loss, they can be detected even at long distances from their source locations. Our observations in a tropical country India, over three decades, on VLF propagation reveal some interesting phenomena under realistic conditions. The variations of the overall intensity of radio atmospheric signal at the time of sunrise and sunset as well as by the so called cosmic ray layer formation after sunrise have been critically examined besides investigations on geomagnetic effects producing long period fading in sferics level. The effects of land and sea thunderstorms in the sferics level occurring over this area have also been taken into consideration. As the wave propagates over greater distances, the excitation and attenuation vary significantly with the parameters of the waveguide, such as reflection height, anisotropy etc. which change over time and also over the typically long VLF propagation path.

1. INTRODUCTION

Lighting, an electric discharge occurring in the atmosphere of the earth, has a total length of tens of kilometers and there are about 100 discharges at each second throughout the globe. The electromagnetic radiation radiated during lighting discharges may be received by sensitive receivers as sferics. Such radiation is produced by the acceleration of electric particles and may provide information about the associated atmospheric noise. The radiation in the form of pulse when split up shows a wide range of frequencies. Long distance propagation of such electromagnetic waves is controlled largely by an invisible layer of charged particle, known as ionosphere. The electromagnetic waves at frequencies in the VLF band are reflected by the ground as well as by the conducting layer of the ionosphere and thus they are guided efficiently around the earth [1]. As in this earth-ionosphere wave guide sferics propagate with low loss, they can be detected even at long distances from their source locations. Knowledge of the nature of VLF sferics propagation in the tropics is of special interest for a number of reasons. One important reason is that fifty percent of the surface of the globe lies between locations 30°N and 30°S and over a third of the world's population inhabits tropical lands.

2. SELECTION OF FREQUENCY BAND

The pressing need for the long-range navigational aids and worldwide frequency standard initiated interest in investigating characteristics of VLF propagation. The VLF sferics, which impose a fundamental limitation on the efficiency of a VLF radio communication system reveal a spectral distribution of power [2] with a peak at about 10 KHz. Studies of the sferics variation from its regular behavior can have importance in two ways: (i) it provides information about the nature and intensity of the noise energy at the observation site and thereby the nature of the electrical activity in the clouds, (ii) it helps in the design of radio communication systems appropriate for use at a particular frequency. Lack of data on VLF sferics in the eastern India prompted us to make an elaborate study of its propagating characteristics in the earth-ionosphere wave guide so as to meet the needs of communication service.

3. OBSERVATIONS

With a view to investigating the characteristics of VLF sferics, we examined our round-the-clock records obtained in Calcutta (now renamed as **Kolkata**) (Lat., 22°34′N; Long., 88°24′E) going

back to 1976 and then at Kalyani (Lat., 22°58′N; Long., 88°28′E), 40 Km away from Kolkata. The graphical locations of both Kolkata and Kalyani are very close to the Bay of Bengal (Fig. 1), where clouds initially developed during the monsoon (June to September) and post monsoon months (October to February) and, in fact, the sferics originating in these local clouds at such times have a dominant role over other widely distributed weak sources as confirmed by radar observations. On the contrary, during the pre monsoon months (March to May), the cumulus clouds formed initially in the Asansol (Lat., 23°40′N; Long., 87°01′E) and Hazaribag (Lat., 23°58′N; Long., 85°26′E) area when approaching towards Kolkata converted to a full fledge cumulonimbus clouds, thus producing vigorous sferics activity in its path of journey towards our observing stations. The activity of both monsoon/post monsoon clouds as well as that of pre monsoon thunderclouds becomes maximum at afternoon hours of the day while the effect of sea thunderstorms is maximum at pre midnight hours. The sferics level in the sunrise and sunset hours of the day are mainly controlled by the long distance sferics sources because of the poor local cloud activity at these hours. It is the purpose of the paper to focus the variations of the overall intensity of radio atmospheric signal at the time of sunrise and sunset as well as the effect in the noise level by the so called cosmic ray layer formation after sunrise. Long-period fading in sferics during severe meteorological disturbances and associated solar geophysical phenomena are reported in this paper. The effects of land and sea thunderstorms in the sferics level showing characteristic differences with the fading pattern have also been pointed out. The receivers used for the collection of VLF sferics data were constructed to have a large dynamic range with a view to handling a fairly wide range of field intensities owing to the activity of thunderclouds in the neighboring regions. The receivers have an overall bandwidth of 1 KHz with charging and discharging time constants 3 s and 25 s respectively.
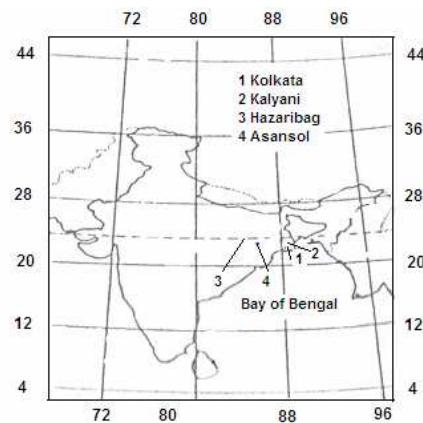

Figure 1: Sketch showing the locations of Kolkata, Kalyani, Hazaribag, Asansol and Bay of Bengal.

4. ANALYSES AND RESULTS

A photograph of typical record of sferics at three harmonically related frequencies 10, 20 and 30 KHz as recorded over Kolkata is shown in Fig. 2, which exhibits all the regular variations [3], *viz.*, sunrise effect (A), first minimum (B), recovery effect (C), morning minimum (D), afternoon maximum (E), late afternoon minimum (F) and night maximum (G). In the early morning and late afternoon hours remarkable changes occur in the ionospheric layers. Our studies exhibit a stepped-falls during sunrise of which the first two steps observed before ground sunrise is relatively weak while the last one called the main stepped-fall observed around ground sunrise is prominent. During sunset hours, the onset times of the main step are before the local ground sunset but for the second and third steps, the onset times are after that. Typical records of sferics showing three steps, two steps and one step fall of the sunrise effect are presented in Fig. 3. The geometry of the problem can be explained using Fig. 4 where we have shown an east-west section through the station and the center of the earth. For any position of the receiver R, there exist sferic sources towards east and west from which radio waves can be received by transmission. As absorption takes place over the sunlit part of the ray path, before sunrise the sources present on the eastern side are ineffective but those on the western side are received by the multi hop paths. It appears that the steps in the sunrise fall are produced by the successive elimination of the effective western sources through the switching off of the 3-hop, 2-hop and 1-hop rays. The sloping fall second and first hop rays. The
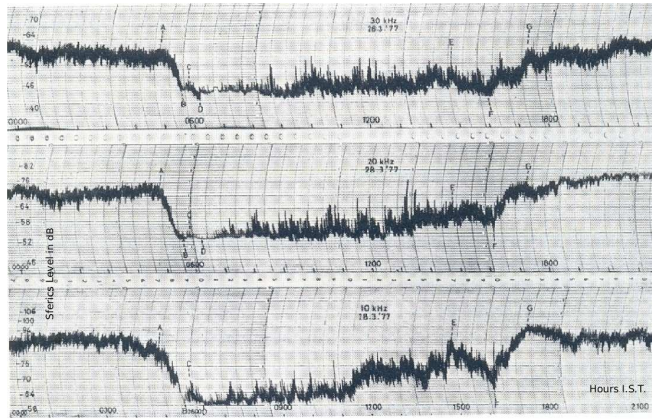
Figure 2: Photograph of the typical records of the integrated field intensity of sferics at 10, 20 and 30 KHz over Kolkata (ordinates are in dB above $1\,\mu$V/m, bandwidth 1 KHz).
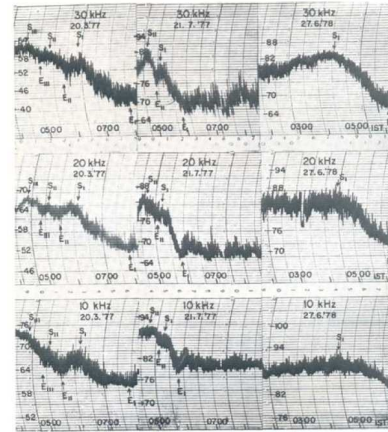

Figure 3: Typical records of sferics showing three step, two steps and one step fall of the sun rise effect.
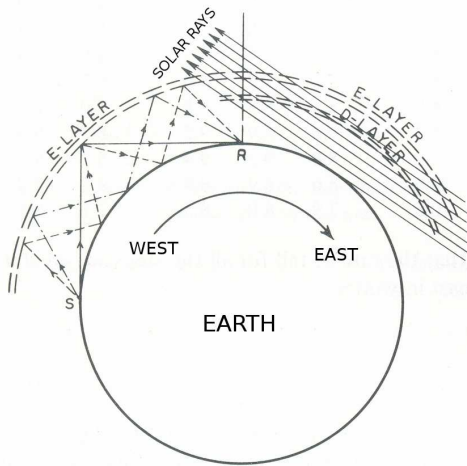

Figure 4: An east-west section through the station and the centre of the earth drawn for explaining the observed steps in the sunrise fall.
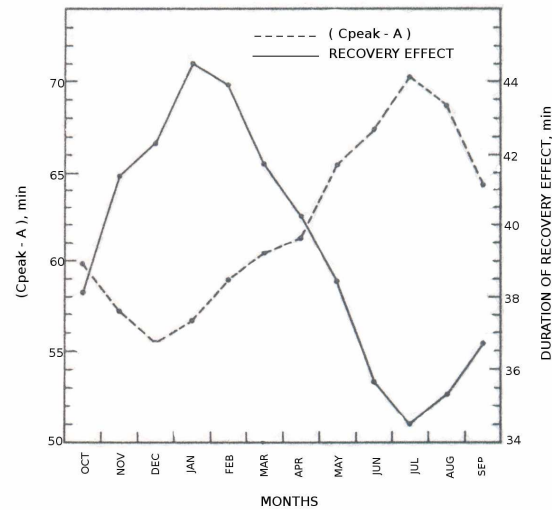

Figure 5: Curves showing the average monthly variation of the (C_{peak}-A) and the mean duration of Recovery Effect.

recovery effect (as marked by C in Fig. 2) in the field intensity of sferics during a particular step represents continually increasing D-layer absorption of waves on there downward path through the newly formed part of D-layer after the last reflection at the E-layer. If this conjecture is true, the horizontal extent of D-region would be increased successively for the third, exhibits a significant seasonal change in the amplitude. Also the time difference between occurrences of the peak value of this effect (taken as C_{peak}) and the initial level of the start of the sunrise fall (A), i.e., (C_{peak}-A) when plotted for different months (Fig. 5) shows a significant variation. In the same figure we have also superposed the mean duration of the recovery effect. It is seen that the curve is almost a mirror image of the former one, with a minimum value during the cloudy monsoon season and a maximum during the clear months. The recovery effect appears to be due to a transient improvement of propagation from nearer sources for which the reflection from E-region is still effective. The D-region, at present, is believed to consist of two layers. The ionization of the lower of the two layers, the so-called 'C-layer' remains constant throughout the day. It appears, therefore, that in daytime and at short distances where the 'C-layer' is penetrable; the ionization only above the 'C-layer' contributes towards the variation of field intensity of VLF sferics [4]. Our sferics records during torrential rainfall caused by violent monsoon and post monsoon depressions, exhibit distinct long period fadings both at day and night hours. The deep depression of 8th November 1995 (Fig. 6)

and associated character figure show that the magnetic character figure (C_p) values from 4th to 8th November were as high as 0.6, 1.0, 1.1, 0.5 and 0.3 respectively. The long period fadings as noticed at 27 KHz sferics at such times over Kalyani are believed to originate from the energetic solar particles and plasma which, in some way, initiate and stimulate the lower atmosphere [6]. The effect of land and sea thunderstorms in the sferics level, however, exhibit an entirely different characteristics like gradual rise, sudden enhancement and steady recovery of the level corresponding to developing, mature and dissipating stages of thunderstorms.
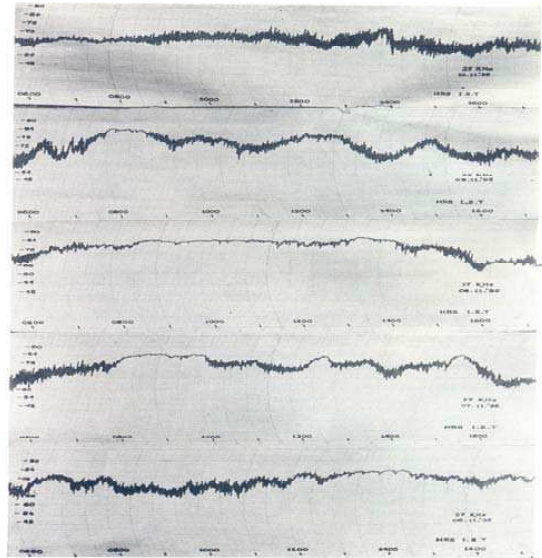

Figure 6: Long period fading at 27 KHz sferics during deep depression.

5. CONCLUSIONS

Summarizing our analyses of sferics data it may be pointed out that as the wave propagates over great distances, the excitation and attenuation significantly vary with the parameters of the wave guide like reflection height, anisotropy, prevalent atmospheric conditions etc. The results further suggest that the west bound wave modes have a higher level of "quasiness" than east bound modes. The variation in the reflection height of the ionosphere are responsible for converting energy from one waveguide mode to another waveguide mode which appears to be more dominant when the path of propagation crosses the day-night or night-day terminators [7].

REFERENCES

1. Inan, U. S. and A. S. Inan, *Electromagnetic Waves*, 1–30, Prentice hall, U.K., 2000.
2. Bhattacharya, A. B., B. K. Datta, and R. Bhattacharya, "Some distinct effect of tropical monsoon clouds as derived from atmospherics," *Theo. Appl. Climatol.*, Vol. 50, 83–92, 1994.
3. Sen, A. K., A. B. Bhattacharya, and S. K. Sarkar, "Very low frequency atmospheric radio noise field intensities," *Arch. Met. Geoph. Biokl. Ser. A*, Vol. 31, 263–268, 1982.
4. Sarkar, S. K., A. B. Bhattacharya, and A. K. Sen, "Some studies on the stepped fall of atmospherics during sunrise," *Arch. Met. Geoph. Biokl. Ser. A*, Vol. 30, 279–285, 1981.
5. Bhattacharya, A. B. and R. Bhattacharya, "Recovery effect in the field intensity of atmospherics," *Indian J. Radio Sp. Phys.*, Vol. 12, 30–32, 1983.
6. Bhattacharya, A. B., M. K. Chatterjee, and S. K. Kar, "Fading at 27 KHz atmospherics due to winter depression," *Indian J. Radio Sp. Phys.*, Vol. 26, 200–203, 1997.
7. Powar, S. D. and A. K. Kamra, "Recovery curves of the surface electric field after lighting discharges occurring between the positive charge pocket and negative charge center in a thundercloud," *Geophys. Res. Let.*, Vol. 29, 2108–2111, 2002.

A Roadmap for Detecting Extraterrestrial Intelligent Life

A. B. Bhattacharya[1], **S. S. Banerjee**[1], **and R. Bhattacharya**[2]

[1]Department of Physics, University of Kalyani, Kalyani 741235, West Bengal, India
[2]Department of Environmental Science, University of Kalyani, Kalyani 741235, India

Abstract— Search for extraterrestrial intelligence is indeed a great challenge to astronomers since its method of communication, direction of propagation as well as spectral pattern are all unknown beforehand. Due to all of these various limitations, the problem have become complex and thus the science has become a 'hard science'. As galactic societies would be transitory, an obvious solution is an interstellar communication network, or type of library mostly consisting of automated systems to store the cumulative knowledge of vanished civilizations and thereby to communicate the knowledge through the galaxy. This is called 'interstellar internet', where the different automated systems act as network servers. Intercepting such signals is highly difficult and extremely complex. The paper reviews the modern techniques and instruments for cosmic search. Theories and hypothesis developed for investigating life in the universe including some exciting findings have been critically examined. Finally mythology and exotheology about the universe have been pointed out focusing the future role of where is everyone.

1. INTRODUCTION

There are an estimated 10^{22} stars in the universe but only the Sun is known to have a planet, the Earth, which harbors life. Naturally, the question that appears is that are we unique and alone with a cosmic curiosity or are there other abodes of life. An answer would be of profound significance and is becoming an increasingly important phase of radio astronomy. The question that arises first in mind is where should we search; whose immediate answer is as the Sun is a yellow G-type star so in the absence of other information it would be reasonable to begin searches of sunlike stars like F, G and K-type stars [1]. But the most important question is concerned with the selection of wavelength when searching is to be attempted from radio astronomical point of view. Some initial efforts and the consequent limitations put a burning question to the radio astronomers "are there some wavelengths that would be better for gathering information than the usual one we using at present?"

2. SELECTION OF WAVELENGTH

For a suitable selection of wavelength, we have to consider what kind of signal one might expect from a distant civilization. If any such civilization transmits a signal intentionally, it will most likely be narrow-band as that signal is supposed to travel a long path for a given power. Besides the consideration of band, there should be a preference of wavelengths that would be better. The wavelength for which natural radiation from the galaxy is lowest and for which waves will move with least absorption may be apparently taken as the most appropriate one. When these two points get priority, wavelengths between a few millimeters and 30 centimeters would appear to be the most ideal one.

3. SKY NOISE DIAGRAM AND THE APPROACH

Figure 1 exhibits a sky noise diagram, showing plot of sky noise temperature against wavelength. The variation of noise due to earth's atmosphere, the galactic noise and the photon noise are also shown in the diagram. The hydrogen line and the water line are indicated in the figure, in addition to both Earth window and Cosmic window [2]. It is evident from the figure that the hydrogen (H) line is at 21 cm while the hydroxyl radical (OH) line is at 18 cm. When hydrogen and hydroxyl radical are combined it makes water (H_2O) whose wavelength obviously lies between 18 cm and 21 cm. This window may be taken as the first water hole which can be exploited by galactic civilizations as well as by our own civilization to make radiolink with each other. Over and above, it seems highly interesting to predict from Fig. 1 that there is a 2 mm line of start of the water molecule which extends up to 14 mm line and this might be regarded as bracketing a second "water hole".
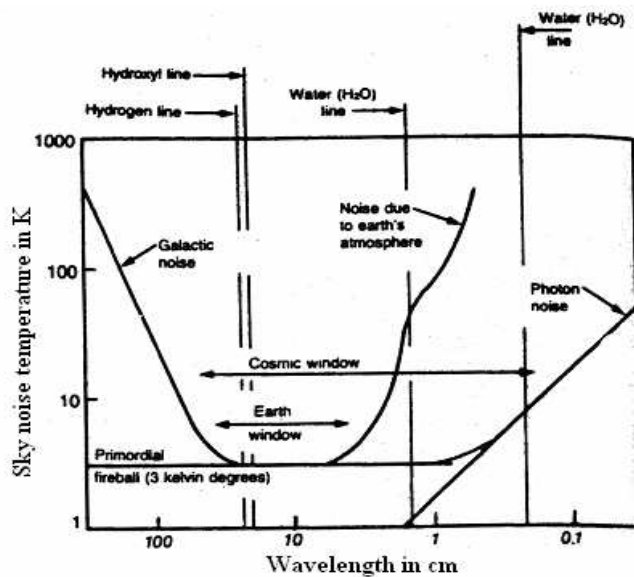
Figure 1: Sky noise diagram illustrating that the two sky noise windows, referred to as the "water holes", are most appropriate for interstellar communications.

4. THE RADIO COMMUNICATION LINK

Let a radio transmitter installed at earth's surface radiates a power P_t isotropically over a transmitter bandwidth B_t (Fig. 2).
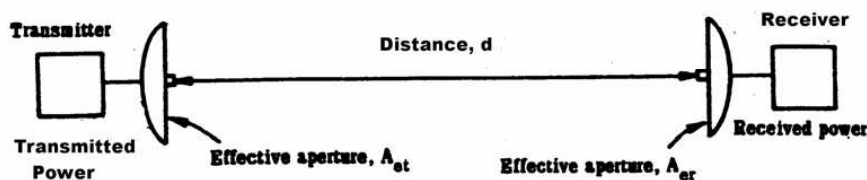

Figure 2: Radio communication link between transmitter to receiver and back.

The distance of the radio communication link can be derived from,

$$d = \sqrt{\frac{P_t A_{et} A_{er}}{k T_{sys} \left(S/N\right) \lambda^2 B_r}} \tag{1}$$

where, k = Boltzmann constant, T_{sys} = system temperature, S/N = signal-to-noise ratio and λ = wave length.

Equation (1) indicates that smaller the receiver bandwidth B_r, the larger is the range d. Under marginal conditions, the communication thus favors the use of small bandwidth. From some critical considerations, Drake and Helou [3] suggested that due to time broadening by multiple scattering of the interstellar medium, a minimum usable bandwidth is 0.1 Hz approximately.

5. EARLY SEARCHES

The first attempt to search extra terrestrial life was made by Drake [4] in 1960 with a radio telescope of antenna diameter 26 m at N.R.A.O., Green Bank. It was a 'targeted search' for two stars, Tau Ceti and Epsilon Eridani, at the hydrogen line corresponding to 1.4 GHz with a single tunable channel of 100 Hz bandwidth. Dixon [5] conducted a long search over 10 years with 110 m radio telescope at Ohio State University. Almost at the same time Horowitz and Forster [6] at Harvard University made an ultra-narrow band multi-channel search with 26 m radio telescope. Both these

searches were near the hydrogen line and those were all sky search instead of targeted search. A summary of all these three early attempts are presented in Table 1.

Table 1: Summary of early attempts of interstellar communication.

Type of search	Location	Scientist	Year	Antenna Diameter
Targeted search	N.R.A.O., Green Bank	Frank Drake	1960	26 m
All sky search	Ohio State University	Robert Dixon	1985	110 m
All sky search	Harvard University	Paul Horowitz & J. Forster	1985	26 m

6. PRESENT APPROACH

If we consider a hypothetical interstellar communication link to an extraterrestrial civilization which transmits 10^6 W with a 110 m dish, then the effective apertures of the antennas,

$$A_{et} = A_{er} = 0.5\pi \,(\text{Diameter of the dish})^2 = 0.5\pi \,(110)^2 = 4450 \,\text{m}^2$$

Taking, $S/N = 1$, $k = 1.38 \times 10^{-23} \,\text{JK}^{-1}$, $T_{sys} = 10 \,\text{K}$ and $B_r = 0.1 \,\text{Hz}$, we get using Eq. (1),

$$d = \frac{127.9 \times 10^{16}}{\lambda} \tag{2}$$

For the selection of frequency, we choose the first and second water holes with a view to consider a more realistic situation. The wavelengths, distances of the radio link and the corresponding frequencies for the two holes are given in Table 2.

Table 2: Water holes and the corresponding frequencies.

Feature	Wavelength	Distance of radio link (in m); using Eq. (2)	Frequency (in GHz)
First water-hole	18 cm	710×10^{16}	1.67
	25 cm	609×10^{16}	1.43
Second water-hole	2 mm	63950×10^{16}	150.00
	14 mm	9136×10^{16}	21.43

It appears from the table that the two water holes have their frequencies 1.43 to 1.67 GHz and 21.43 to 150 GHz respectively. In the calculation of d, we have assumed that the receiving antenna on the earth has a dish of identical diameter and the aperture efficiency of both the antennas is 50% with the system temperature at the earth station as 10 K with a bandwidth of 0.1 Hz. If we put $1 \,\text{Ly} = 10^{16} \,\text{m}$, then the corresponding distances become 710 Ly, 609 Ly, 63950 Ly and 9136 Ly respectively. If a uniform star density of 0.01 per Ly^3 is accepted, for simplicity of calculation, then determining the volume for different radius and the associated directivity of the antennas (which may be considered the same as the number of objects an antenna can resolve), the number of stars in the antenna beam at any time regardless of the antenna pointing direction can be calculated. The numbers of stars for different wavelengths are presented in the Table 3.

Table 3: The number of stars for different radius and the associated wavelengths.

Wavelength, λ (in m)	Radius (in Ly)	Number of stars = Volume/Directivity
0.18	710	8
0.21	609	7
0.002	63950	734
0.014	9136	10486

The table indicates that instead of pointing the antenna at known stars (known as targeted search), an all sky search would be more appropriate. In our approach, the transmitter power, antenna size, system temperature etc are all typical and considered on the basis of our present ground-based technology. We suggest searches giving emphasis to the wavelengths for both the water-holes for detecting and identifying any signal and also subsequently for its successful interpretation.

With so much uncertainty for handling various parameters as to what signal characteristics to expect, search-strategy planning becomes a very complex one. Philosophy, science, literature and religion all contain strong statements of belief about the possibility of intelligent lifeforms on other planets. In fact, it is a profoundly interesting question as it raises issues about the origins of life as well as our place in the physical and spiritual universe [7]. At the same time, it has inspired a new discipline named as exotheology whose purpose is to examine the theological and mythological issues as they pertain to extraterrestrial intelligence. Its existence is still largely hypothetical and the range of speculative forms of extraterrestrial life is considered from sapient beings to life at the scale of bacteria. The hypothesis regarding the origin of extraterrestrial life is mainly based on two aspects. One proposes that its emergence occurred quite independently at different places in the universe. The second hypothesis suggests that life emerges in one location and then spreads between habitable planets. Any study for the search of extraterrestrial life in the present scenario, including the development of a new technique, as suggested in this report, may therefore be considered as highly significant.

REFERENCES

1. Bhattacharya, A. B., S. Joardar, and R. Bhattacharya, *Astronomy and Astrophysics*, 44–59, Infinity Science Press, Massachusetts, USA, 2008.
2. Kraus, J. D., *Radio Astronomy, Cygnus-quasar Books*, 12-0–12-6, USA, 1986.
3. Drake, F. D. and G. Helou, "The optimum frequencies for interstellar communication as influenced by minimum bandwidths," N.A.I.C. (Cornell), Rept. 76, 1977.
4. Drake, F. D., "A reminiscence of ozma," *Cosmic Search*, Vol. 1, 10–15, 1979.
5. Dixon, R. S., "A search strategy for finding extraterrestrial radio beacons," *Icarus*, Vol. 20, 187–199, 1973.
6. Horowitz, P. and J. Forster, "Ultra-narrowband SETI," *Proc. I.A.U. Symp.*, No. 112, Reidel, 1985.
7. Troxel, D. K., "Intelligent life in the universe and exotheology in Christianity and Bahi's writings," file: //I: \universe, 1–29, 2008.

Analysis of Collided Signal Waveform on the Long Transmission Line of UART-CSMA/CD Control Network

Chuzo Ninagawa[1] and Yasumitsu Miyazaki[2]

[1]Mitsubishi Heavy Industries, Ltd., Japan
[2]Aichi University of Technology, Japan

Abstract— In this paper, waveforms of collided digital signal bits are analyzed by superposition of step responses of a distributed constant circuit for a transmission line. Our analysis indicates narrow voltage spike on the leading edge of bit pulses in collided waveforms. The width and peak of the spike calculated by our analysis were matched with those of experiments. Since the spike might cause signal collision detection failure and cause deterioration of communication performance, our analysis is effective to estimate the allowable maximum length and/or bit rate of a transmission line for control communication network.

1. INTRODUCTION

UART-CSMA/CD (Universal Asynchronous Receiver Transmitter-Carrier Sense Multiple Access with Collision Detection) is a type of communication protocol for building facility control networks [1]. This protocol is based on detection of signal collision that is simultaneous transmission from more than two network nodes. However, in the case of an extremely long transmission line, collision detection becomes difficult because collided signal waveforms include severe distortion due to propagation delay and line terminal reflection. This distortion of waveform might cause collision detection failure resulting in deterioration of communication performance.

So far, many research works on analysis and simulation for signal waveform distortion on the long transmission line have been carried out [2–5]. However, analytical methods for calculating waveforms of signal collision on the distributed constant transmission line have not been published. Although the collided waveforms are easily obtained using circuit simulation tools, analytical methods are still valuable to understand the mechanism of distortion in collided waveforms on the long transmission line.

In this paper, waveforms of collided digital bits are analyzed with superposition of step responses of a distributed constant circuit for transmission line. The waveforms calculated by our analysis have been compared with those of experiments. Our analysis has predicted a narrow spike on each leading edge of bit pulses of collided waveforms. In sever cases of combination of transmission length and bit width, collision detection failure might occur. We have investigated the relationship between communication throughput and collision detection failure rate.
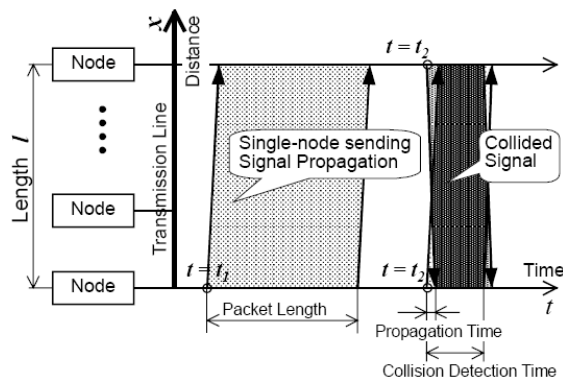

Figure 1: Signal propagation on the transmission line.

2. SIGNAL COLLISION OF UART-CSMA/CD

In a typical control network system for building facilities, there are hundreds of controllers, which are called nodes from a view point of network. The total line length often reaches almost 1000 m

throughout the building. Since control networks have to be installed in the most inexpensive way, any network terminators are not installed. It is well known that impedance mismatching at the terminals adversely affects on signal waveform in the case of a significantly long line.

Figure 1 shows UART-CSMA/CD signal propagation on the transmission line using time and distance axes. In the case of CSMA/CD, each node can start sending a signal at arbitrary timing. If more than two nodes happen to start sending at the exactly same timing, a signal collision occurs. At the moment when sending nodes detect the collision, the nodes immediately stop sending, wait for a random period, and restart sending the previous signal again. In the case of UART-CSMA/CD, the collision detection is carried out by comparing the logical sending signal and the actual receiving signal from the line. Therefore, in the case of a long transmission line, it is difficult to distinguish the collision because actual waveforms are also distorted in the case of non-collision sending.

Figure 2 shows our model of signal collision on the transmission line. This is a case that Node1 and Node2 located at the both terminals ($x = 0$ and $x = l$), and they start sending a series of two bits with opposite logical levels of "HL" and "LH". Here, "H" and "L" mean logical high and low levels, respectively. At the location of Node1 ($x = 0$), the signal $v_{HL}^{(1)}(0, t)$ sent by Node1 and the signal $v_{LH}^{(2)}(0, t)$ propagated from the Node2 coexist. In the case of a long line, propagation time can not be neglected. Even if Node2 simultaneously starts sending the signal $v_{LH}^{(2)}(l, t)$, receiving signal $v_{LH}^{(2)}(0, t)$ by Node1 at $x = 0$ will be shifted on the time axis by propagation time.
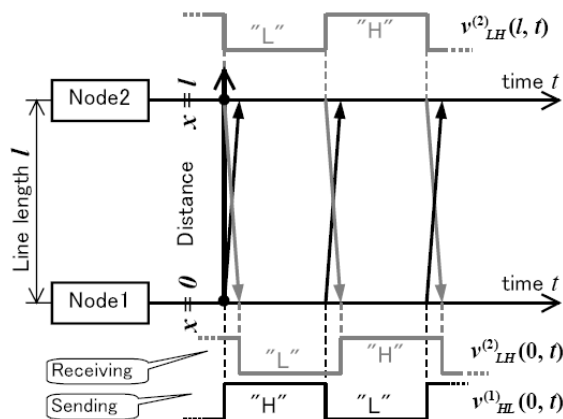

Figure 2: Signal collision of opposite logical bits propagating on the transmission line.

Here, an important point is that each of $v_{HL}^{(1)}(0, t)$ and $v_{LH}^{(2)}(0, t)$ is distorted by the terminal reflections. Consequently, the waveform of the collision is not a simple superposition of logical levels with the propagation delay, but it becomes a complicated aggregation of distorted waveforms. Therefore, analysis of distortion elements is important to investigate affection of the line length and termination impedance for practical system design of a control network.

3. SIGNAL WAVEFORM ANALYSIS

3.1. Analysis Model

In order to calculate the superposition of terminal reflections on the transmission line, a Laplace transformation method for transient analysis of a distributed constant circuit is used. For a step response of signal voltage on a lossless transmission line, voltage $v_{ST}(x, t)$ at the position x and time t is given by the following equation [4, 6],

$$v_{ST}(x, t) = [E_1 Z_0 / (Z_1 + Z_0)]\{u[t - x/w] + \Gamma_2 u[t - (2T_{OW} - x/w)] + \Gamma_1 \Gamma_2 u[t - (2T_{OW} + x/w)]$$
$$+ \Gamma_1 \Gamma_2^2 u[t - (4T_{OW} - x/w)] + \Gamma_1^2 \Gamma_2^2 u[t - (4T_{OW} + x/w)] + \ldots\}, \tag{1}$$

where $u[t]$ is a unit step function, E_1 is signal source voltage, $Z_0 = (L_0/C_0)^{1/2}$ is characteristic impedance, L_0 and C_0 are characteristic inductance and capacitance of unit length, Z_1 is source impedance, Z_2 is load impedance, $T_{OW} = l/w = l(L_0 C_0)^{1/2}$ is one way trip time, l and w are line length and propagation velocity, $\Gamma_1 = (Z_1 - Z_0)/(Z_1 + Z_0)$ is source reflective coefficient, and

$\Gamma_2 = (Z_2 - Z_0)/(Z_2 + Z_0)$ is load reflective coefficient. Both Z_1 and Z_2 are assumed to be pure resistance.

The sending voltage $v_{HL}^{(1)}(0,t)$ for a signal "HL" transmitted by Node1 located at $x = 0$ is defined using the step response Equation (1) as follows,

$$v_{HL}^{(1)}(0,t) = v_{ST}(0,t) + \overline{v_{ST}(0, t - T_b)} = v_{ST}(0,t) - v_{ST}(0,t)u[t - T_b], \tag{2}$$

where the bar in the above equation means the complementary logical bit "L". Similarly, voltage $v_{LH}^{(2)}(0,t)$ received by Node1 for the signal "LH" sent by the Node2 located at $x = l$ is also defined as follows.

$$v_{LH}^{(2)}(0,t) = \overline{v_{ST}(l,t)} + v_{ST}(l, t - T_b) = -v_{ST}(l,t) + v_{ST}(l,t)u[t - T_b] \tag{3}$$

Here, in the case of signal collision, both nodes are in the signal transmission mode. Therefore, defining as $E \equiv E_1 = E_2$, $Z \equiv Z_1 = Z_2$, $\Gamma \equiv \Gamma_1 = \Gamma_2$, the step responses $v_{ST}(0,t)$ and $v_{ST}(l,t)$ become the following equation.

$$v_{ST}(0,t) = [EZ_0/(Z + Z_0)]\{u[t] + \Gamma u[t - 2T_{OW}] + \Gamma^2 u[t - 2T_{OW}] + \ldots\}, \tag{4}$$

$$v_{ST}(l,t) = [EZ_0/(Z + Z_0)]\{u[t - l/w] + \Gamma u[t - (2T_{OW} - l/w)] + \Gamma^2 u[t - (2T_{OW} + l/w)] + \ldots\}. \tag{5}$$

Therefore, a collided waveform expression of signal "HL" from Node1 and signal "LH" form Node2 at the location of Node1, i.e., $x = 0$, is given by superposition of the step responses using the Equations (2)–(5) as the following.

$$v_C(0,t) = v_{HL}^{(1)}(0,t) + v_{LH}^{(2)}(0,t) = \{v_{ST}(0,t) - v_{ST}(l,t)\}\{1 + u[t - T_b]\} \tag{6}$$

Thus, we have obtained an analytical expression of signal voltage waveform of collided bit series of opposite logical levels of "HL" and "LH" as superposition of step responses.

3.2. Numerical Calculation and Experiment

We have calculated collided signal waveforms using typical parameters for an actual UART-CSMA/ CD control network system. The parameters were assumed as follows. The bit rate was $R_b = 9.6\,\text{kbps}$, i.e., the bit pulse width was $T_b = 104\,\mu\text{s}$. The transmission line length was $l = 100\,\text{m}$ or $500\,\text{m}$. The signal source voltage $E = 3.0\,\text{V}$ and impedance $Z = 28\,\Omega$ were assumed for parameters of a typical RS-485 driver/receiver IC. The type of the cable was assumed to be a polyvinyl chloride (PVC) twisted cable with $2\,\text{mm}^2$ (AWG14) wires. The values of cable parameters were measured as $L_0 = 0.59\,\text{mH/km}$ and $C_0 = 0.12\,\mu\text{F/km}$ resulting in the characteristic line impedance $Z_0 = (L_0/C_0)^{1/2} = 70\,\Omega/\text{km}$.

Figure 3 shows an example of numerical calculation of collided signal waveform $v_C(0,t)$ using our analytical Equations (2), (3), and (6) with the above-mentioned parameters. The original waveforms of sending and receiving signals, $v_{HL}^{(1)}(0,t)$ and $v_{LH}^{(2)}(0,t)$, are also superimposed on the same figure. The distortions due to the terminal reflections appear at the leading edge of each bit pulse. In this example, the large portion of each bit was stable at the "H" or "L" voltage
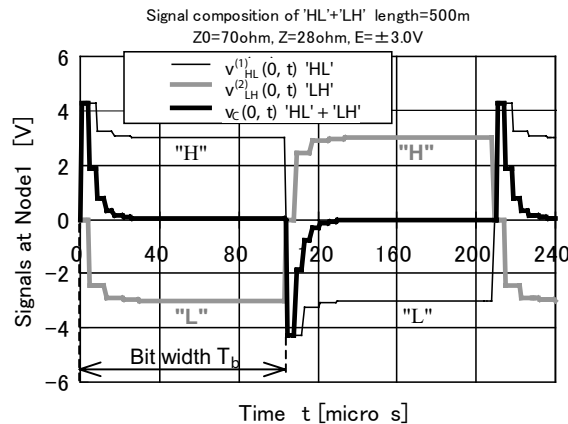

Figure 3: Sending, receiving, and superposition signal waveforms at Node1.

level. Consequently, voltage level in the almost all portion of the collided waveform became "H" + "L" $= 0$ V. However, a voltage spike appeared at the leading edge of each bit pulse of the collided waveform.

We have confirmed the above-mentioned results of our analysis by comparing with those of experiments. Figs. 4 and 5 show the results of waveforms calculated using our equations and those observed by an oscilloscope. Fig. 4 is a case of line length of $l = 100$ m. As predicted by our analysis, the similar spikes appeared in the results of experiment observation. The peak voltage of the spikes was 1.9 V for the calculation and 2.4 V for the observation, respectively. Fig. 5 is a case of the line length of 500 m. The peak voltages were 4.3 V for the analysis, and 4.0 V for the observation, respectively. Although there were small differences in the peak values, overall shapes of the collided waveforms calculated using our analytical equations have matched with those of experiments. Thus, our analytical method of calculating collided signal waveforms on the long transmission line has been verified.
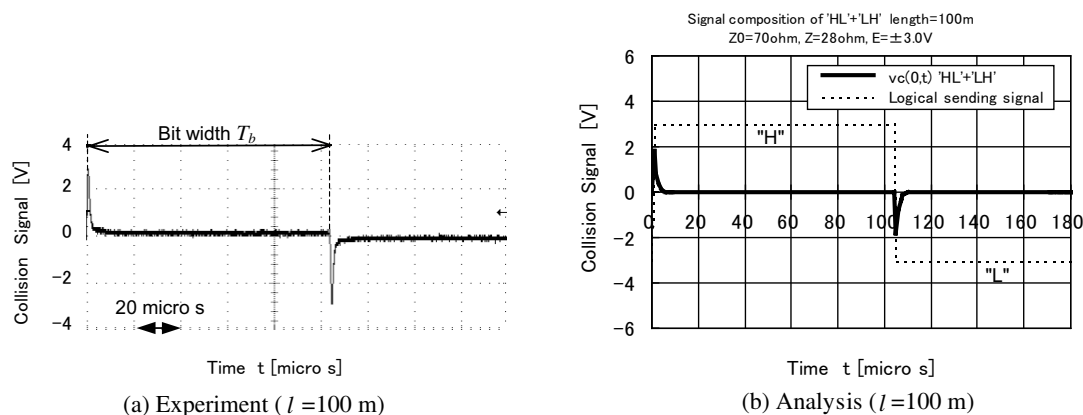

(a) Experiment (l =100 m)   (b) Analysis (l =100 m)

Figure 4: Collision signal waveform (Sort line).

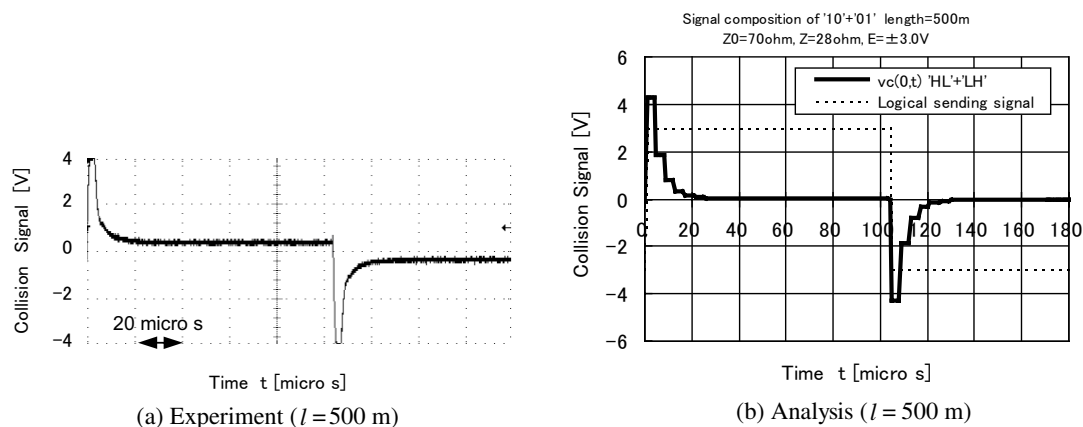(a) Experiment (l =500 m)   (b) Analysis (l = 500 m)

Figure 5: Collision signal waveform (Long line).

4. COMMUNICATION PERFORMANCE

Theoretical average throughput $S(\sigma)$ of the UART-CSMA/CD protocol is given by the following expression for an offered load of σ, where σ is the number of signal-sending-trials per discrete time unit τ for each node [1, 7],

$$S(\sigma) = \frac{\sum\limits_{i=0}^{M} \pi_i \cdot P_s(i) \cdot T}{\sum\limits_{i=0}^{M} \pi_i \cdot \{1/(1 - \delta_i) + T_f + P_s(i) \cdot T + [1 - P_s(i)] \cdot \gamma\}}. \tag{7}$$

Here, M is the total number of nodes, π_i is the probability that the number of nodes with a signal is i, $P_s(i)$ is the probability that only one node starts sending at a discrete time slot, T is the signal packet length in the unit of discrete time τ, $1/(1-\delta_i)$ is the average idle time before sending, and T_f is the length of the constant spacing time after sending. Here, γ is the sending duration time in the case of successful collision detection, and γ is smaller than the packet length T because the node stops sending at the detection moment.

However, in the case of collision detection failure, the duration time becomes the full packet length T because the nodes do not stop sending. If the rate of collision detection failure is P_F, the average sending duration time becomes $(1-P_F)\cdot\gamma+P_F\cdot T$ instead of γ. Therefore, the average throughput $S'(\sigma)$ for P_F is given by the following equation.

$$S'(\sigma) = \frac{\sum_{i=0}^{M} \pi_i \cdot P_s(i) \cdot T}{\sum_{i=0}^{M} \pi_i \cdot \{1/(1-\delta_i) + T_f + P_s(i) \cdot T + [1 - P_s(i)] \cdot [(1 - P_F) \cdot \gamma + P_F \cdot T]\}}. \tag{8}$$

Since $(1-P_F)\cdot\gamma+P_F\cdot T$ is larger than γ, $S'(\sigma)$ becomes smaller than $S(\sigma)$. In other words, the throughput reduces along with collision detection failure rate P_F.

Figure 6 shows throughput curves calculated by the Equation (8) with 0%, 50%, and 100% of P_F. The horizontal axis of the figure is the normalized offered load of the entire network $G = (\sigma/\tau) \cdot M \cdot T_p$, where T_p is the signal packet length in the unit of second. The vertical axis is the normalized throughput, i.e., the number of non-collided transmissions per packet length time. As shown in Fig. 6, the throughput $S(G)$ deteriorates along with P_F for range of G greater than about 0.7. At the maximum load of $G = 2$ in this case, $S(G)$ has reduced from $S(2) = 0.68$ for $P_F = 0\%$ to $S'(2) = 0.43$ for $P_F = 100\%$ by approximately 40% down. Thus, performance of UART-CSMA/CD control network is significantly affected by collision detection failure.

There is a possibility of collision detection failure due to the spikes in signal waveforms. It might be necessary to eliminate the spikes depending the scale of spike and bit width. One of the possible means for elimination of the spike is implementing a noise filter in the receiving circuit of a node as shown in Fig. 7. The filter will be based on low pass filter circuit that eliminates high-frequency components of the spike but passes low-frequency components of basic shape of original bit pulse. For designing this filter circuit parameters, our analytical method will contribute to case studies in the system design process.
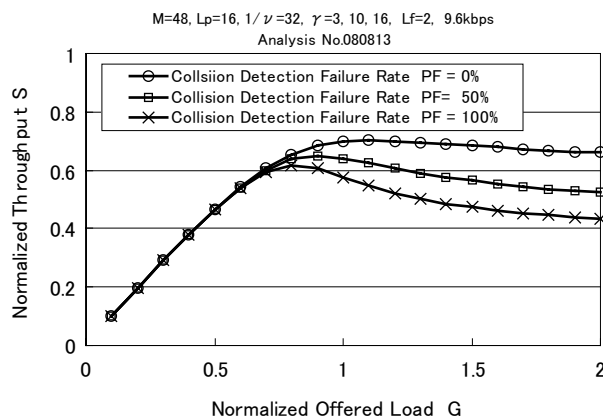

Figure 6: Influence of collision detection failure rate on throughput of UART-CSMA/CD.
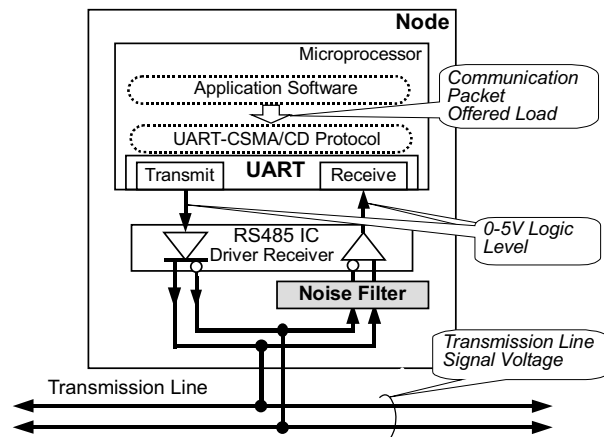

Figure 7: Filtering spike noise in receiving circuit of UART-CSMA/CD node.

5. CONCLUSION

In this research, we have proposed an analytical method of calculating collided signal waveforms on the long transmission line. Our analysis has successfully predicts existence of the spike voltage in collided waveforms. These spikes might cause collision detection failure in the UART-CSMA/CD communication protocol. This will cause deterioration of throughput of the CSMA/CD protocol.

The prediction of waveform of collided signal waveforms before actual system installation is important for avoiding the throughput deterioration. Our analytical method is effective to estimate the allowable combination of the line length, terminal impedance mismatching and bit rate of the transmission line.

ACKNOWLEDGMENT

REFERENCES

1. Ninagawa, C., K. Yokohama, F. Aoi, and Y. Miyazaki, "Performance analysis on air-conditioner control CSMA/CD protocol by embedded microcontroller UART collision detection," *IEE J. Trans. IA*, (in Japanese), Vol. 126, No. 5, 665–672, 2006.
2. Magnusson, P., G. Alexander, V. Tripathi, and A. Weisshaar, *Transmission Lines and Wave Propagation*, 4th Edition, CRC Press, Boca Raton, 2001.
3. Canright, R., "Simulating and controlling the effects of transmission line impedance mismatches," *IEEE 23rd Conf. on Design Automation*, 778–785, 1986.
4. Ninagawa, C., K. Yokohama, F. Aoi, H. Otake, and Y. Miyazaki, "Transmission characteristics of UART-CSMA/CD control network with one-chip microcontroller and RS-485 driver IC," *PIERS Proceedings*, 271–275, Tokyo, Aug. 2006.
5. Ninagawa, C., K. Yokohama, F. Aoi, and Y. Miyazaki, "Signal waveform distortion on terminatorless transmission line of UART-CSMA/CD control network," *PIERS Proceedings*, 1694–1698, Beijing, Mar. 2007.
6. Matick, M., "Transmission lines for digital and communication networks," An IEEE PRESS Classic Reissue, Picataway, NJ, 1995.
7. Tobagi, F. and V. Hunt, "Performance analysis of carrier sense multiple access with collision detection," *Computer Networks*, Vol. 4, 245-259, 1980.

Recognition of Wavelength-multiplexed Labels with Acoustooptic Waveguide Circuit for Hierarchical Photonic Routing

Nobuo Goto[1] and Yasumitsu Miyazaki[2]

[1]The University of Tokushima, Japan
[2]Aichi University of Technology, Japan

Abstract— High-speed optical processing for packet routing can overcome bottleneck in large-capacity photonic networks. We have studied on collinear acousto-optic (AO) switches and applications to optical label recognition. Since parallel combination of collinear AO switches can handle wavelength-division-multiplexed (WDM) optical pulses, recognition for optical labels encoded in spectral and time domains can be realized. In this report, we discuss recognition of layer-structure labels for hierarchical routing control. By employing WDM time-series labels including identifying bits, the labels corresponding to each layer can be recognized. The recognition characteristics are confirmed by computer simulation.

1. INTRODUCTION

Broadband photonic network has been extensively studied to realize high-speed and low-power consumption routing system. In optical label routing network, label processing at routers can be faster by using optical signal processing. Various types of optical label recognition and swapping systems have been reported [1, 2]. Hierarchical networking has also been studied for effective routing [3, 4].

We have investigated acousto-optic (AO) devices [5] and their applications to label processing circuit [6, 7]. In particular, we have reported circuits for optical labels encoded in time and spectral domains to use effectively the unique nature of collinear AO devices, that is, optical multiple wavelengths can be independently processed by frequency multiplexed surface acoustic waves (SAWs).

In this report, we consider a label recognition circuit with a new label structure for hierarchical routing. The principle of the proposed waveguide circuit is discussed with computer simulations.

2. HIERARCHICAL NETWORK AND ROUTING

Figure 1(a) shows a two-layer network as an example of hierarchical networks. By employing layer-structure labels, routing processing can be simpler and faster. We consider a code m for M-layer routing network as represented by

$$\boldsymbol{C}_{code}^{m} = (C_{L1}^{m}, C_{L2}^{m}, \ldots, C_{LM}^{m}), \tag{1}$$

where, C_{Li}^{m} represents a code in the ith layer. A label router for $M = 2$ is shown in Fig. 1(b). The extracted label is processed to find whether the 1st-layer label C_{L1}^{m} is matched with the label C_{L1}^{k} of the local network. If these labels are matched, the 2nd-layer label is examined to find its destination output port. If the 1st labels are not matched, this packet is forwarded to a border router by cutting through this router.

3. LABEL STRUCTURE AND RECOGNITION CIRCUIT

A proposed label structure is shown in Fig. 2(a). The routing code in each layer consists of N_t pulses with the interval of Δt_p. Each pulse has N_λ WDM components which are different among layers. The ith-layer code consists of $\lambda_{i,j}$, $(j = 1, \ldots, N_\lambda)$. In advance of the routing code, M identifying (ID) pulses having interval of $N_t \Delta t_p$ are placed. The whole label can be written as

$$\boldsymbol{C}_{label}^{m} = (C_{id}, C_{L1}^{m}, C_{L2}^{m}, \ldots, C_{LM}^{m}), \tag{2}$$

where, the ID pulse train is written as

$$C_{id} = (\boldsymbol{c}_{id,1}, 0, \ldots, 0, \boldsymbol{c}_{id,2}, 0, \ldots, 0, \boldsymbol{c}_{id,M}, 0, \ldots, 0) \tag{3}$$

with each ID pulse

$$\boldsymbol{c}_{id,n} = (1, 0, \ldots, 0)^{t} \quad n = 1, \ldots, M. \tag{4}$$

The WDM pulse train for routing codes is written for $M = 3$ as

$$
C_{code}^m = (\boldsymbol{c}_1^m,\ \boldsymbol{c}_2^m,\ \ldots,\ \boldsymbol{c}_{3N_t}^m) = \begin{bmatrix} \begin{matrix} c_{1,1,1}^m c_{1,2,1}^m c_{1,3,1}^m \\ c_{1,1,2}^m c_{1,2,2}^m c_{1,3,2}^m \\ c_{1,1,3}^m c_{1,2,3}^m c_{1,3,3}^m \end{matrix} & 0 & 0 \\ 0 & \begin{matrix} c_{2,1,1}^m c_{2,2,1}^m c_{2,3,1}^m \\ c_{2,1,2}^m c_{2,2,2}^m c_{2,3,2}^m \\ c_{2,1,3}^m c_{2,2,3}^m c_{2,3,3}^m \end{matrix} & 0 \\ 0 & 0 & \begin{matrix} c_{3,1,1}^m c_{3,2,1}^m c_{3,3,1}^m \\ c_{3,1,2}^m c_{3,2,2}^m c_{3,3,2}^m \\ c_{3,1,3}^m c_{3,2,3}^m c_{3,3,3}^m \end{matrix} \end{bmatrix}. \quad (5)
$$

Figure 1: (a) Hierarchical routing network and (b) optical router.

An integrated-optic processor consisting of parallel AO switches and delay lines is shown in Fig. 2(b). The incident label is divided into $N_t + 1$ pulse trains. The partial labels or the whole label to be matched in this device are represented by frequency-multiplexed SAWs. The label pulse trains are wavelength-selectively switched and the outputs are balanced detected with photodiodes (PDs). The electric output signals are electrically multiplied. The M matched output pulses correspond to label matching of each layer codes. Since the SAWs representing a code to be recognized do not need to be changed during recognition, the slow switching speed of AO devices does not restrict the processing speed for label recognition.

Figure 2: (a) Structure of the hierarchical routing label and (b) label recognition AO processor.

4. PRINCIPLE OF CODE RECOGNITION

We consider the electric field of the incident optical pulse train representing code m as given by

$$E^m(t) = E_0 \sum_{n=1}^{M} c_{id,n} g_{in,0}(t-t_0) \exp[j\omega_1(t-t_0)] + E_0 \sum_{n=1}^{M}\sum_{l=1}^{N_t}\sum_{i=1}^{N_\lambda} c_{n,l,i}^m g_{in,0}(t-t_1) \exp[j\omega_{n,i}(t-t_1)], \quad (6)$$

where, $g_{in,0}(t)$ denotes the pulse envelope of a single pulse, ω_i is the angular frequency of wavelength λ_i, and t_0 and t_1 are defined by

$$t_0 = (m-1)N_t\Delta t_p, \quad t_1 = [(M+n-1)N_t + l - 1]\Delta t_p. \quad (7)$$

An incident pulse train into the l'th AO switch after passing the divider and a delay waveguide is written as

$$E_{l'}^m(t) = \frac{1}{\sqrt{N_t+1}} E^m(t - (N_t - l')\Delta t_p). \quad (8)$$

When a code to be recognized is code k, the SAW strain in the l'th AO switch is written as

$$S_{l'}^k(t, z) = \sum_{n=1}^{M}\sum_{i=1}^{N_\lambda} c_{n,l',i}^k s_0 \cos(\Omega_{n,i}t - K_{n,i}z), \quad (9)$$

where, s_0 is strain components related to AO switching, $\Omega_{n,i}$ and $K_{n,i}$ are the angular frequency and the propagation constant of the SAWs. The switched optical signal has the output

$$E_{out1,l'}^{m,k}(t) = \frac{E_0}{\sqrt{N_t+1}} \left\{ \sum_{n=1}^{M} c_{id,n} c_{1,l',1}^k g_{out,0}(t-t_3) \exp[j(\omega_{n,1} - \Omega_{n,1})(t-t_3)] \right.$$
$$\left. + \sum_{n=1}^{M}\sum_{l=1}^{N_t}\sum_{i=1}^{N_\lambda} c_{n,l,i}^m c_{n,l',i}^k g_{out,0}(t-t_2) \exp[j(\omega_{n,i} - \Omega_{n,i})(t-t_2)] \right\}, \quad (10)$$

where, t_2 and t_3 are defined by

$$t_2 = T_1 + (N_t - l')\Delta t_p + l_{SW} N(n,i)/c, \quad t_3 = (nN_t - l')\Delta t_p, \quad (11)$$

where, c is the optical velocity, $N(n,i)$ is the effective index for the guided wave at wavelength $\lambda_{n,i} = 2\pi c/\omega_{n,i}$, $g_{out,0}(t)$ is the envelope of the switched pulse, and l_{SW} is the AO interaction length. A common phase term that changes through propagation along the waveguides is omitted for simplicity. The optical field exitting from the unswitched port is given by

$$E_{out2,l'}^{m,k}(t) = \frac{E_0}{\sqrt{N_t+1}} \times \left\{ \sum_{n=1}^{M} c_{id,n} \bar{c}_{1,l',1}^k g_{out,0}(t-t_3) \exp[j(\omega_{n,1} - \Omega_{n,1})(t-t_3)] \right.$$
$$+ c_{id,n} c_{1,l',1}^k g_{out,0}^{res}(t-t_3) \exp[j\omega_{n,1}(t-t_3)]$$
$$\left. + \sum_{n=1}^{M}\sum_{l=1}^{N_t}\sum_{i=1}^{N_\lambda} \left[c_{n,l,i}^m \bar{c}_{n,l',i}^k g_{out,0}(t-t_2) + c_{n,l,i}^m c_{n,l',i}^k g_{out,0}^{res}(t-t_1) \right] \exp[j(\omega_{n,i} - \Omega_{n,i})(t-t_2)] \right\}, (12)$$

where, $g_{out,0}^{res}(t)$ denotes the residual component due to incompleteness of switching, and $\bar{c}_{n,l',i}^k = 1 - c_{n,l',i}^k$.

When we assume the AO switching to be ideal as given by

$$g_{out,0} \simeq g_{in,0}, \quad g_{out,0}^{res} \simeq 0, \quad (13)$$

the electric differential current from the PDs is written as

$$I_{l'}^{m,k}(t) \simeq \frac{I_0}{N_t+1} \left\{ \left[\left(c_{id,n} c_{1,l',1}^k \right)^2 - \left(c_{id,n} \bar{c}_{1,l',1}^k \right)^2 \right] g_{in,0}^2(t-t_3) \right.$$
$$\left. + \sum_{n=1}^{M}\sum_{l=1}^{N_t}\sum_{i=1}^{N_\lambda} \left[\left(c_{n,l,i}^m c_{n,l',i}^k \right)^2 - \left(c_{n,l,i}^m \bar{c}_{n,l',i}^k \right)^2 \right] g_{in,0}^2(t-t_2) \right\}. \quad (14)$$

By multiplying output currents from N_t AO switches with the electrical multiplier, we obtain

$$I^{m,k}(t) = \prod_{l'=1}^{N_t} I_{l'}^{m,k}(t) = \left(\frac{I_0}{N_t+1}\right)^{N_t} \prod_{l'=1}^{N_t} \sum_{n=1}^{M} \sum_{l=1}^{N_t} \sum_{i=1}^{N_\lambda} \left[(c_{n,l,i}^m c_{n,l',i}^k)^2 - (c_{n,l,i}^m \bar{c}_{n,l',i}^k)^2\right] g_{in,0}^2(t-t_2). \quad (15)$$

To evaluate correlation in each-layer code, ID bit pulses are used. The pulse train passed through only the delay waveguide, which is the lowest waveguide in the AO processor, is converted to electric current with the PD as

$$I_{out,ID}^m(t) = \frac{I_0}{|E_0|^2}\overline{|E_{out,ID}^m|^2} = \frac{1}{N_t+1}\left\{\sum_{n=1}^{M} c_{id,n}^2 g_{in,0}^2(t-t_4) + \sum_{n=1}^{M}\sum_{l=1}^{N_t}\sum_{i=1}^{N_\lambda}(c_{n,l,i}^m)^2\right.$$

$$\left. \times g_{in,0}^2(t-t_1 - [(M+2)N_t - 1]\Delta t_p - l_{SW}N_1/c)\right\}. \quad (16)$$

Here, we assume $t_2 = t_4$ and $N(n,i) = N_1$ resulting in $l = l'$, and t_2 is rewritten as

$$t_2 = (M+n)N_t - 1 + l_{SW}N_1/c. \quad (17)$$

By multiplying the current from the AO switches and the current for identifying pulse train, we obtain

$$I_{out}^{m,k}(t) = I^{n,k}(t) \times I_{out,ID}^m(t)$$

$$= \left(\frac{1}{N_t+1}\right)^{N_t+1}\left\{\sum_{n=1}^{M}\sum_{l=1}^{N_t}\sum_{i=1}^{N_\lambda} c_{id,n}^2 \left[(c_{n,l,i}^m c_{n,l,i}^k)^2 - (c_{n,l,i}^m \bar{c}_{n,l,i}^k)^2\right] g_{in,0}^4\right.$$

$$\left.(t - (M+n)N_t - 1 + l_{SW}N_1/c)\right\}. \quad (18)$$

5. COMPUTER SIMULATION OF CODE RECOGNITION

In the previous analysis, the ideal switching characteristics given by Eq. (13) were assumed. In this section, we evaluate the code recognition taking the swiching response of AO switches into consideration. When the bit rate of the optical pulses becomes higher, broader frequency spectrum is accompanied in the optical signal. The filtering characteristics of the AO switches affect the switching response for high bit-rate pulses.

We assume that the incident pulse train consists of 9 pulses ($M = 3$, $N_t = 3$) with bandwidth of 160 GHz, pulse period of 12 ps and pulse width of 5.9 ps. Labels in each layer consist of WDM pulses of $N_\lambda = 3$. We also assume the interaction length of the AO switches to be $l_{SW} = 16$ mm.

We consider OOK orthogonal codes as each WDM pulse. For $N_\lambda = 3$, we employ

$$c_i = \{a_1, a_2, a_3\} = \left\{\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}\right\}. \quad (19)$$
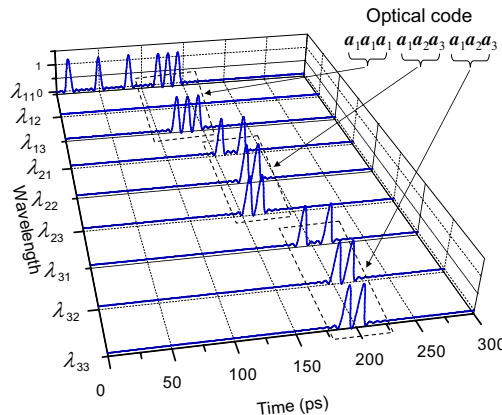

Figure 3: Wavelength components of electric fields of optical incident pulse train for $\boldsymbol{C}^m = (\boldsymbol{a}_1, \boldsymbol{a}_1, \boldsymbol{a}_1, \boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3, \boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3)$.

As an example of optical incident code m, we consider $\boldsymbol{C}^m = (\boldsymbol{a}_1, \boldsymbol{a}_1, \boldsymbol{a}_1, \boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3, \boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3)$. Fig. 3 shows the wavelength components of the label consisting of the code m and ID pulses at wavelength $\lambda_{1,1}$.

When the AO processor is set to recognize code k of $\boldsymbol{C}^k = (\boldsymbol{a}_1, \boldsymbol{a}_1, \boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_2, \boldsymbol{a}_2, \boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3)$ with SAWs, the output appears at the first and the third ID pulses as shown in Fig. 4(a). The second layer code is unmatched in this case. If all the three layer codes are matched, three pulses are obtain.

Next, we consider code matching for a single layer code. To recongnize the second layer code, the AO processor is assumed to be set for $\boldsymbol{C}^k = (- \; - \; - \; \boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3 \; - \; - \; -)$. The output current has a positive pulse at the second ID pulse and negative pulses at other ID pulses as shown in Fig. 4(b).

When the second layer code is unmatched by setting the AO processor with $\boldsymbol{C}^k = (- \; - \; - \; \boldsymbol{a}_2, \boldsymbol{a}_2, \boldsymbol{a}_2 \; - \; - \; -)$, the output current is obtained as shown in Fig. 4(c).
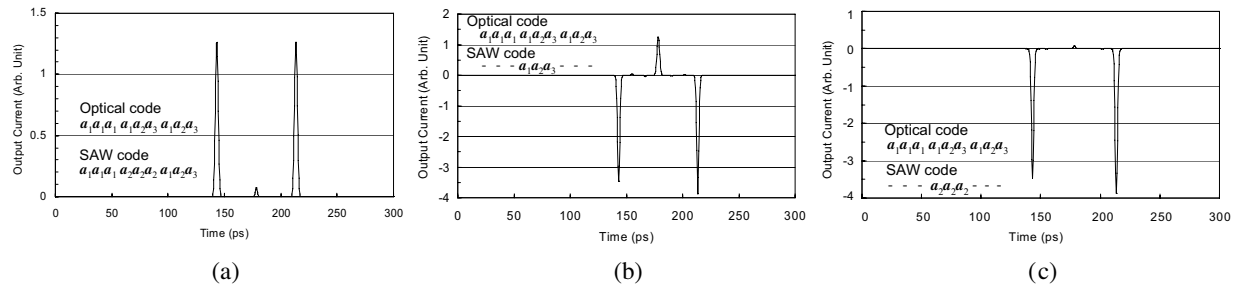
Figure 4: Output current from multiplier for SAWs representing (a) $\boldsymbol{C}^k = (\boldsymbol{a}_1, \boldsymbol{a}_1, \boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_2, \boldsymbol{a}_2, \boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3)$, (b) $\boldsymbol{C}^k = (- \; - \; - \; \boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3 \; - \; - \; -)$, and (c) $\boldsymbol{C}^k = (- \; - \; - \; \boldsymbol{a}_2, \boldsymbol{a}_2, \boldsymbol{a}_2 \; - \; - \; -)$.

6. CONCLUSION

We discussed optical label matching using an integrated-optic device and an electrical multiplier. By introducing WDM address bits and ID bits, partial label matching is performed. We will further investigate the layer-structure routing.

REFERENCES

1. Blumenthal, D. J., B. E. Olsson, G. Rossi, T. E. Dimmick, L. Rau, M. Masanovic, O. Lavrova, R. Doshi, O. Jerphagnon, J. E. Bowers, V. Kaman, L. A. Coldren, and J. Barton, "All-optical label swapping networks and technologies," *J. Lightwave Technol.*, Vol. 18, No. 12, 2058–2075, 2000.
2. Kitayama, K., N. Wada, and H. Sotobayashi, "Architectural considerations for photonic IP router based upon optical code correlation," *J. Lightwave Technol.*, Vol. 18, No. 12, 1834–1844, 2000.
3. Seddighian, P., S. Ayotte, J. B. Rosas-Fernandez, J. Penon, L. A. Rusch, and S. LaRochelle, "Label stacking in photonic packet-switched networks with spectral amplitude code labels," *J. Lightwave Technol.*, Vol. 25, No. 2, 463–471, Feb. 2007.
4. Yamanaka, N., K. Shiomoto, and E. Oki, *GMPLS Technologies — Broadband Backbone Networks and Systems*, CRC Press, Boca Raton, 2006.
5. Kondo, T., Y. Miyazaki, and Y. Akao, "Optical tunable switched directional couplers consisting of two thin-film waveguides using surface acoustic waves," *Jpn. J. Appl. Phys.*, Vol. 17, No. 7, 1231–1243, 1978.
6. Goto, N. and Y. Miyazaki, "Recognition of wavelength-division-multiplexed time-series optical coded labels using collinear acoustooptic devices without time gating for photonic routing," *Jpn. J. Appl. Phys.*, Vol. 46, No. 7B, 4602–4607, 2007.
7. Goto, N. and Y. Miyazaki, "Recognition characteristics of layered code for optical time-series wavelength-division-multiplexed labels using collinear acoustooptic switch arrays," *Jpn. J. Appl. Phys.*, Vol. 47, No. 5, 3974–3979, May 2008.

Design and Development of A FMCW Ground Based Imaging Radar System

Y. K. Chan, **C. Y. Ang**, **V. C. Koo**, **and C. S. Gan**

Multimedia University, Malaysia

Abstract— This paper describes the design and development of a ground based Frequency Modulated Continuous Wave (FMCW) radar system in Multimedia Univeristy (MMU), Malaysia. In this project, a ground-based fully polarometric, C-band, high bandwidth linear FM-CW and real time imaging radar system is to be designed and constructed. The system should have the capability to measure the complex scattering matrices of distributed targets using FMCW system and obtained fully polarimetric signals which can be used to provide more accurate identification and classification of the geophysical media.

The purposed system hardware consists of four major sections: an antenna system, a radio frequency (RF) subsystem, an intermediate frequency (IF) electronic, and a data acquisition unit (DAU). The RF section is constructed in-house from several RF components, which include voltage-controlled oscillator, high power amplifier, directional coupler, RF switches, band-pass filter, isolators, and mixers. In RF section, the received signal is mixed with a portion of the transmitted signal to produce low frequency IF signals. The IF signals are pre-processed in IF section, before they are digitised in DAU. A mobile personnel computer with analog to digital converter card is used to store the measurement data and process the data on real time basics.

In paper, the high level design will be discussed and detail design parameters will be presented. It followed by radar electronics design, which outlined the detail in radar transmitter and receiver.

1. INTRODUCTION

When an earth terrain is illuminated by an electromagnetic wave, the characteristics of the scattered wave are found to be related to the physical properties of the terrain. The wave received by a radar, commonly known as the backscattered signal, is therefore the information carrier from which the dielectric and geometrical properties of the terrain may be retrieved. In recent years, polarimetric technique has proven to be a promising tool for geophysical remote sensing [1–5]. Fully polarimetric radar signals carry additional phase information of the backscatterer and thereby provide more accurate identification and classification of terrain types.

A device which measures the full polarisation response of the scattered wave is called a polarimeter. Polarimetric radar systems differ from conventional radar system in that they are capable of measuring the complete scattering matrix of the remotely sensed media. The complete scattering matrix consists of four vector (amplitude and phase) quantities that may be determined from four different polarization combinations of the transmitting and receiving antennas. Typically, horizontally and vertically polarized antennas are employed to achieve the necessary polarisation diversity. However, any two orthogonally polarised antennas could be used [6]. By having the knowledge of the complete scattering matrix, it is possible to calculate the backscattered signal for any given combination of the transmitting and receiving antennas. This process is called the polarisation synthesis, which is an important technique used in terrain classification [1, 3].

Radar polarimetry deals with full polarimetric information of scattered waves from a target. It has been attracting attention in many application areas such as observation of the earth [1, 2], surveillance system for disaster, and highly advanced radar sensing. Therefore a polarimetric FMCW radar system is purposed to realize classification of various targets.

2. DESIGN CONSIDERATION

The primary goal of this project is to design and development of a ground based high bandwidth linear FMCW, fully polarometric and real time imaging radar system. High-level system design and subsystem level requirements have been carefully considered. High-level design consideration include:

2.1. Operating Frequency and Polarization

For remote sensing applications, frequency range from 1G to 30G Hz is normally used. In the 1 GHz to 10 GHz range, the transmissivity through air approaches 100%. Thus, a radar operating

in this frequency range is always able to image the earth's surface independent of the cloud cover or precipitation. Our system is designed to operate at C-band (5.3 GHz or 0.057 cm wavelength), which is within the allowable spectrum defined by International Telecommunication Union (ITU) for Earth Exploration Satellite System (EESS) [7].

Full polarization is chosen because a fully polarimetric radar signals carry additional phase information of the backscatterer and thereby provide more accurate identification and classification of target.

2.2. Modulation and Measurement Range

In radar remote sensing, there are two widely used configurations, i.e pulse and FM-CW radar. The FM-CW configuration is employed due to its simplicity in design and construction. In addition, by using the FM-CW arrangement, it is possible to obtain large independent samples within a single illumination footprint — which is the key to improve the precision of the measurement. This radar system will be used for near range measurement which range from 20 m to 100 m.

2.3. Operation Platform and Antenna

A simple ground based radar system is to be designed in this project. Thus a simple mobile unit utilized a trolley will be employed in the field measurement. Typical radar system antenna has the gain of 17 dB to 28 dB. Four 25 dBi gain horn antenna with single polarization will be used in this system design for ease of implementation and increasing isolation between the transmitting antenna and receiving antenna; and between two cross polarization.

2.4. Signal Processor

There are many polarimetric systems that utilise vector network-analyser as the signal conditioner and processor. With the advancement in computer technology, however, computer-based systems are getting popular in recent years. A typical example is the C/X-band system developed by Gogineni et al. [8], which is an inexpensive polarimetric scatterometer. Another interesting design is the C/X-band helicopter-borne scatterometer [9] that uses a digital signal processor, together with a microcomputer for real-time signal processing. In our project, a personal computer (PC) is used as the signal processor and storage device. Compared to the network analyser based system, the PC-based system is more flexible in handling various sources of data.

2.5. Dynamic Range of Backscattering Coefficient σ°

The radar system sensitivity is determined based on the various categories of earth terrain to be mapped such as man made target, ocean, sea-ice, forest, natural vegetation and agriculture, geological targets, mountain, land and sea boundary. From the open literatures, the typical value of σ° falls in the range of +20 dB to −40 dB [10, 11]. For vegetation the typical value of σ° vary from +0 dB to −20 dB. In our system, a dynamic range of 50 dB is targeted from +20 dB to −30 dB in order to facilitate the measurement of various types of target.

The design considerations are summarised in Table 1:

Table 1: Design considerations.

System Parameter	Design Considerations
System configuration	FM-CW
Operating frequency	C-band
Quantities measured	Magnitude and phase of the received signal in full polarisation capability
σ° dynamic range	Adequate dynamic range for targets of interest > 50 dB
Signal Processor	PC-based system
Operating platform	Ground base (using trolley)

3. DESIGN PARAMETER

This C-band radar system is proposed to operate at 5.3 GHz. The FMCW waveform with bandwidth of 400 MHz is selected in our design.

Since the radar system is planned to operate within the range of 20 m to 100 m, and assuming a signal to noise ratio of 10 dB is sufficient for detection of various target, for the lowest value

of backscattering coefficient, $\sigma_o = -30\,\mathrm{dB}$ and the system losses is assume as 6 dB, the minimum average power required to be transmitted will be 0.023 watt. Therefore a high power amplifier that has a 1-dB compression level larger than +20 dBm can be employed.

The Nyquist Criterion states that in order to construct a band-limited signal from its samples, the signal must be sampled at least twice the highest frequency. In practice, the signal is over-sampled at a rate higher than the Nyquist by 25% in order to account for non-ideal filter behaviour. From the minimum and maximum operation range, the beat signal from both range will be 6.4 KHz and 32 KHz respectively. Therefore in our design, sampling rate of 100 KHz is chosen. The data rate for single ADC channel (assuming 12 bits per sample for greater dynamic range) will be 1.2 Mbits/second. Therefore the total data rate can be calculated by multiplying the data rate for single channel with the number of receive channels which is equal to 150 kbyte/second. Design Parameters of the FMCW Polarimetric radar system is listed in Table 2.

Table 2: Design parameters of the polarimetric radar system.

System Parameter	Selected Value
System configuration	FM-CW
Operating Frequency	5.3 GHz (C-band)
Transmit power, P_t	20 dBm
Measurement range	20 to 100 m
σ° dynamic range	+20 dB to -30 dB
Received power, P_r	-41 dBm to -107 dBm
IF bandwidth	100 KHz
Minimum signal-to-noise ratio	10 dB

4. PROPOSED RADAR SYSTEM

The proposed block diagram of the fully polarometric, C-band, high bandwidth linear FM-CW imaging radar system is shown in Fig. 1. The system is based on a superheterodyne design. It consists of antenna system (2 horn antennas with vertical polarization, 2 horn antennas with horizontal polarization), a radar electronics subsystem and a data acquisition system.

The microwave source is generated from a voltage control oscillator (VCO). VCO is used to generate the required FM-CW signal. Portion of the signal from VCO is couple to the down-converter mixer to act as the reference signal for return echo. The major portion of VCO output is routed to a solid-state high power amplifier with 20 dB gain. The amplified signal is then radiated
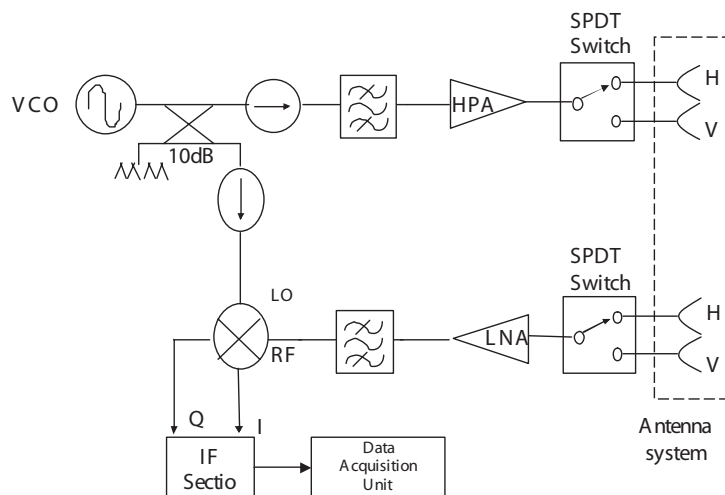
Figure 1: Proposed block diagram of C-band radar system.

through the antenna via a SPDT switch. The SPDT switch will determine the polarization of the transmitted signal. Thus a timing control is needed to ensure the full polarimetric signal is being received. The transmitted waveform is centered at 5.3 GHz with 400 MHz bandwidth.

The first stage of the receiver is a low noise amplifier (LNA) and followed by a band-pass filter. The down-converter mixer is used to convert the received signal to an intermediate frequency (IF). The signal from the mixer will be routed to IF section which consists of IF filter and amplifier. This radar system is proposed to employ a PC-based digital signal processing system for data acquisition. It consists of a high-speed 12-bits 100 KHz analogue-to-digital converter (ADC). The ADC is capable of converting the down-converted radar echoes into digital signals and stores them into high-density digital disk for future processing.

5. CONCLUSION

The conceptual design of a FMCW, fully polarometric imaging radar system system has been presented. This radar system can be used as a tool for monitoring and classification of simple target. This low cost system can be achieved by using simple RF subsystem, commercial component for chirp generation, PC based data acquisition and processing system.

ACKNOWLEDGMENT

REFERENCES

1. Nghiem, S. V., M. Borgeaud, J. A. Kong, and R. T. Shin, "Polarimetric remote sensing of geophysical media with layer random medium model" *Progress In Electromagnetics Research*, PIER 3, 1–73, 1990.
2. Ulaby, F. T. and C. Elachi, *Radar Polarimetry for Geoscience Applications*, Artech House, California, 1990.
3. Zebker, H. A., J. J. van Zyl, and D. Held, "Imaging radar polarimetry from wave synthesis," *Journal of Geophysical Research*, Vol. 92, No. B1, 683–701, 1987.
4. Boerner, W. M., B. Y. Foo, and H. F. Eom, "Interpretation of polarimetric co-polarization phase term in radar images obtained with JPL airborne L-band SAR system," *IEEE Transactions on Geoscience and Remote Sensing*, GE-25, 77–82, 1987.
5. Winebrenner, D. P., L. D. Farmer, and I. R. Joughin, "On the response of L-band polarimetric SAR signatures at 24 cm wavelength to the thickness of arctic sea ice in leads," *Radio Science*, 1994.
6. Boerner, W. M., M. B. El-Arini, C. Y. Chan, and P. M. Mastoris, "Polarization dependence in electromagnetic inverse problems," *IEEE Transactions on Antennas and Propagation*, Vol. 29, 262–271, 1981.
7. International Telecommunication Union's (ITU) World Radiocommunication Conference (WRC-97), 1997.
8. Gogineni, S., F. A. Hoover, and J. W. Bredow, "High performance, inexpensive polarimetric radar for in-situ measurements," *IEEE Transaction Geoscience and Remote Sensing*, Vol. 28, No. 4, 450–455, 1990.
9. Hallikainen, M., J. Hyyppa, J. Haapanen, T. Tares, P. Ahola, J. Pulliainen, and M. Toikka, "A helicopter-borne eight-channel ranging scatterometer for remote sensing: Part I: System description," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 31, No. 1, 161–169, 1993.
10. Hyyppa, J., J. Pulliainen, K. Heiska, and M. T. Hallikainen, "Statistics of backscattering source distribution of boreal coniferous forests at C- and X-band," *Proceeding of the 1986 International Geoscience and Remote Sensing Symposium*, Vol. 1, 241–242, 1986.
11. Pulliainen, J. T., K. Heiska, J. Hyyppa, and M. T. Hallikainen, "Backscattering Properties of boreal forests at the C- and X-bands," *IEEE Trans. on Geosc. and Remote Sensing*, Vol. 32, No. 5, 1041–1050, 1994.

Design and Development of a Low Cost Chirp Generator for Airborne Synthetic Aperture Radar

Y. K. Chan and S. Y. Lim
Faculty of Engineering & Technology, Multimedia University
Jalan Ayer Keroh Lama, Bukit Beruang, Melaka 75450, Malaysia

Abstract— An airborne C-band Synthetic Aperture Radar (SAR) has been designed and developed over the past few years as a key geometric data source for environmental monitoring by the Centre for Applied Electromagnetic of Multimedia University (MMU), Malaysia. Several modifications and enhancement are underway on top of the successful construction of the current SAR prototype. The highlight of this continuation work is the hardware implementation of the airborne SAR, specifically on the modifications of the existing SAR sensor microwave transceiver, such as the design and development of one high-speed dual-channel chirp generator using a digital approach. The digital approach is selected over an equivalent analogue solution for its stability, repeatability and flexibility following breakthrough and advancement in the world of digital electronics. This paper focuses on the design and development issues concerning the chirp generator.

1. INTRODUCTION

Synthetic Aperture Radar (SAR) is fast becoming a favourite among researches in the field of remote sensing ever since the nineties, overriding real aperture Side-Looking Airborne Radar (SLAR) as the sole imaging radar on a textbook case while attracting worldwide participation in the research and development of itself. The advantages of SAR have been detailed in many books and journals, which record the concrete proof and support behind the blossoming of SAR systems not only nationwide as in Malaysia but worldwide. Among them includes fine resolution achievable that made headline when the technique first came to light, often credited to Carl Wiley of Goodyear Aerospace in 1951 [1]. The onset of SAR has since significantly done away with many limitations revolving SLAR [2], like offering much longer apertures to improve the along-track resolution. Also, it permits round the clock operation as it is capable of providing its own illumination and therefore not dependent on light from the sun. Besides, being an active system, it can work under any weather conditions as clouds, fog and precipitation have no remarkable effect on microwaves. All these merits of SAR can be collectively phrased into one sentence that reads "SAR is an all-weather imaging tool that achieves fine along-track resolution by taking the advantage of radar motion to synthesize a large antenna aperture".

Nearly forty years since the emergence of SAR, Multimedia University, Malaysia has also stepped into this research field, developing theoretical modelling and image processing techniques on SAR images. In the meantime, a radar system of C-band with single polarization and linear FM has been designed and developed to serve as a test-bed for demonstrating SAR technology and acquiring data for the development of radar processing techniques and applications [3–5]. The current transmitter consists of an exciter, a high power amplifier (HPA), a STALO, a modulation circuitry, a chirp mode gate and some RF accessories such as RF cables and isolators; whereas the receiver consists of radio frequency (RF) section, intermediate frequency (IF) section, and data acquisition unit (DAU).

One breakthrough in technology made up the primary motivation of this research work, which is to take a digital approach towards designing and developing the exciter of SAR. The idea of taking the digital approach is rooted in the belief that, as for most contemporary radar systems designs, digital electronics offer better stability, repeatability, and flexibility over an equivalent analogue upshot. This revolution in radar designs has over the decades justified the common perception of many radar designers about radar system, that unlike early radar systems that consisted entirely of analogue circuits, digital techniques can now be employed too for optimization purpose [6].

2. DESIGN AND SYSTEM IMPLEMENTATION

2.1. SAR Design Consideration

Having surveyed the state-of-the-art in both SAR transmitter and digital chirp generator, it is determined that RF electronics is favoured while RAM/ROM based method is taken. RF electronics

method is chosen over all other methods because it is deemed the most practical yet most economical approach. On the other hand, RAM/ROM based method by comparison outshines all others in that it offers high bandwidth and linear ramps yet puts no threat to the coherency of SAR system.

2.2. Chirp Simulations

Typically, to construct a chirp generator there are two aspects involved, namely, software simulation and hardware implementation. For software simulation, an original source code is written on MATLAB to generate chirp signal prior to the actual implementation of hardware. Data is then obtained from the simulation and converted from the numerical data in decimal into machine code readable by the components using an assembler. Variation of the chirp rate and the data format, which is the bit of DAC, allows for viewing of chirp signals by simulation before decision is made as to which set of data to be used. After software simulation, direct hardware implementation is headed.

In ramp sweep, the output sine wave frequency is increased from a start frequency to a stop frequency. This produces a linear frequency versus time plot. Its accuracy, sweep time, and frequency resolution of the source are usually specified. The chirp generation can be described as a Cosine function with quadratic time sample multiply with the number of sample generated by the DAC. By interchanging between the Cosine and Sine function, depending on the application for dual channels operation, with In-Phase and Quadrature (IQ) Sampling being the driving force behind the dual-channel design. Table 1 lists down the specifications for running simulation of chirp signal on MATLAB.

Table 1: The specifications for running simulation of chirp signal on MATLAB.

Parameters	Values
Start Frequency, f_s	0 Hz
Stop Frequency, f_1	40 MHz
Start Time, t_s	0 s
Stop Time, t_1	20 μs
Bandwidth, B	40 MHz
Chirp Pulse Duration, τ_P	20 μs
Chirp Rate, K (K=B/τ_P)	2×10^{12}
Data format of DAC, m	8-bit
Number of Time Samples, nt	1000
Interval between Time Samples	2×10^{-8}
Functions	Sine
	Cosine

2.3. Hardware Construction

The chirp generator is built on the concept of Direct Digital Synthesis (DDS), first introduced by Barry and Fenwick in year 1965, which sparked the reports of a variety of digital generation methods in the literature with varying degree of success. Figure 1 shows the construction of the chirp generator. It consists of a microcontroller, a counter, a crystal oscillator, 2 UV-PROMs, and 2 DACs.

2.4. Output Chirp Analysis

Since the output chirp signal appears differently in the time domain and in the frequency domain, both results are presented. An Analogue-to-Digital Card (ADC) is used to digitise the analogue chirp signals into digital data for viewing and processing of the chirp signals. Figure 2 shows the In-Phase and Quadrature chirp signals of 40 MHz bandwidth (20 MHz on the left and right each) captured by the ADC. They are all together 1000 number of time samples and with each number constituting 20 ns, a concept illustrated by uniform time sampling, the obtained chirp pulse duration is 20 μs. It is observed that as the frequency of the chirp signal increases towards its left and right, the amplitude of the chirp signal reduces slightly. Since this phenomenon does not occur on the oscilloscope, its occurrence on the analysis in this part can be explained by the insufficient
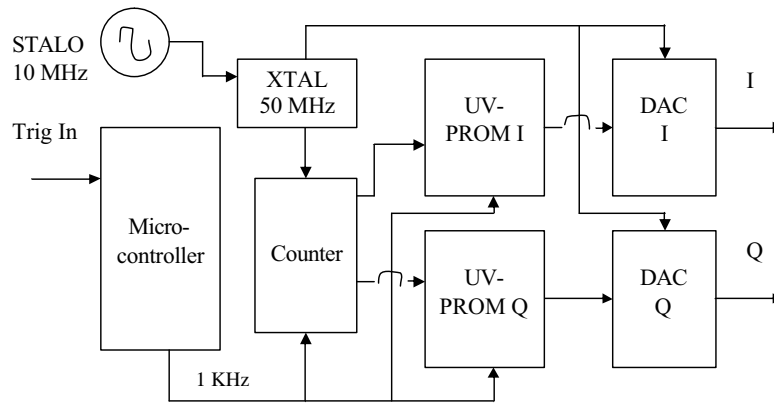
Figure 1: Basic hardware construction of a high-speed chirp generator.
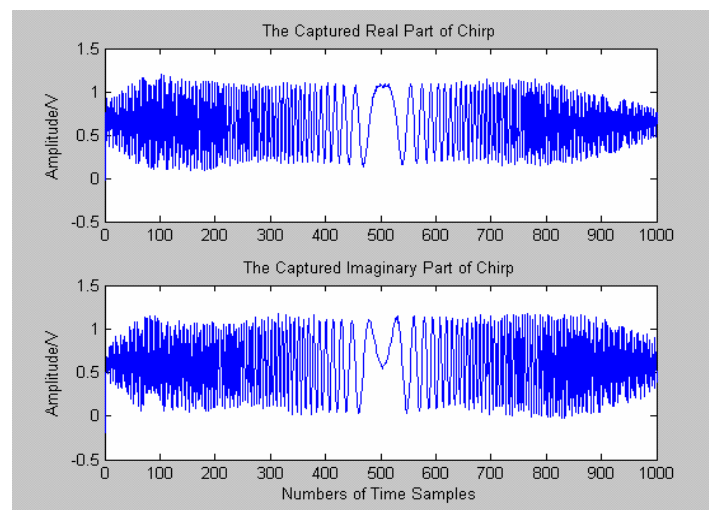

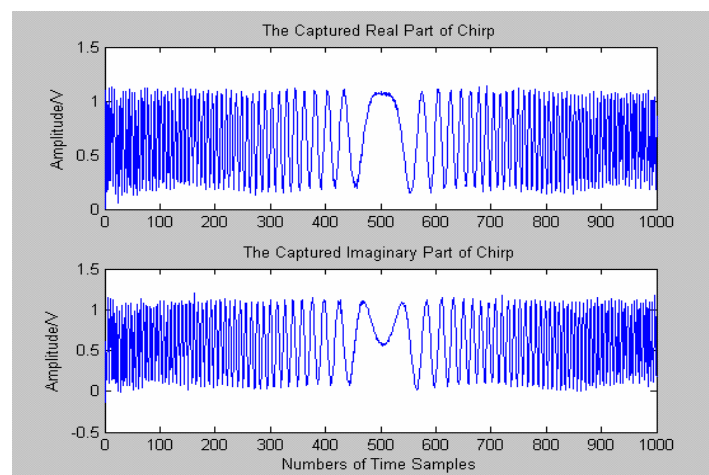Figure 2: The captured In-Phase and Quadrature chirp signals with 40 MHz bandwidth.

Figure 3: The captured In-Phase and Quadrature chirp signals with 20 MHz bandwidth.

sampling rate and the limited bandwidth of the ADC. The ADC has 50 MHz of bandwidth and its sampling rate is 100 MS/s (max) for single channel and 50 MS/s (max) for dual channels. Because dual-channel is used concurrently for I and Q channels, the sampling rate is limited to 50 MS/s only. In Figure 3, when the bandwidth is reduced to 20 MHz (10 MHz on the left and right each), this phenomenon disappears.

Figure 4 is obtained by modulating the In-Phase and Quadrature chirp signals together on MATLAB. A straight vertical line at the centre is seen in Figure 4 as a result of negative number elimination. On Matlab simulation, when chirp signal spreads from −128 to 128, no such vertical line appears but when it is pushed up by 128 that give 0 to 256, the vertical line appears on its spectrum. For software modulation, a resultant total bandwidth of 40 MHz is presented as one contiguous spectrum after applying *fftshift* on MATLAB. For hardware modulation of the In-Phase and Quadrature chirp signals, an I/Q modulator is required.
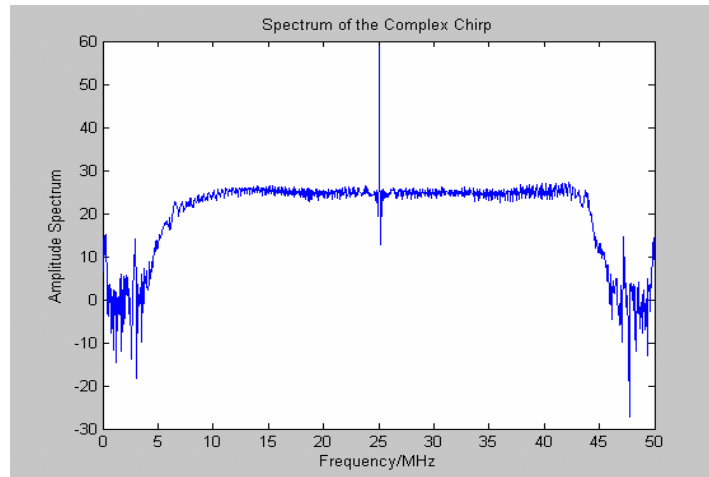
Figure 4: The spectrum of the complex chirp signal features a total bandwidth of 40 MHz.

3. CONCLUSION

This paper incorporates the simulation, development, and measurement of the output chirp signals of a newly constructed digital chirp generator, the name of which has been personalized from a common term "exciter" to the chirp generator. The primary function of the chirp generator is to generate a coded pulse waveform from a crystal oscillator that allows the input of reference signal from continuous tone STALO output. The pre-stored waveform approach taken to develop the chirp generator is proven at the end of the research to give maximum performance for short pulses.

REFERENCES

1. Wiley, C. A., "Synthetic aperture radar — A paradigm for technology evolution," *IEEE Trans. Aerosp. Electron. Syst.*, Vol. 21, 440–443, 1985.
2. Curlander, J. C. and R. N. McDounough, *Synthetic Aperture Radar, Systems and Signal Processing*, John Wiley & Sons, New York, 1991.
3. Chan, Y. K., M. K. Azlindawaty, V. Gobi, B. K. Chung, and H. T. Chuah, "The design and development of airborne synthetic aperture radar," *Proc. Igarss'00*, Vol. 2, 518–520, 2002.
4. Chan, Y. K., B. K. Chung, and H. T. Chuah, "Transmitter and receiver design of an experimental airborne synthetic aperture radar sensor," *Progress In Electromagnetic Research*, PIER 49, 203–218, 2004.
5. Koo, V. C., Y. K. Chan, V. Gobi, T. S. Lim, B. K. Chung, and H. T. Chuah, "The MASAR project: Design and development," *Progress In Electromagnetic Research*, PIER 50, 279–298, 2005.
6. Brandon, D. and J. Kornblum, "Synchronized synthesizers aid multichannel systems," *Microwaves & RF*, 57–68, 2005.

Electromagnetic Scattering Theory of Car Body Imaging Using Scanning Millimeter Wave Radar

Yasumitsu Miyazaki

Department of Media Informatics, Aichi University of Technology
50-2 Manori, Nishihasama-cho, Gamagori 443-0047, Japan

Abstract— Technical subjects of automatic driving system and safety transport system are very important problems. Automatic driving and cruising systems using microwave and millimeter radars in the cars are sensor systems for measurements of distances among driving cars, and recognition of road environments. Millimeter wave radar of short wavelengths using short pulses of millimeter wave carriers yields precise distance measurement for distance control system between driving automobiles, and auto-breaking system. Image information is given by reflected waves by scanning millimeter beam waves, like optical vision information. Comparing with optical waves, millimeter waves have less attenuation characteristics due to rains, mist, fog, and snows, and can give fairly good size resolution. Shape image recognition systems of car bodies with scattered and reflected waves in scanning millimeter radars, are very useful for ITS. Temporal and spatial characteristics of electromagnetic scattering and reflection by driving car bodies are studied by beam spectral function method. Fundamental characteristics and application of scanning millimeter wave radars of synthetic aperture type and multi-wavelengths for image recognition of car body are shown.

1. INTRODUCTION

In the ITS, Advanced Information Transport Systems, technical and social subjects of automatic driving system, and safety transport and traffic system are very important problems. Automatic driving and cruising systems using microwave and millimeter radars in the cars are sensor systems for measurements of distances among driving cars, detection of obstacles, and intelligent recognition of road environments [1–3]. Millimeter wave radar of short wavelengths and high frequency using short pulses of millimeter wave carriers yield precise distance measurement and image processing. ITS applications of millimeter wave radar are distance control system between driving automobiles, and auto-breaking system. Distance measurement using short pulse millimeter carriers are derived by time differences between transmitted pulse and received pulse, and image information are given by reflected waves by scanning millimeter beam waves, like optical vision information. Comparing with optical waves, millimeter waves have less attenuation characteristics due to rains, mist, fog, and snows, and can give millimeter size resolution. Shape image recognition systems of car bodies receiving scattered and reflected waves in scanning millimeter radars, are very useful for ITS. Temporal and spatial characteristics of electromagnetic scattering and reflection by driving car bodies are studied. Fundamental characteristics and application of scanning millimeter wave radars of synthetic aperture type and multi-wavelengths for image recognition of car body are shown.

Radiating millimeter beam packets from scanning millimeter radars and reflected beam packets of temporal Gaussian pulse form and spatial Gaussian beam form are expressed as Fourier components for time coordinates and as spectral functions expanded by Hermite-Gaussian functions for space coordinates [4]. Incident, reflected and scattered fields are studied using beam mode expansions derived by Hermite-Gaussian spectral functions. Boundary conditions of car bodies are applied to derive reflected and scattered field for spectral components. Car body imaging is analyzed by the Synthetic Aperture Radar Method. Correlation functions by transmitted and received fields yield image processing information for transverse image patterns at each driving time. Velocities of cars are given by frequency variations of received fields. Scanning millimeter wave radar in the front of car radiating short Gaussian pulse and forward Gaussian beam can construct spatial image at each time analyzing reflected and scattered waves by other cars and objects on the road. Based on this fundamental theory, safety driving systems may be constructed [5–10].

2. BODY SHAPE RECOGNITION BY SCANNING MILLIMETER WAVE RADAR

Fundamental Architectures of automobile driving millimeter wave radars are shown in Figs. 1 and 2. Distances among cars are given by pulse time difference measurement of transmitting millimeter wave pulse and receiving pulse. Driving situation of two automobiles A and B, and the millimeter wave radar system for collision protection are shown in Fig. 3.
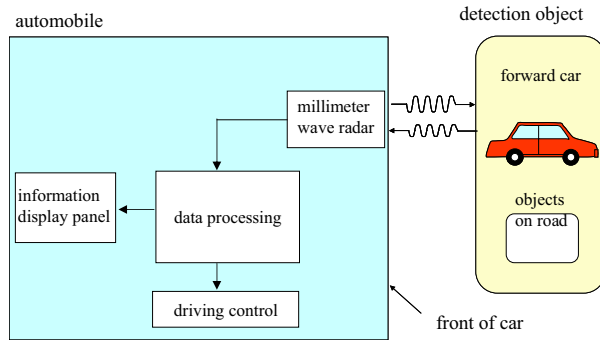
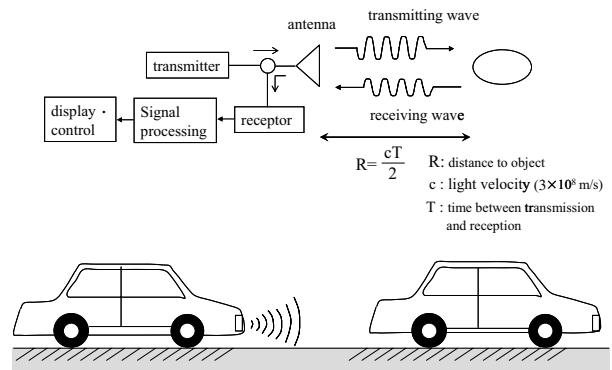Figure 1: Fundamental architecture of automobile driving millimeter wave radar.

Figure 2: Car detection using car collision protection radar.

Figure 4 shows car body shape recognition systems using scanning millimeter wave radar. S_{ij} are radiation apertures of scanning millimeter pulse radar and f_{ij} are multiple millimeter wave frequencies.
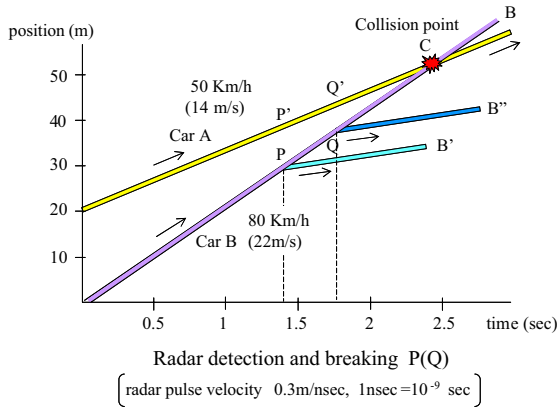

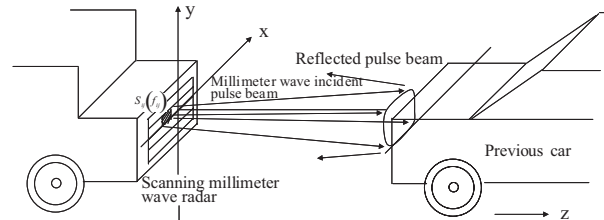Figure 3: Car collision protection millimeter wave radar system.

Figure 4: Scanning millimeter wave radar and car body imaging.

3. REFLECTION AND SCATTERING BY MILLIMETER WAVE PULSES

Beam waves with beam center $(-x_0, -y_0)$ of millimeter wave pulses are radiated from scanning millimeter wave antenna at position $z = -z_0$ as shown in Figs. 5 and 6. Body surfaces of forward car at time t are given by $z = f(x, y) + z_t(t)$ of position center z_t.
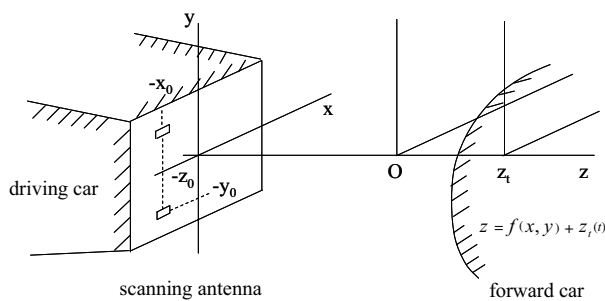
Figure 5: Reflection and scattering of millimeter wave pulse beam.
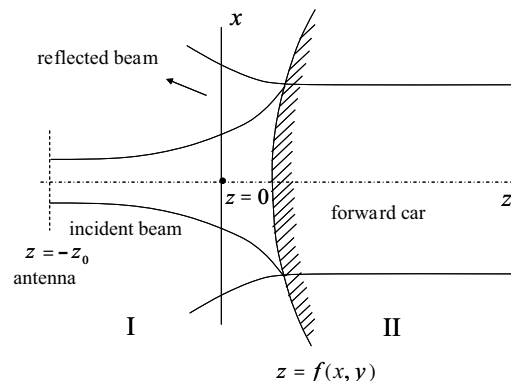
Figure 6: Incident beam and car body.

Incident pulses as shown in Fig. 7 are

$$\mathbf{E}^{(i)}(\mathbf{r}, t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \tilde{V}(\omega)\tilde{\mathbf{E}}^{(i)}(\mathbf{r}, \omega)e^{j\omega t}d\omega \tag{1}$$

where temporal function $V(t)$ and spectral function $\tilde{V}(\omega)$ are

$$V(t) = e^{-\left(\frac{2t}{T}\right)^2}e^{j\omega_0 t} = \frac{1}{2\pi}\int \tilde{V}(\omega)e^{j\omega t}d\omega, \quad \tilde{V}(\omega) = \frac{\sqrt{\pi}T}{2}e^{-\left(\frac{(\omega-\omega_0)T}{4}\right)^2} \tag{2}$$

when plane waves are incident, $\tilde{\mathbf{E}}^{(i)}(\mathbf{r}, \omega) = E_0 e^{-j\beta z}\mathbf{i}_\tau$ and using $\beta_0 = \omega_0/c$

$$\mathbf{E}^{(i)}(\mathbf{r}, t) = E_0 e^{j\omega_0 t - j\beta_0 z}e^{-\left(\frac{2}{T}\right)^2\left(t - \frac{z}{c}\right)^2}\mathbf{i}_\tau \tag{3}$$

and when the fundamental beam mode with y polarization is incident, in case of two dimensional space,

$$\tilde{\mathbf{E}}^{(i)} = E_0\mathbf{i}_y e^{-j\beta(z+z_0)}\frac{1}{\sqrt{1-j\zeta}}e^{-\frac{a^2 x^2}{2(1-j\zeta)}} \tag{4}$$

where $\zeta = \frac{(z+z_0)}{\beta}a^2$.

Car surface $z = f(x)$ with body center $x = 0$ sepates the region of free Space I and region of car Body II as shown in Figs. 6 and 8(a). Material constants of Regions I and II are ε_1, μ_1, σ_1 and ε_2, μ_2, σ_2. Incident wave, reflected wave and transmitted wave are $\tilde{\mathbf{E}}_i$, $\tilde{\mathbf{E}}_r$, and $\tilde{\mathbf{E}}_t$. In the Region I, $\tilde{\mathbf{E}}^{(I)} = \tilde{\mathbf{E}}_i + \tilde{\mathbf{E}}_r$, and in the Region II, $\tilde{\mathbf{E}}^{(II)} = \tilde{\mathbf{E}}_t$.

$$\mathbf{E}_n(\mathbf{r}_j, t) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\tilde{\mathbf{E}}_n(\mathbf{r}_j, \omega)e^{j\omega t}d\omega, \quad \tilde{\mathbf{E}}_n(\mathbf{r}_n, \omega) = \int_{-\infty}^{\infty}\hat{\mathbf{E}}_n(\beta_{tn})e^{-j\beta_{tn}x_n - j\sqrt{\beta_n^2 - \beta_{tn}^2}z}d\beta_{tn} \tag{5}$$

Wave numbers β_n are $\beta_i = \beta_r = \beta_1 = \omega\sqrt{\varepsilon_1^*\mu_1}$, $\beta_t = \beta_2 = \omega\sqrt{\varepsilon_2^*\mu_2}$, and $\varepsilon_i^* = \varepsilon_i - j\sigma_i/\omega$. β_{tn} are β_{ti}, β_{tr} and β_{tt}. If β_{tn} are small, we can use approximation of $\sqrt{\beta_n^2 - \beta_{tn}^2} \cong \beta_{tn} - \frac{1}{2}\frac{\beta_{tn}^2}{\beta_n} - \frac{1}{8}\frac{\beta_{tn}^4}{\beta_n^3}$, $\hat{\mathbf{E}}_n(\beta_{tn})$ can be expanded as series of Hermite-Gaussian functions

$$\hat{\mathbf{E}}_n(\beta_{tn}) = \sum_{m=0}^{\infty} a_{mn}\phi_m(\beta_{tn})\tau_n \tag{6}$$

where

$$\phi_m(\beta_{tn}) = \frac{1}{\sqrt{\alpha_n(2^m m!\sqrt{\pi})^{\frac{1}{2}}}}e^{-\frac{1}{2}\beta_n'^2}H_m(\beta_n') \tag{7}$$

$$\beta_n' = \beta_{tn}/\alpha_n$$
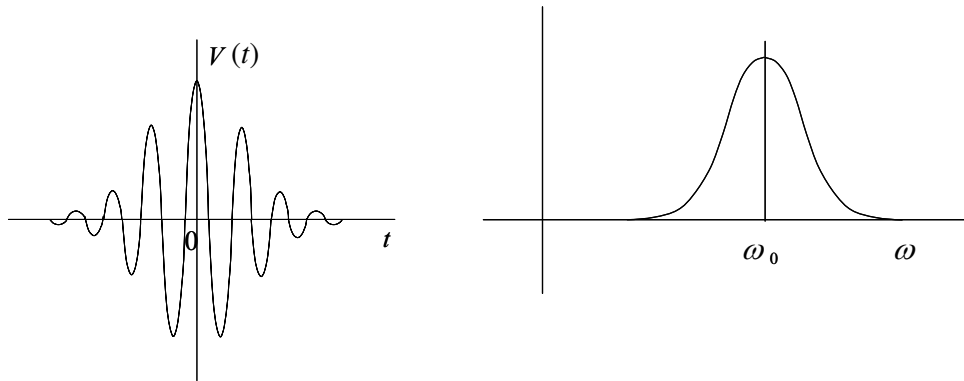
(a) Temporal characteristics of millimeter   (b) Frequency characteristics of millimeter
wave radar   wave pulse

Figure 7: Pulse characteristics of incident beam.

and coefficients a_{nm} are

$$a_{nm} \cong \int_{-\infty}^{\infty} \hat{\mathbf{E}}_n(\beta_{tn})\phi_m(\beta_{tn})\tau_n d\beta_{tn} \tag{8}$$

When incident beam wave is beam wave with the polarization of y-direction, using incident coordinates (x_i, z_i), incident beam waves are

$$\tilde{E}_{iy} = \int_{-\infty}^{\infty} \hat{E}_i(\beta_{ti})e^{-j\beta_{ti}x_i - j\sqrt{\beta_1^2 - \beta_{ti}^2}(z_i + z_{oi})}d\beta_{ti},$$

$$\tilde{H}_{ix} = \int_{-\infty}^{\infty} \hat{E}_i(\beta_{ti})\frac{(j\omega\varepsilon_1 + \sigma_1)\left(-j\sqrt{\beta_1^2 - \beta_{ti}^2}\right)}{-\beta_1^2}e^{-j\beta_{ti}x_i - j\sqrt{\beta_1^2 - \beta_{ti}^2}(z_i + z_{oi})}d\beta_{ti} \tag{9}$$

$$\tilde{H}_{iz} = \int_{-\infty}^{\infty} \hat{E}_i(\beta_{ti})\frac{(j\omega\varepsilon_1 + \sigma_1)(-j\beta_{ti})}{-\beta_1^2}e^{-j\beta_{ti}x_i - j\sqrt{\beta_1^2 - \beta_{ti}^2}(z_i + z_{oi})}d\beta_{ti}$$

Transmitted waves are obtained as Eq. (9), using transmission coordinates (x_t, z_t). Reflected waves are using $z_{0r} = z_{0i}$

$$\tilde{E}_{ry} = \int_{-\infty}^{\infty} \hat{E}_r(\beta_{tr})e^{-j\beta_{tr}x_r + j\sqrt{\beta_1^2 - \beta_{tr}^2}(z - z_{or})}d\beta_{tr}, \quad \tilde{E}_{ty} = \int_{-\infty}^{\infty} \hat{E}_t(\beta_{tt})e^{-j\beta_{tt}x_t + j\sqrt{\beta_2^2 - \beta_{tt}^2}(z - z_{ot})}d\beta_{tt} \tag{10}$$

Boundary conditions on car body surface $z = f(x)$ are

$$\mathbf{n} \times \left(\mathbf{E}^{(1)} - \mathbf{E}^{(2)}\right) = 0, \quad \mathbf{n} \times \left(\mathbf{H}^{(1)} - \mathbf{H}^{(2)}\right) = 0 \tag{11}$$

Here, $\mathbf{E}^{(1)} = \mathbf{E}_i + \mathbf{E}_r$ and $\mathbf{E}^{(2)} = \mathbf{E}_t$, $\eta_i = \frac{\omega\varepsilon_i + \sigma/j}{k_i^2}$, the following reflected and transmitted waves are obtained

$$\hat{E}_r(\beta_t) = \frac{\eta_1\sqrt{k_1^2 - \beta_t^2} - \eta_2\sqrt{k_2^2 - \beta_t^2}}{\eta_1\sqrt{k_1^2 - \beta_t^2} + \eta_2\sqrt{k_2^2 - \beta_t^2}}\frac{B_i(\beta_t)}{B_t(\beta_t)}e^{-j\sqrt{k_1^2 - \beta_t^2}z_o}\hat{E}_i(\beta_t),$$

$$\hat{E}_t(\beta_t) = \frac{2\eta_1\sqrt{k_1^2 - \beta_t^2}}{\eta_1\sqrt{k_1^2 - \beta_t^2} + \eta_2\sqrt{k_2^2 - \beta_t^2}}\frac{B_i(\beta_t)}{B_t(\beta_t)}e^{-j\sqrt{k_1^2 - \beta_t^2}z_o}\hat{E}_z(\beta_t) \tag{12}$$

In case of rectangular target on car body as

$$f(x) = \begin{cases} 0 & (|x| > W/2) \\ h & (|x| < W/2) \end{cases} \tag{13}$$

and defining $\varphi^{(0)} = |\tilde{E}_i(x, 0)|$, we obtained asymptotically

$$B_r(\beta_t) = \int_{-\infty}^{\infty} e^{\mp j\sqrt{k_1^2 - \beta_t^2}f(x)}e^{-j\beta_t x}\varphi^{(0)}(x)dx, \quad B_t(\beta_t) = \int_{-\infty}^{\infty} e^{-j\sqrt{k_2^2 - \beta_t^2}f(x)}e^{-j\beta_t x}\varphi^{(0)}(x)dx \tag{14}$$

From Eq. (12), back scattered and reflected waves are obtained. Pulse responses due to back scattering by car body are derived by Eq. (12).

When incident fundamental beam expressed in Eq. (4) is reflected by car body surface of convex creature with the radius R at $z = 0$, reflected beam is

$$\tilde{\mathbf{E}}^{(r)} = E_0\mathbf{i}_y e^{+j\beta(z + z_0 - z_r)}\frac{1}{\sqrt{1 - j\zeta_r}}e^{-\frac{a_r^2 x^2}{2(1 - j\zeta_r)}} \tag{15}$$

where $\zeta_r = \frac{2(z - z_r)}{\beta}a_r^2$, $z_r = \frac{R}{2 + R/z_0}$ and $a_r = \frac{R}{2z_0 + R}a$.

Received pulse responses are, when reflection coefficient of car body is R_s

$$\mathbf{E}^{(i)}(\mathbf{r}, t) = E_0 R_s e^{j\omega_0 t - j\beta_0(z + z_0 - z_r)}e^{-\frac{a_r^2}{2}x^2}e^{-(\frac{2}{T})^2(t - \frac{\ell}{c})^2}\mathbf{i}_\tau \tag{16}$$

where $\ell(x, z_0, z_r, R)$ is given by ray path of incident beam and reflected beam.

Body shape recognition is performed by image processing of received signals given by multiple scanning radar beams, as shown in Figs. 8(b) and (c), where I_{xz} and $I(x, y, z)$ are image characteristics given by azimuth compression $C_{RA}(x, t)$.
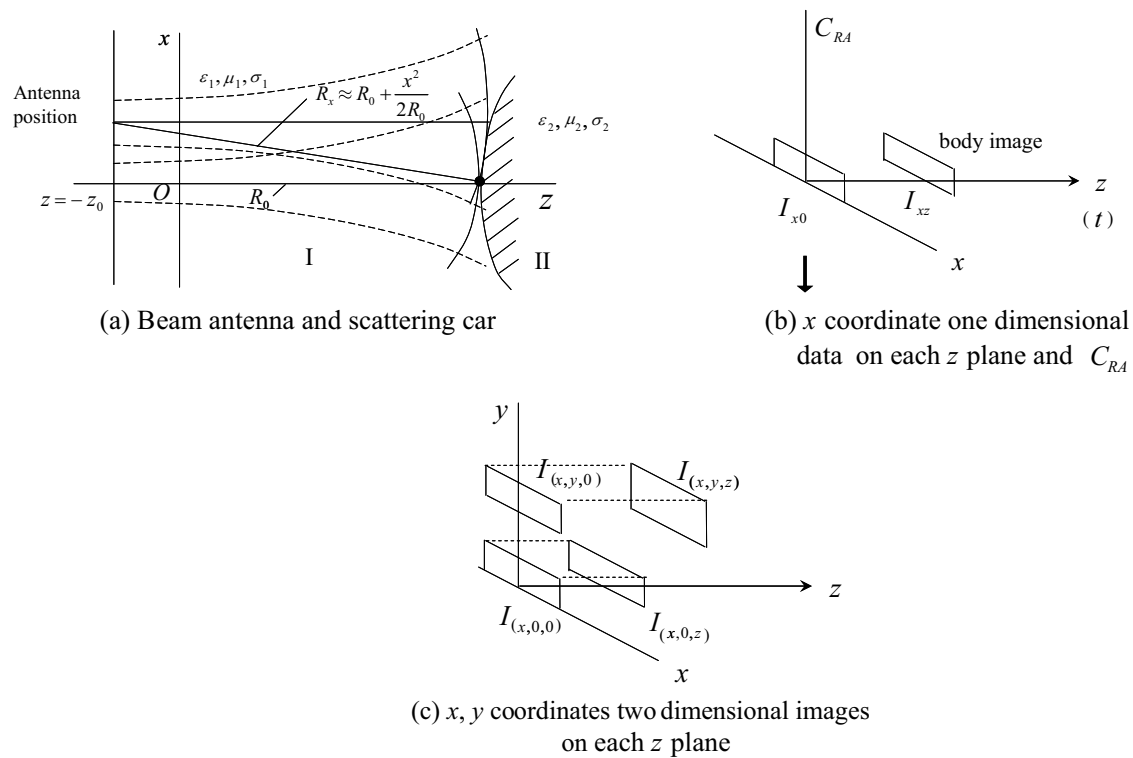
(a) Beam antenna and scattering car

(b) x coordinate one dimensional data on each z plane and C_{RA}

(c) x, y coordinates two dimensional images on each z plane

Figure 8: Scanning millimeter wave radar and synthetic aperture processing.

4. SUMMARIES

In the ITS, automatic driving systems using millimeter wave scanning radar are very useful. Particularly, body shape detection and recognition systems are extremely important. Body shape detection systems by using scanning millimeter wave radar are discussed by electromagnetic field theory with spectral functions of Gaussian functions. Pulse responses of very short millimeter waves are shown for reflected and scattered beam waves by automobile body. Synthetic aperture radar system can be performed based on this fundamental theory for body shape detection in ITS systems.

REFERENCES

1. Tanaka, M. and Y. Miyazaki, "Microwave simulation characteristics of electromagnetic wave scattering by automobile," *Trans. of IECE, Japan*, Vol. J66-B, No. 8, 1005–1012, 1983.
2. Miyazaki, Y., "Space-time impulse responses of electromagnetic backscatterings of beam waves for sensing," *Proc. of OFSET 2000*, 361–364, 2000.
3. Takahashi, K. and Y. Miyazaki, "Scattering analysis of millimeter wave radar for collision avoidance using FDTD method," *Trans. IEE Japan*, Vol. 120C-1, No. 1, 111–116, 2000.
4. Miyazaki, Y. and Y. Akao, "Reflection and mode conversion of beam wave through a curved interface between dielectrics and thick dielectric lens," *Trans. of IECE, Japan*, Vol. 51-B, No. 1, 25–31, 1968.
5. Takahashi, K. and Y. Miyazaki, "FDTD computer simulation of millimeter wave driving radar for ITS," *11th World Congress on ITS Nagoya*, Vol. 4088, 6, 2004.
6. Miyazaki, Y., K. Takahashi, N. Goto, and K. Tanaka, "Measurement characteristics of water film sensor using microwaves by FDTD analysis," *11th World Congress on ITS Nagoya*, Vol. 3308, 299, 2004.
7. Tanaka, M. and Y. Miyazaki, "An investigation of radio wave intensity distribution in the automobile body," *Trans. of IECE, Japan*, Vol. J69-B, No. 12, 1804–1810, 1986.
8. Azuma, H. and Y. Miyazaki, "FDTD analysis and simulation experiment of TV signal disturbance characteristics due to airplanes," *Trans. IEE Japan*, Vol. 120C-1, No. 1, 61–67, 2000.
9. Takahashi, K. and Y. Miyazaki, "FDTD simulation of signal detection characteristics of millimeter-wave driving radar for ITS," *Technical Report of IEICE*, MW2006-25, 1–4, 2006.

10. Miyazaki, Y., "Electromagnetic scattering theory of car body imaging using scanning millimeter wave radar," *Technical Report of IEICE*, MW2008-32, 7–12, 2008.

FDTD Analysis of Electromagnetic Wave Propagation for Out-door Active RFID System

Yasumitsu Miyazaki[1], Tadahiro Hashimoto[2], and Koichi Takahashi[1]

[1]Department of Media Informatics, Aichi University of Technology
50-2 Manori, Nishihasama-cho, Gamagori 443-0047, Aichi, Japan
[2]Broadband Technology Department, Synclayer, Inc.
1-20 Himegaoka, Kani 509-0249, Gifu, Japan

Abstract— In recent years, RFID systems have received much attention in security, logistics and medical fields. However, most of these systems are used at in-door and the RF tags are passive tags that are controlled by reader. We consider to use active RFID system of very weak UHF electromagnetic waves at out-door. In this paper, we describe the characteristics of electromagnetic wave propagation by RF tags using FDTD method. In transmit and receive points of the weak electromagnetic wave, receive characteristics are influenced by the established environment. We consider the road models as this environment. RF tags are carried by the human on the road, and the readers are installed on the electric pole and connected to the cable network. As out-door models, we consider road models of three types, the straight road, the T-type road and the cross road and numerical results of maximum receiving intensity of the electric fields in these models are shown.

1. INTRODUCTION

The reader of active RFID system receives periodic weak electromagnetic wave pulses of UHF carrier band from multiple tags and recognizes multiple tags. We considered to build human identification system for elementary school students by active RFID using UHF band as shown in Fig. 1. The active RFID is able to transmit the electromagnetic wave tens of meters far from the transmitter. Therefore, the area where the reader can recognize the presence of tags is wider than the area of passive RFID. In transmitting and receiving points of the weak electromagnetic wave, receive characteristics are influenced by the established environment. We consider the road models as this environment. RF tags are carried by the human on the road, and the readers are installed on the electric pole and connected to the cable network. Therefore, understanding of out-door propagation characteristics of RFID is necessary to develop high performance antenna of reader. In this paper, FDTD method is applied to show the propagation characteristics of RFID tags. The radiation field generated at tag locations on the road, where the reader can receive weak intensity of electric fields from active RFID are investigated [1–4].
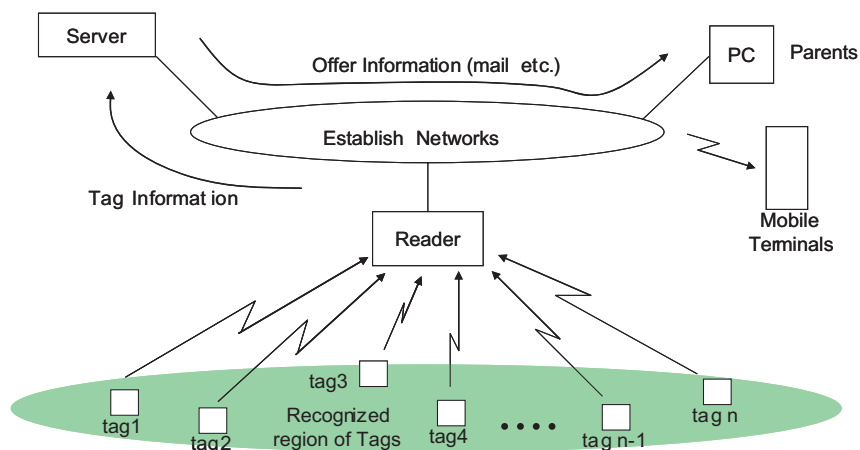
Figure 1: Components of RFID system.

2. ANALYSIS OF RF PROPAGATION IN RFID SYSTEM

Electromagnetic fields of RF propagation in RFID system of many radiation sources of RF tags can be studied by Maxwell's equations and boundary conditions. For RFID system, we consider electromagnetic waves radiated at m current sources of positions $\mathbf{r}_s$ and radiation time at t_s.

$$\nabla \times \mathbf{E}(\mathbf{r},\, t) = -\frac{\partial \mathbf{H}(\mathbf{r},\, t)}{\partial t} \tag{1}$$

$$\nabla \times \mathbf{H}(\mathbf{r},\, t) = \varepsilon \frac{\partial \mathbf{E}(\mathbf{r},\, t)}{\partial t} + \sum_{s=1}^{m} \mathbf{J}_s(\mathbf{r},\, t) \tag{2}$$

where, position vector is $\mathbf{r} = x\mathbf{i} + y\mathbf{j}$, and $\mathbf{J}_s$ are source currents corresponding to RF tags of each user at $\mathbf{r} = \mathbf{r}_s$. The source currents of Equation (2) are represented by the following equations for spatial and temporal functions.

$$\mathbf{J}_s(\mathbf{r},\, t) = \mathbf{J}(\mathbf{r} - \mathbf{r}_s,\, t - t_s) = \mathbf{J}(\mathbf{r} - \mathbf{r}_s)f_0(t - t_s) \tag{3}$$

$$\mathbf{J}(\mathbf{r} - \mathbf{r}_s) = \mathbf{J}_0\delta(\mathbf{r} - \mathbf{r}_s) \tag{4}$$

$$f_0(t - t_s) = \cos 2\pi f_{cs}(t - t_s)u_s(t) \tag{5}$$

Signal of each tag in RFID is FSK signal with currier frequency f_{cs}. The signal wave $f_0(t)$ for source current of Equation (3) and wave packet pulse of (5) is represented in Fig. 2, where f_c is the carrier frequency and $u_s(t)$ is pulse function $u_s(t) = 1$ (or 0) $t_s \leq t \leq t_{s+1}$, 0 (or 1) $t < t_s$, $t > t_{s+1}$.
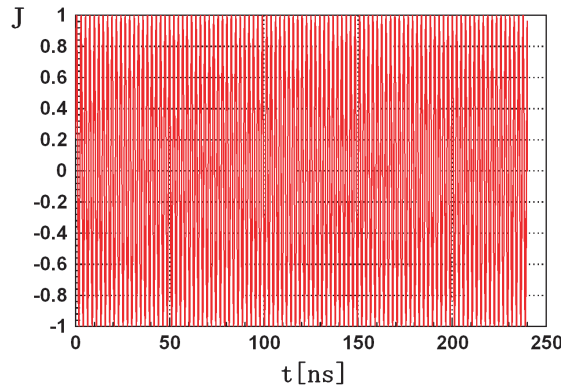
Figure 2: Transmiting pulse from tag.

3. NUMERICAL ANALYSIS OF PROPAGATION BY FDTD METHOD

In this paper, propagation characteristics are investigated by numerical computation method of FDTD. Electromagnetic fields of RFID system for complicated environment as urban status containing many houses, trees and cars on the roads, can not be easily studied by analytical methods.

(a) Straight road model   (b) T-type road model   (c) Cross road model
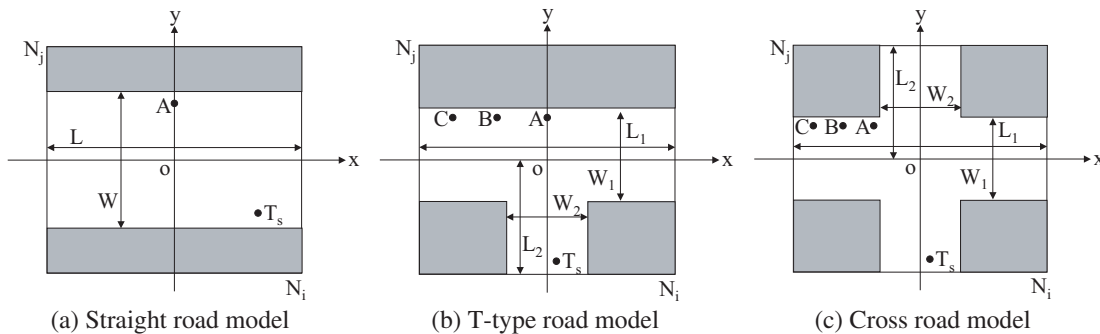
Figure 3: Analysis models for FDTD method.

The analysis models for several types of roads are shown in Figs. 3(a)–(c). These models consist of the straight road (a), the T-type road (b) and the cross road (c). The parameters of the straight road are length L and width W, and the parameters of the T-type and cross roads are horizontal length L_1 and width W_1 and vertical length L_2 and width W_2. In Fig. 3, the points A, B and C indicate the positions of receiving reader antennas and the point T_s indicates the position of one tag. N_i and N_j are the number of divisions in x and y direction. The walls of the road are considered to be perfect conducting plane or concrete.

The FDTD method is formulated by discretization of Maxwell's equations, if we define space coordinates and time parameters as $x = i\Delta s$, $y = j\Delta s$, $t = n\Delta t$ in Equations (1) and (2), as follows, for electric and magnetic fields $\mathbf{E}$, $\mathbf{H}$ and current sources of tags $\mathbf{J}_s$.

$$H_x^{n+\frac{1}{2}}\left(i, j+\frac{1}{2}\right) = H_x^{n-\frac{1}{2}}\left(i, j+\frac{1}{2}\right) - C_1\left\{E_z^n(i, j+1) - E_z^n(i, j)\right\} \tag{6}$$

$$H_y^{n+\frac{1}{2}}\left(i+\frac{1}{2}, j\right) = H_y^{n-\frac{1}{2}}\left(i+\frac{1}{2}, j\right) + C_1\left\{E_z^n(i+1, j) - E_z^n(i, j)\right\} \tag{7}$$

$$E_z^n(i, j) = C_2 E_z^{n-1}(i, j) - C_3 \sum_{s=1}^{m} J_s^{n-\frac{1}{2}}(i_s, j_s)$$

$$+ C_4\left\{H_x^{n-\frac{1}{2}}\left(i, j+\frac{1}{2}\right) - H_x^{n-\frac{1}{2}}\left(i, j-\frac{1}{2}\right) - H_y^{n-\frac{1}{2}}\left(i+\frac{1}{2}, j\right) + H_y^{n-\frac{1}{2}}\left(i+\frac{1}{2}, j\right)\right\} \tag{8}$$

where,

$$C_1 = \frac{\Delta t}{\Delta s \mu(i, j)} \qquad\qquad C_2 = \frac{1 - \Delta t \sigma(i, j)/2\varepsilon(i, j)}{1 + \Delta t \sigma(i, j)/2\varepsilon(i, j)}$$

$$C_3 = \frac{\Delta t/\varepsilon(i, j)}{1 + \Delta t \sigma(i, j)/2\varepsilon(i, j)} \qquad\qquad C_4 = \frac{\Delta t/\Delta s \varepsilon(i, j)}{1 + \Delta t \sigma(i, j)/2\varepsilon(i, j)}$$

i, j indicate the position in x-y plane, and n is number of time steps. $\Delta s = \Delta x = \Delta y$ is space increment, and Δt is time increment. For stability of the FD-TD solution, Δs and Δt should satisfy the condition, $\Delta t \leq \Delta s/\sqrt{2}c$. c is the propagation velocity of light in the free space. $\sigma(i, j)$ and $\varepsilon(i, j)$ are conductivity and dielectric constant of air space and road wall at the coordinate (i, j), respectively.

J_s is equivalence source and corresponds to a transmitting antenna of tag. (i_s, j_s) is the position coordinates of tag source current that represent by T_s in Fig. 3. Temporal signal wave is assumed to be $f_0(t) = \cos 2\pi f_{cs}$ $(t \geq 0)$.

For examples of Active RFID systems, Low Power Radio and Weak Radio waves are considered. Numerical parameters for FDTD analysis are shown in Tables 1 and 2, respectively. Recognized
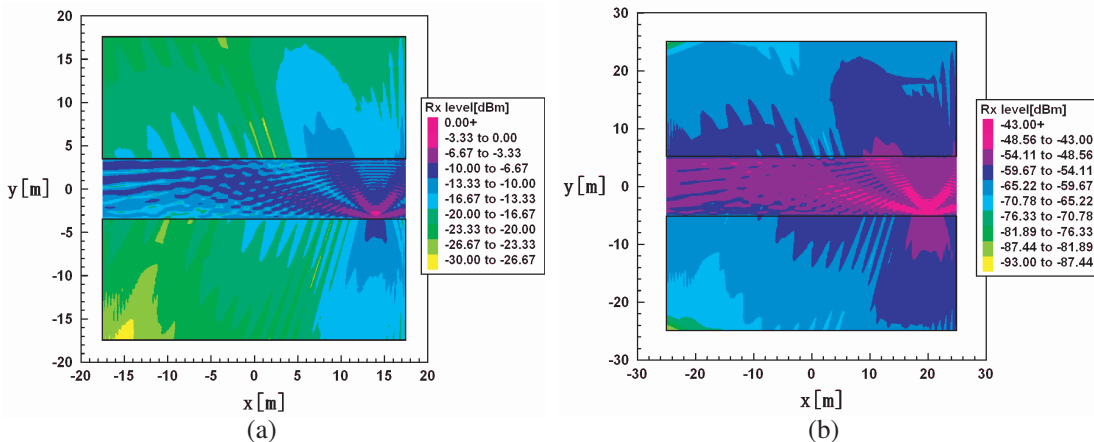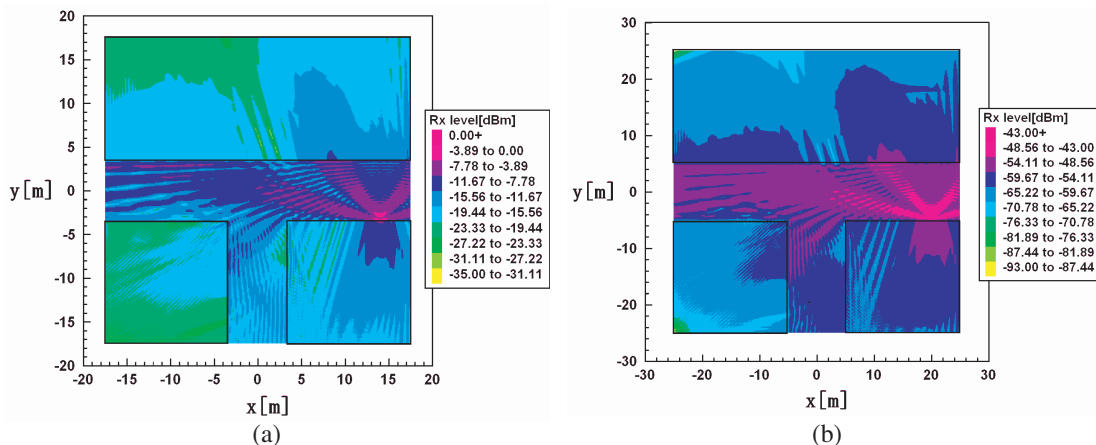

Figure 4: Maximum received intensity of constant state (case of straight road). Position of tag: $x = +14$ [m], $y = -2.8$ [m], concrete wall ($\varepsilon = 6.0$). (a) Case of low power radio, exchanged power level 1 [mW] $\to 0$ [dBm]. (b) Case of weak radio, exchanged power level 50 [nW] $\to -43$ [dBm].

level of reader is $-100\,\text{dB/m}$ in Low Power Radio wave and $-90\,\text{dB/m}$ in Weak Radio wave. Numerical results of the maximum received intensity in these models are shown in Figs. 4–7.

Table 1: Numerical parameters for case of low power radio waves.

Numbers of discrete $N_i \times N_j$	1001×1001	Horizontal road width W_1	$7\,[\text{m}]$
Space increment d_s	$0.035\,[\text{m}]$	Vertical road width W_2	$7\,[\text{m}]$
Time increment d_t	$0.08\,[\text{ns}]$	Horizontal road length L_1	$35\,[\text{m}]$
Carrier frequency f_c	$426\,[\text{MHz}]$	Vertical road length L_2	$17.5\,[\text{m}]$
Radiation times t_s	$240\,[\text{ns}]$	Position of Tag T_s	$+14\,[\text{m}],\ -2.8\,[\text{m}]$

Table 2: Numerical parameters for case of weak radio waves.

Numbers of discrete $N_i \times N_j$	1001×1001	Horizontal road width W_1	$10\,[\text{m}]$
Space increment d_s	$0.05\,[\text{m}]$	Vertical road width W_2	$10\,[\text{m}]$
Time increment d_t	$0.1\,[\text{ns}]$	Horizontal road length L_1	$50\,[\text{m}]$
Carrier frequency f_c	$300\,[\text{MHz}]$	Vertical road length L_2	$25\,[\text{m}]$
Radiation times t_s	$300\,[\text{ns}]$	Position of Tag T_s	$+20\,[\text{m}],\ -4\,[\text{m}]$

Figure 5: Maximum received intensity of constant state (case of T-type road). Position of tag: $x = +14\,[\text{m}]$, $y = -2.8\,[\text{m}]$, concrete wall ($\varepsilon = 6.0$). (a) Case of low power radio, exchanged power level $1\,[\text{mW}] \to 0\,[\text{dBm}]$. (b) Case of weak radio, exchanged power level $50\,[\text{nW}] \to -43\,[\text{dBm}]$.
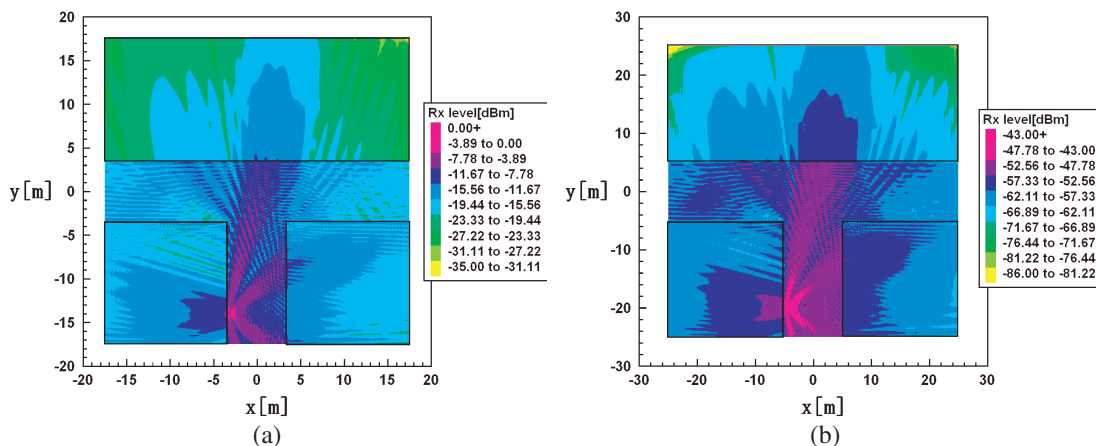
Figure 6: Maximum received intensity of constant state (case of T-type road). Position of tag: $x = -2.8\,[\text{m}]$, $y = -14\,[\text{m}]$, concrete wall ($\varepsilon = 6.0$). (a) Case of low power radio, exchanged power level $1\,[\text{mW}] \to 0\,[\text{dBm}]$. (b) Case of weak radio, exchanged power level $50\,[\text{nW}] \to -43\,[\text{dBm}]$.
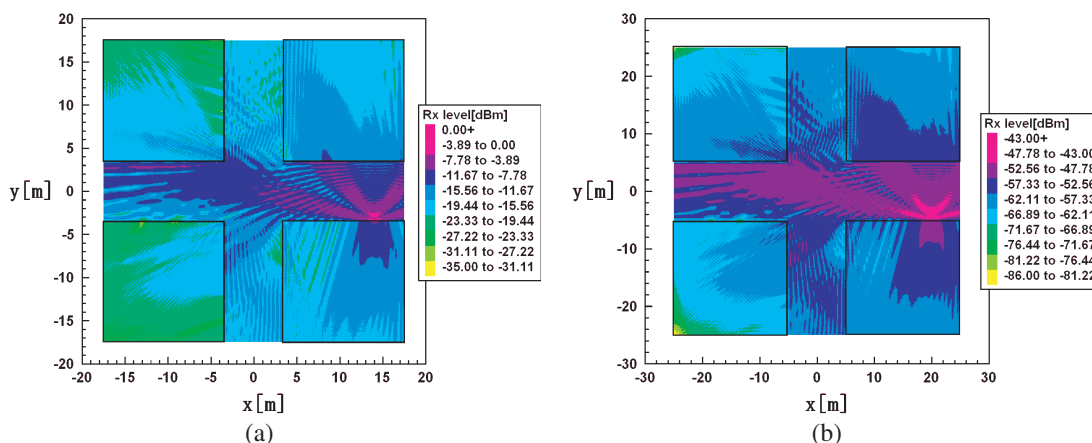
Figure 7: Maximum received intensity of constant state (case of cross road). Position of tag: $x = +14 \, [\text{m}]$, $y = -2.8 \, [\text{m}]$, concrete wall ($\varepsilon = 6.0$). (a) Case of low power radio, exchanged power level $1 \, [\text{mW}] \rightarrow 0 \, [\text{dBm}]$. (b) Case of weak radio, exchanged power level $50 \, [\text{nW}] \rightarrow -43 \, [\text{dBm}]$.

In case of straight road, line-of-site communication is provided and strong received intensity is observed at almost all area as shown in Fig. 4. It seems easy for the reader to communicate with tags in the straight road. In case of T-type road, two cases of different position of tags are simulated. In Fig. 5, the position of tag is horizontal road in Fig. 5, and vertical road in Fig. 6. The area where the received intensity becomes weak is observed behind walls. In case of cross road model, weak received intensity is observed in vertical road when the tag is placed at horizontal road. These results show basic foundation for the design of RFID systems.

4. CONCLUSIONS

For the optimum design of RFID system, we analyzed the characteristics of electromagnetic wave propagation in out-door active RFID system consisting of multiple tags and reader using FDTD method. In this paper, we formulated FDTD method for electromagnetic fields radiated from source current of RFID Tag. We study computer simulation using the above FDTD formulation, and make clear the characteristics of propagation. As a next step, we perform field experiments using actual tags and compact reader in out-door and compare numerical results with the experimental results. These results yield basic foundation for the design of RFID system, and system design of resolving collision problems for multiple pulses.

REFERENCES

1. Rodriguez, G., Y. Miyazaki, and N. Goto, "Matrix-based FDTD parallel algorithm for big areas and its applications to high-speed wireless communications," *IEEE Trans. Antennas and Propagation*, Vol. 54, No. 3, 785–796, 2006.

2. Masuda, T. and Y. Miyazaki, "Scattering characteristics of electromagnetic environment by construction on the ground," *Trans. IEE Japan*, Vol. 118-C, No. 1, 112–117, 1998.

3. Taki, K. and Y. Miyazaki, "Input impedances and current distributions for meander line antennas with planar coupled parasitic meander element," *PIERS Proceedings*, 1700–1703, Beijing, China, March 26–30, 2007.

4. Masuda, T., Y. Miyazaki, and Y. Kashiwagi, "Analysis of electromagnetic wave propagation in out-door active RFID system using FD-TD method," *PIERS Online*, Vol. 3, No. 6, 937–939, 2007.

The Effect of Shorting Post on Axe-shaped Circular Antenna Miniaturization

Jingxian Liu[1], Salman Naeem Khan[2], and Sailing He[1,3]

[1]Centre for Optical and Electromagnetic Research
Zhejiang University, Zijingang Campus, Hangzhou 310058, China
[2]Department of Physics, COMSATS Institute of Information Technology
Defense Road, Off Raiwind Road, Lahore, Pakistan
[3]Division of Electromagnetic Engineering, School of Electrical Engineering
Royal Institute of Technology, S-100 44 Stockholm, Sweden

Abstract— An edge shorted axe-shaped circular antenna is studied for miniaturization. The effect of the location and number of shorting post on the resonance frequency, impedance match and radiation pattern are studied. By inserting a shorting post near the edge of circular axe-shaped patch, the antenna size (resonates at 1.36 GHz) is reduced about 72% compared to the antenna without shorting post (resonates at 2.58 GHz) and 91% compared to the conventional circular antenna (resonates at 4.6 GHz). The effects of the distance between feeding point and shorting post, the number of shorting posts and the post radius are analyzed. The results show that by putting a shorting post near the edge of the axe-shaped circular patch, the resonant frequency is lower than a post near the center. Meanwhile, the feed should be excited near the shorting post for better impedance match. The effect of increasing the number of shorting post is also studied. If the number of shorting posts is increased to 3, the resonant frequency shifts to 1.48 GHz, which is higher than that with only one shorting post. However, the design offers stable return loss and broadside radiation pattern.

1. INTRODUCTION

Microstrip patch antennas are well known for their compactness and low profile. Nowadays, the rapid development of wireless communication has aroused extensive studies on the miniaturization of patch antennas. Various methods are used to reduce the antenna size such as increasing the height of substrate, etching U or E-shaped slots on the patch [1, 2], employing stacked structure [3] and recently metamaterials [4]. Among these methods, cutting patch shape and inserting shorting posts are two most commonly ways for their simplicity in fabrication. In our previous study, two circular arcs were cut from the circular antenna to reduce the patch size [5]. Furthermore, one shorting post is inserted to realize dual band application [6].

This paper presents an advanced design on the axe-shaped circular antenna by inserting shorting posts for low frequency design. Parametric analyses of single and multiple posts are presented and compared with measured results.

2. DESIGN CONSIDERATION AND EXPERIMENTAL RESULTS

The antenna scheme is shown in Fig. 1. The shorting post and feed point are along Y axis with their coordinate $Y_p = -7.8$ mm, $Y_f = -6.5$ mm. The radius of circular patch (R_0) is fixed to be 8.42 mm with arc radius $R_c = 4$ mm. The patch is built on FR4 material with relative permittivity of 4.4 and height of 1.6 mm ($\sim 0.007\lambda_0$, where λ_0 represents the wavelength of antenna at 1.36 GHz). The ground size is 24 mm × 24 mm for a total size minimization.

2.1. Single Post Antenna

The simulated and measured results of single post antenna are presented in Fig. 2. The simulation is obtained using HFSS10 with simulated resonance at 1.36 GHz. The antenna size is $1 - (1.36/2.58)^2 = 72\%$ smaller than axe shaped antenna and $1 - (1.36/4.6)^2 = 91\%$ of the conventional circular antenna, where 2.58 GHz and 4.6 GHz are the resonant frequency of axe shaped antenna and conventional antenna with the same R_0 and R_c respectively.

The measured resonance is 1.46 GHz, The shift of resonance and impedance may be caused by the induced parasitic current on the outer conductor of the coaxial line because the feed is positioned near the edge of patch.

The radius of shorting post R_{pin} is demonstrated in Fig. 3. With R_{pin} increases from 0.22 mm to 0.52 mm, the resonant frequency of antenna increases from 1.32 GHz to 1.42 GHz. When fixing
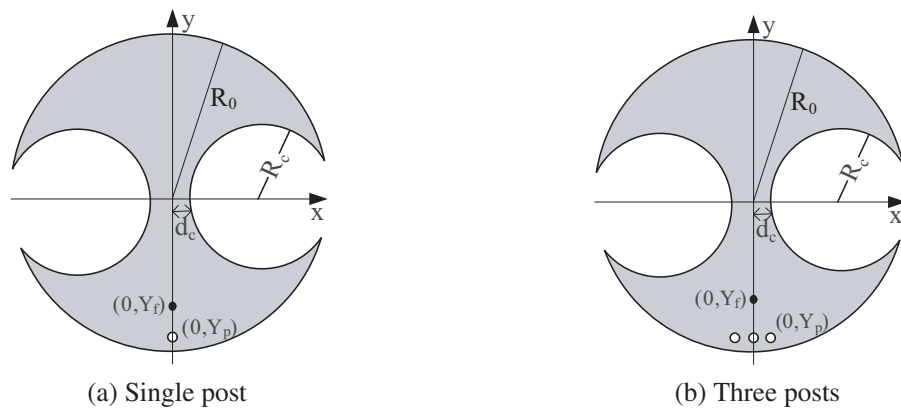
(a) Single post

(b) Three posts

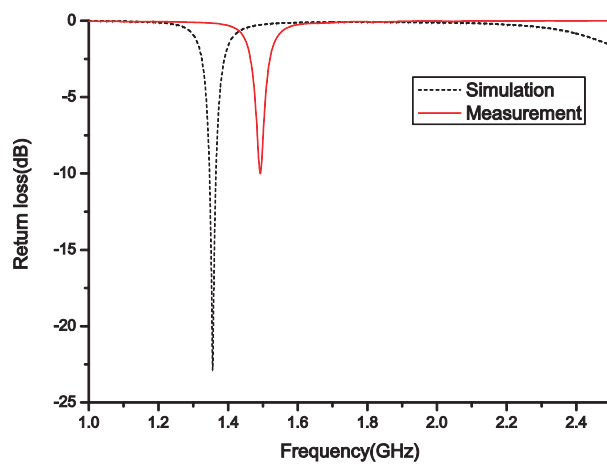Figure 1: Scheme of post shorted axe-shaped antenna.

Figure 2: Return loss of single post axe-shaped antenna.

$Y_f = -6.5$ mm, moving the post (Y_p) from -7.8 mm to 0.2 mm, the resonance is also shifting upwards, as shown in Fig. 4. In other words, the antenna size can be reduced by smaller post radius and larger $|Y_p|$. The best impedance match is achieved by setting the feed close to the post.
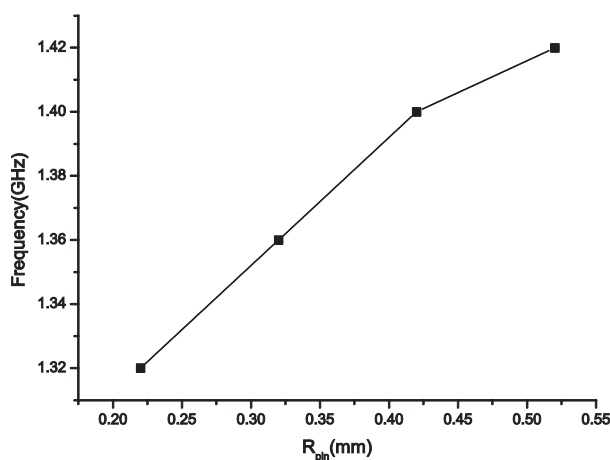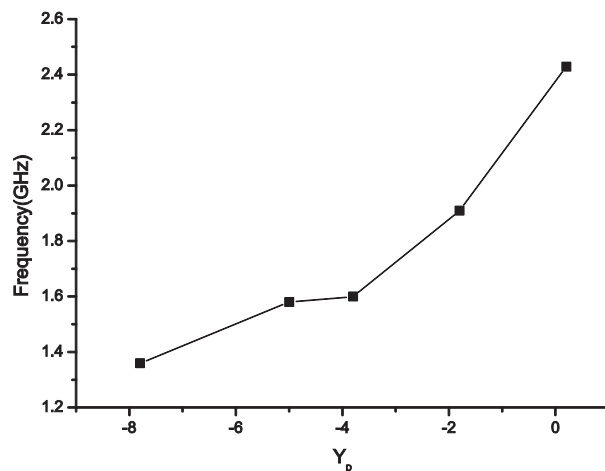

Figure 3: effect of $R_{\rm pin}$ on the resonance.

Figure 4: Effect of Y_p on the resonance.

2.2. Three Posts Antenna

For comparison, an axe-shaped antenna with three shorted posts is proposed in Fig. 1(b). The radius of each post is 0.1 mm, which is 1/3 of that of single post antenna. The space between adjacent posts is 1.5 mm. Other parameters are the same as that of single post antenna discussed above.

The simulated and measured return losses are shown in Fig. 5. The simulated antenna resonates at 1.48 GHz. The measured and simulated results match well with each other. The measured result seems more immune to circumstance than the single shorted antenna.
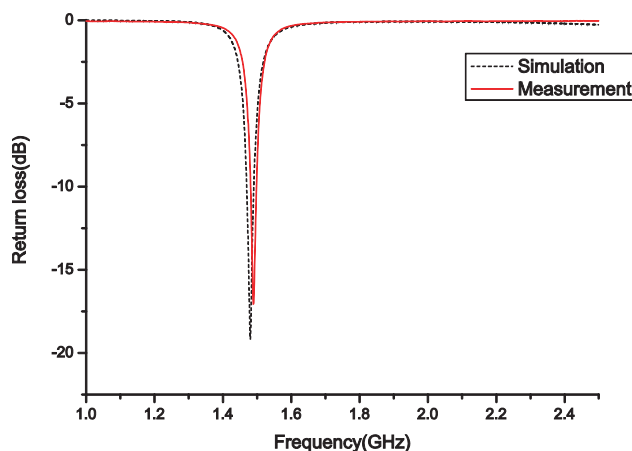
Figure 5: Return loss of axe-shaped antenna by three shorting post.

Figure 6 shows the measured radiation pattern of two discussed antennas. The experiment is carried out on flat roof of a building to resemble the free space. The slight slope of the radiation pattern is caused by the asymmetric position of feed. The multipath lobes and concaves might due to the undesired reflection of open environment. The three posts antenna shows better broadside property than the single post antenna in both E plane and H plane.

(a) Single post   (b) Three posts
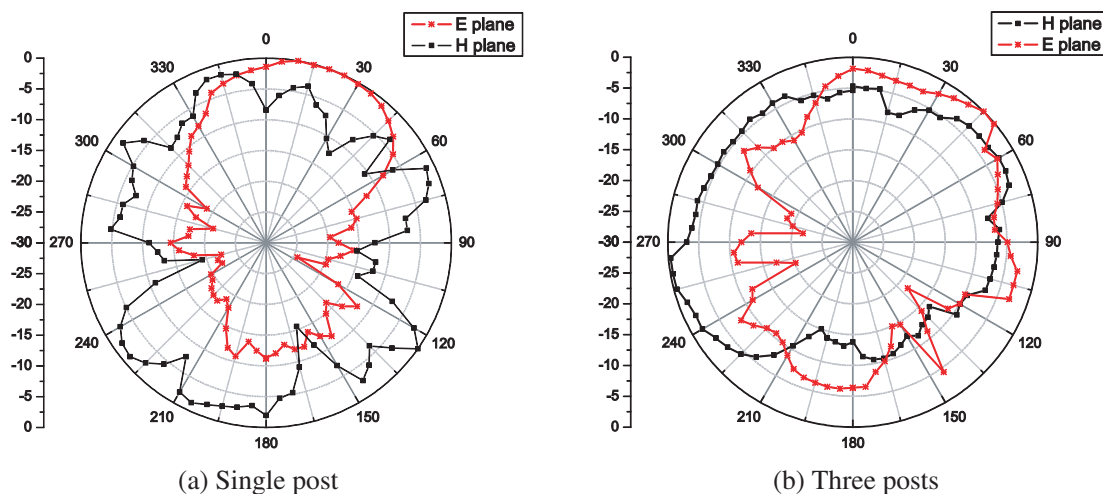
Figure 6: Radiation pattern of edge shorted axe-shaped antenna.

3. CONCLUSIONS

The idea of using shorting post on axe-shaped circular antenna to design compact patch antenna is successfully demonstrated. The resonant frequency is reduced to 52% with substrate thickness of $0.007\lambda_0$ at the center frequency. The frequency gets lower with smaller post radius. Good impedance match is achieved by shortening the distance between feed and shorting post. By using three shorting posts instead of one single post, the measured return loss seems more stable. Meanwhile, the antenna with three posts shows better broadside properties than single post antenna

in their radiation pattern, which can be an alternative choice for special requirements on radiation pattern.

ACKNOWLEDGMENT

REFERENCES

1. Choi, W., N. S. Seong, J. M. Kim, "A planar inverted-F antenna (PIFA) to be attached to metal containers for an active RFID tag," *Antennas and Propagation Society International Symposium*, Vol. 1B, 508–511, IEEE, 2005.
2. Chair, R., K. F. Lee, and C. L. Mak, "Wideband half U-slot antennas with shorting pin and shorting wall," *Antennas and Propagation Society International Symposium*, Vol. 4, 4132–4135, IEEE, 2004.
3. Wong, K. L., *Compact and Broadband Microstrip Antennas*, John Wiley & Sons, Inc., New York, 2002.
4. Bilotti, F., A. Alu, and L. Vegni, "Design of miniaturized metamaterial patch antennas with μ-negative loading," *IEEE Transactions on Antennas and Propagation*, Vol. 56, No. 6, 1640–1647, 2008.
5. Hu, J., S. N. Khan, J. Xiong, and S. He, "Circular axe-shaped patch antenna for compact and dual band wireless applications," *Journal of Electromagnetic Waves and Applications*, Vol. 22, 1655–1660, 2008.
6. Liu, J. X., S. N. Khan, and H. Li, "Shorted pin circular axe-shaped patch antenna for compact and dual band applications," *Journal of Electromagnetic Waves and Applications*, Vol. 22, 2105–2112, 2008.

Microstrip Slot Antenna with a Finite Ground Plane for 3.1–10.6 GHz Ultra Wideband Communication

Huan-Cheng Lien[1], Yung-Cheng Lee[1], Wen-Fei Lee[2], and Huei-Chiou Tsai[2]

[1]Department of Security Management, WuFeng Institute of Technology, Taiwan
[2]Department of Electrical Engineering, WuFeng Institute of Technology, Taiwan

Abstract— This paper introduces an Ultra Wideband microstrip Slot Antenna conformed to craving for the band of IEEE802.15.3a UWB (3.1 GHz $\sim$ 10.6 GHz) communication. In this paper, the design for enhancing impedance bandwidth of wideband microstrip slotting antenna with a finite ground plane is proposed and studied. With this design, a matching impedance bandwidth (SWR $\leq$ 2) about of more than 146% was achieved; as a result, the bandwidth can be located at the wireless communications from 2.3 GHz to 12 GHz. The variations of gain are changing from 0.2 to 6.98 dBi. In addition, the design has been a planar profile and it can easily be integrated in small mobile units; besides, it also can be in the laptops or various remote-sensing devices etc..

1. INTRODUCTION

In the communication systems, to develop small size, lightweight, low profile, broad bandwidth, and proper polarization are fundamental demands in the antenna design for the miniaturization of the communication equipment. Microstrip antennas have many desirable features, such as low profile, lightweight, and are usually fabricated by a photolithographic etching process or a mechanical milling process of these kinds antennas, making the construction relatively easy and inexpensive. These features make microstrip antennas are one of the most widely used types of antennas in the microwave frequency range, and they are often useful for many applications in the satellite communication and mobile communication systems. However, the major drawback of these kinds of antennas inherently has limited impedance bandwidth (VSWR? 2), if the narrow bandwidth of the microstrip antenna can be widened, then it can serve as a dual antenna for second- and third-generations of mobile communications systems. Therefore, developing broadband techniques to enhance the bandwidths of the microstrip antennas is very important.

Recently, most of the research on microstrip antennas focused on methods to increase their bandwidth. Slot antennas exhibit wider bandwidth, lower dispersion and lower radiation loss than microstrip antennas, and when feeding by a coplanar waveguide they also provide an easy means of the parallel and series connection of active and passive elements that are required for improving the impedance matching and gain [1]. The U-slot antenna, which achieves a relatively broad bandwidth without a parasitic patch, has been reported [2]. A lot of slot antennas for enhancing impedance bandwidth have been investigated [3–5].

A broader bandwidth, obtained using an improved feeding method, has also been reported [6]. It is very important to choose a suitable feeding circuit since it controls the antenna performance in terms of bandwidth. A bandwidth of 10 GHz ultra wide bands square planar metal-plate monopole antennas has been proposed propos ing in a 3D geometric configuration [7]. However, this 3D geometry designs needs more space, which is not suitable for mobile terminals, where the space is very limited.

The recent allocation of the 3.1–10.6 GHz frequency spectrums by the Federal Communications Commission (FCC) for Ultra Wideband (UWB) radio applications has had presented challenges for the antenna designers. For many UWB wireless communications, the successful transmission and reception of UWB pulses, sufficient impedance matching, the characteristics of the antenna with omni directional radiation pattern, high radiation efficiency, and easy manufacturing are required.

In this paper, a novel design of microstrip slot antenna with a finite ground plane for the desired band of IEEE802.15.3a UWB (3.1 GHz $\sim$ 10.6 GHz) is studied. This design of the proposed antenna is different from that of the other slot antennas with a tuning stub to enhance impedance bandwidth [8–12], and successfully implemented and the simulated results show reasonable agreement with the measured results. From experimental results, the proposed antenna shows that geometries can a significant increase in the impedance bandwidth obviously and, with respect to 10-dB impedance, the widest band obtained was 95%. Radiation patterns and gains are also examined.

2. DESIGN AND REALIZATION

The configuration of the ultra wideband microstrip slot patch antenna is shown in Fig. 1. The antenna has compact dimensions of $30.5 \times 35.3\,\text{mm}^2$, which consists of wide-slot ground plane with the same four small corners, fork-like U-type radiating element, a 50 O microstrip feed line, a small parasitical ground plane and a finite ground plane. Where the fork-like U-type radiating element and the 50 O microstrip feed line is coplanar with place in the back of the wide-slot ground plane, and printed on the substrate FR4 of thickness 1.6 mm and relative permittivity 4.4.

(a) Wide-slot ground plane   (b) Cross section
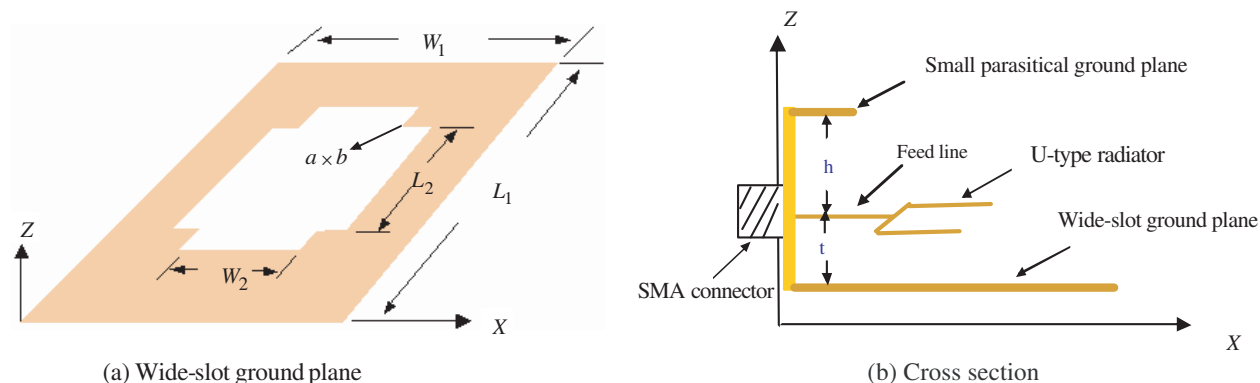
Figure 1: Configurations of the ultra wideband microstrip slot antenna with a small parasitical ground plane.

The internal and external wide-slot ground plane has dimensions of L_1 ($L' 2$) wide, W_1 ($W' 2$) long, respectively. The dimension of the small corner is Xa by Yb. The fork-like U-type radiating element is divided into three parts., These dimensions are ℓ_1 by w_1, ℓ_2 by w_2, and ℓ_3 by w_3, respectively. The small parasitical ground plane finite ground plane has dimensions of x wide y long, and the thickness of d, that is suspended at h mm under the fork-like U-type radiating element, and connected to the 50 ohms SMA connector. By properly adjusting the dimensions of the small parasitical ground plane finite ground plane, the radiator element, and the spacing h of between the small parasitical finite ground planes relative to the radiating element, a better impedance matching can be achieved. The design optimized parameters of all the components are listed in the Table 1. A photograph of the fabricated rectangular patch antenna is shown in Fig. 2.

Table 1: The relative parameters of the proposed antenna (Unit: mm).

1	U-type radiating element	$c = 0.9$, $d = 10$, $e = 10.4$, $f = 1.1$
2	Rectangular microstrip feed line	3×4.12
3	Wide-slot ground plane	$L_2 = 23.9$, $W_2 = 30.5$ $L_1 = 35.3$, $W_1 = 42.8$
4	small parasitical ground plane	$x = 16$, $y = 8$, thickness $= 0.35$
5	SMA connector	50 ohm
6	small corner	$a = 2.5$, $b = 3.2$
FR4	Dielectric substrate	$\varepsilon_{r1} = 4.4$, $\tan\delta = 0.022$
RO	Dielectric substrate	$\varepsilon_{r2} = 3.38$, $\tan\delta = 0.0025$

3. MEASUREMENTS AND DISCUSSION

The design was also analyzed using Zeland Software's IE3D simulation package. According to the above-mentioned design optimized parameters of all the components are listed in the Table 1, the performance of impedance bandwidth is improved for y is between 8 mm $\sim$ 9 mm. Using an HP8720D vector network analyzer, the return loss curves of both simulated and measured with the finite ground plane proposed antenna is obtained at $2 \sim 12$ GHz, as depicted in Fig. 3. The comparison between the measurements and software predictions are very close. The difference between the two graphs is also due to the fact that the proposed antenna was built on a finite
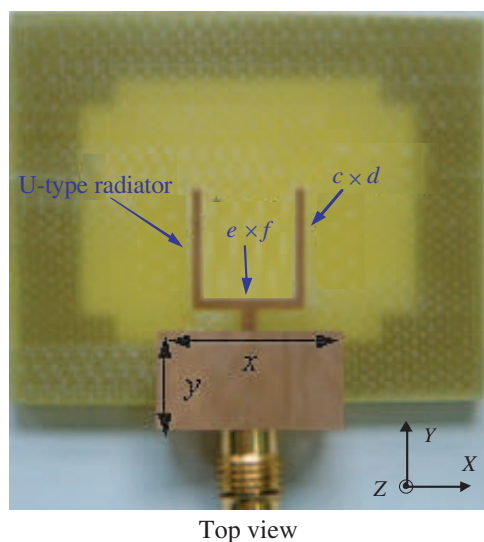
Figure 2: Configuration top view of the wideband slot antenna with a small parasitical ground plane.
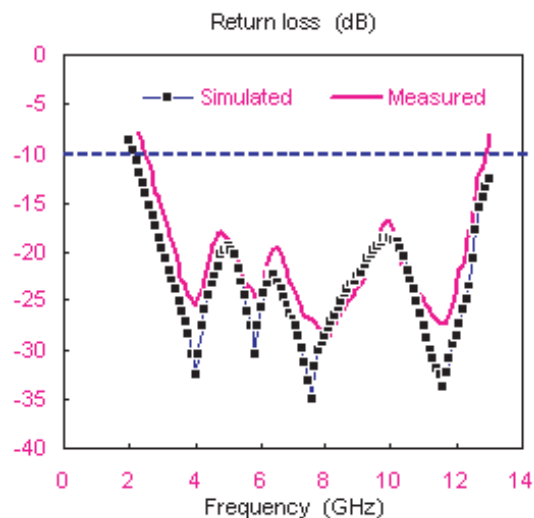

Figure 3: Measured and simulated return loss for the proposed antenna.

finite ground plane ($102\,\mathrm{mm} \times 76.2\,\mathrm{mm}$) and the parasitical ground plane, while the computations assume an infinite ground plane. For the $10\,\mathrm{dB}$ return loss, the measured maximum impedance bandwidth of the proposed antenna is from $2.3\,\mathrm{GHz}$ to $12\,\mathrm{GHz}$, corresponding to an impedance bandwidth of 136.6 percent centered at on $7.1\,\mathrm{GHz}$. Here, the impact of the finite finite ground plane is more pronounced. Therefore, the proposed antenna has good impedance bandwidth for the band of IEEE802.15.3a.

The radiation patterns of the proposed antennas with a finite ground plane are measured. Fig. 4 plot the measured radiation E-plane patterns patterned at on $3.1\,\mathrm{GHz}$ for the proposed antenna. From the measured test results, the finite ground plane does not dominate the radiation power patterns. Fig. 5 shows the measured gain versus the frequency. Within the operating frequency band, the gain of the proposed antenna is $0.04\,\mathrm{dBi}$ to $6.98\,\mathrm{dBi}$ for the antenna with finite ground plane.
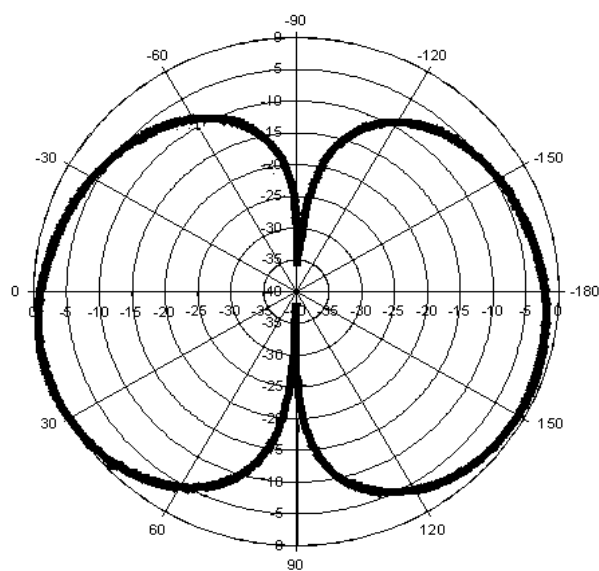
Figure 4: The experimental E-plane patterns of proposed antenna.
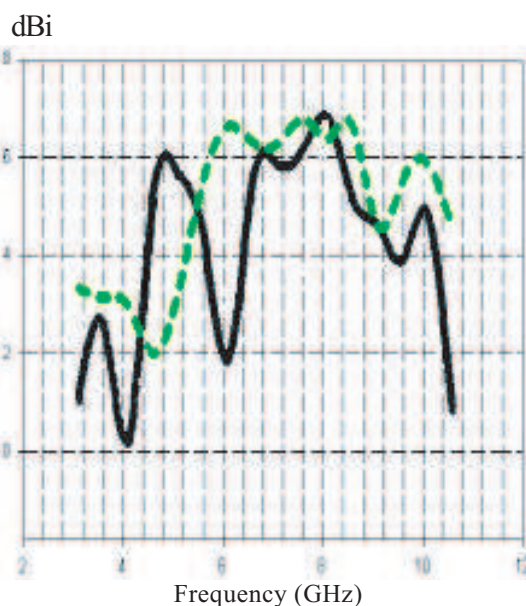
Figure 5: Measured and simulated gains versus frequency.

4. CONCLUSIONS

A new ultra-wideband antenna has been proposed for UWB applications. This The proposed antenna is uses employing a fork-like U-type radiating element structure with a finite ground plane and wide-slot ground plane for bandwidth enhancement. The simulated results conducted by the Zeland Software's IE3D simulation package show reasonable agreement with the measured results. The obtained results show an impedance bandwidth of more than 137% covering the whole IEEE802.15.3a defined UWB frequency band. Acceptable radiation characteristics at on the 3.1, 5.1, and 10.6 GHz, and the measureding antenna gain is 0.04 dBi to 6.98 dBi cover the $2 \sim 12$ GHz operation frequency ranges that makes this class of antennas a good candidate for a variety of the communication applications. Except for the enhancement of impedance bandwidth, the effects of the ground plane to the power patterns and peak gain value are small.

REFERENCES

1. Wong, K.-L., *Compact and Broadband Microstrip Antennas*, John Wiley and Sons Inc., New York, NY, 2002.
2. Lee, K. F., et al., "Experimental and simulation studies of the coaxially fed U-slot rectangular patch antenna," *IEE Proceedings — Microwave Antennas and Propagation*, Vol. 144, No. 5, 354–358, 1997.
3. Kahrizi, M., T. K. Sarkar, and Z. A. Maricevic, "Analysis of a wide radiating slot in ground plane of a microstrip line," *IEEE Trans. Microwave Theory Tech.*, Vol. 41, 29–37, Jan. 1993.
4. Tong, K.-F., K.-M. Luk, K.-F. Lee, and R. Q. Lee, "A broad-band U-slot rectangular patch antenna on a microwave substrate," *IEEE Trans. Antennas Propagat.*, Vol. 48, 954–960, Jun. 2000.
5. Chen, H.-D., "Compact circularly polarized microstrip antenna with slotted ground plane," *Electron. Lett.*, Vol. 38, Jun. 20, 2002.
6. Mak, C. L., K. M. Luk, and K. F. Lee, "Microstrip line-fed L-strip patch antenna," *IEE Proceedings — Microwave Antennas and Propagation*, Vol. 146, No. 4, 282–284, 1999.
7. Wong, K.-L., C. H. Wu, and S. W. (Stephen) Su, "Ultra wide-band square metal-plate monopole antenna with a trident-shaped feeding strip," *IEEE Trans. Antennas Propagat.*, Vol. 53, No. 4, 1262–1269, Apr. 2005.
8. Lin, X.-C. and L.-T. Wang, "A broadband CPW-fed loop slot antenna with harmonic control," *Antennas and Wireless Propagation Letters*, Vol. 2, 323–325, 2003.
9. Chen, H.-D., "Broadband CPW-fed square slot antennas with a widened tuning stub," *IEEE Trans. Antennas Propagat.*, Vol. 51, 1982–1986, Aug. 2003.
10. Leung, K. W. and C. K. Leung, "Wideband dielectric resonator antenna excited by cavity-backed circular aperture with microstrip-line tuning fork," *Electron. Lett.*, Vol. 39, 1033–1035, Jul. 10, 2003.
11. Sze, J.-Y. and K.-L. Wong, "Bandwidth enhancement of a microstrip-line-fed printed wide-wide-slot antenna," *IEEE Trans. Antennas Propagat.*, Vol. 49, 1020–1024, Jul. 2001.
12. Chiou, J.-Y., J.-Y. Sze, and K.-L. Wong, "A broad-band CPW-fed strip-loaded square slot antenna," *IEEE Trans. Antennas Propagat.*, Vol. 51, 719–721, Apr. 2003.

Centerline Longitudinal Shunt Slot Excitation by Parabolic Shaped Single Ridge Waveguide

M. Moradian, **M. Khalaj-Amirhosseini**, **and M. Tayarani**

Department of Electrical Engineering, Iran University of Science and Technology
Narmak, Tehran 16846-13114, Iran

Abstract— The Application of parabolic shaped ridge waveguide is presented for excitation of a centreline longitudinal shunt slot antenna. The proposed method is realized by adding parabolic shaped to the ridge of the single ridge waveguide. The parabolic shaped ridge is placed exactly under the longitudinal slot and the slot is placed along centreline of the waveguide. It is shown that for a typical slot, the slot normalized conductance can be increased by increasing the parabolic shaped depth. The simulation results show that the proposed structure can be considered as a proper candidate for replacing the conventional longitudinal shunt slot.

1. INTRODUCTION

Longitudinal shunt slot antenna has been frequently used due to its pure linear polarization. However, despite its advantages, the conventional longitudinal slot antenna has several shortcomings. Its basic drawback is the appearance of second-order beams off the principal planes which are usually called butterfly lobes [1, 2]. These second-order beams decrease the antenna efficiency and increase the susceptibility of the radar to jamming. Therefore, it is desirable to eliminate these second-order beams.

The second-order beams are associated with the alternating offset of successive slots which is essential in order to avoid the grating lobe appearance. Several methods have been proposed in the literature to suppress the second-order beams [2–5].

Another disadvantage of this antenna is its long width. The width of the conventional longitudinal waveguide slot antenna is about $0.7\lambda_0$ at the design frequency. Therefore, for some applications which are required large scanning angle in E-plane, the conventional slotted waveguide antenna is not a good candidate. This shortcoming has been overcome by utilizing ridge waveguide fed slot antenna [6].

Application of parabolic shaped ridge waveguide fed slot arrays is presented. The proposed structure is realized by adding the parabolic shaped to the ridge of the single ridge waveguide. The parabolic shaped ridge is placed exactly under the slot. It can be shown that the radiation characteristic of the slot can be varied by changing both slot length and parabolic shaped depth. Applying the proposed method not only eliminate the required offset but also has the advantage of ridge waveguide in reducing the required waveguide width. The following sections introduce the geometry of the proposed structure and discuss more about the radiation characteristics of the proposed slot antenna.

2. PROPOSED STRUCTURE

Figure 1 shows the proposed structure which is considered to be a proper candidate for replacing the conventional waveguide longitudinal slot antenna. As the figure shows, a longitudinal narrow slot is cut in the upper broad wall of a ridge waveguide. The slot is placed exactly along the centreline of the waveguide. Parabolic shaped is applied to some part of the ridge which is placed under the longitudinal slot.

The electromagnetic field of the dominant mode and further the current at the upper wall of rectangular waveguide is deviated from the center line due to the parabolic shaped ridge. Therefore, some current interrupt the slot and the displacement current in the slot is formed. The displacement current and its equivalent magnetic current sheet justify the radiation to the outer space.

Figure 2 shows the proposed and conventional slots. As indicated by the figure, the conventional waveguide has a wider width in comparison with the width of the ridge waveguide in the proposed slot. That is due to the effect of the ridge on the characteristics of the ridge waveguide.

The proposed parabolic shaped ridge has middle symmetry and the width of the parabolic shaped ridge is equal to the width of the initial ridge. The ridge height has also some effect on the propagation characteristic of the ridge waveguide and the radiation characteristics of the slot. But in this paper, only a constant value of the ridge height is taken into account. As the height of the

ridge increased the radiated power is also increased, while the power handling capability and phase velocity of the dominant mode are decreased.
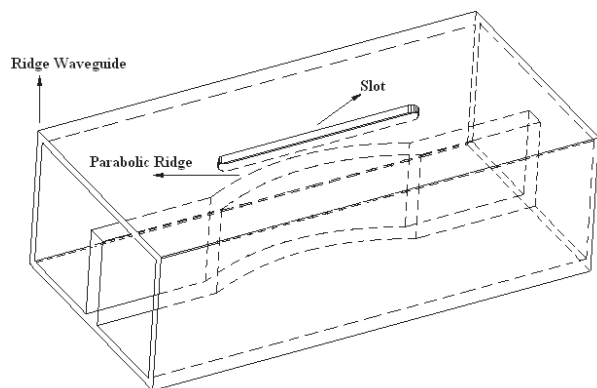

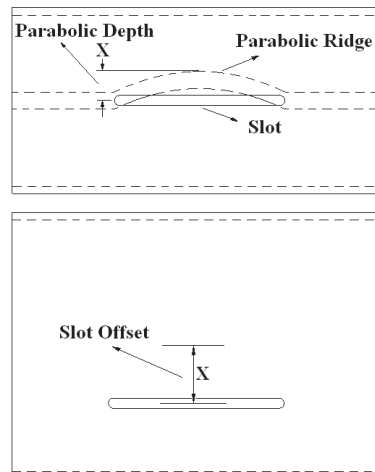Figure 1: Proposed structure for excitation of the centred longitudinal slot.

Figure 2: Top views of the proposed and conventional slot antenna.

3. SIMULATION RESULTS

In order to investigate the characteristics of the proposed structure, the proposed slot antenna with various parabolic shaped depths has been simulated using HFSS. The dimensions of the proposed slot antenna are shown in Table 1. The simulation of the antenna has been done for slot length equal to 45 mm and parabolic shaped depths equal to 6, 8, 12 and 17 mm.

Table 1: Dimensions of the proposed antenna.

Inner waveguide width:	50.4 mm
Inner waveguide height:	31 mm
Wall thickness:	2 mm
Ridge height:	22 mm
Ridge width:	3 mm
Slot width:	3 mm

For comparative reasons, a conventional longitudinal slot with the same specification is also simulated. The slot offsets are selected to be equal to the parabolic shaped depths, i.e., $X = 6$, 8, 12 and 17 mm. The various dimensions of the conventional longitudinal slot are shown in Table 2.

Table 2: Dimensions of the proposed antenna.

Inner waveguide width:	72.14 mm
Inner waveguide height:	34 mm
Wall thickness:	2 mm
Slot width:	3 mm

The normalized conductance and susceptance variation of the conventional and proposed slot versus slot offsets and parabolic shaped depths are shown in Figs. 3 to 6. It is indicated form the figures that the shift of the resonance frequency versus slot offsets in the conventional slot is larger than the shift of the resonance frequency in the proposed slot, while the variation of the magnitudes and the shapes of the normalized conductances are comparable.
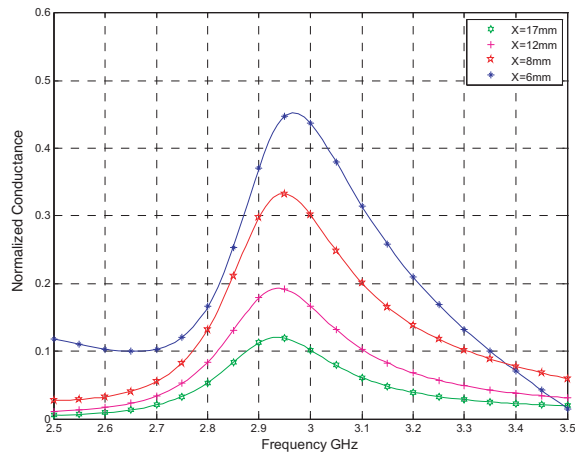
Figure 3: Normalized conductance variation of the proposed slot versus parabolic shaped depth and frequency for slot length equal to 45 mm.
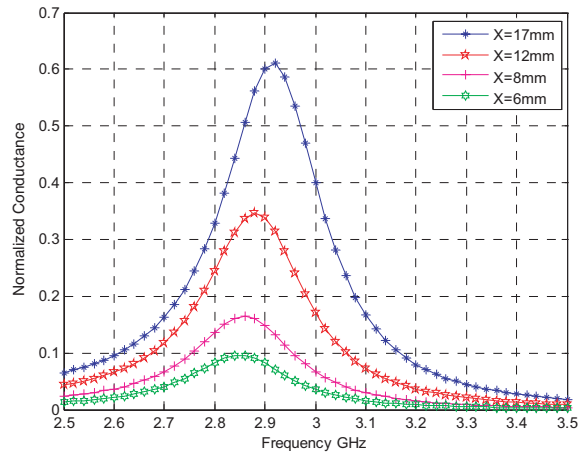

Figure 4: Normalized conductance variation of the conventional longitudinal slot versus slot offset and frequency for slot length equal to 45 mm.
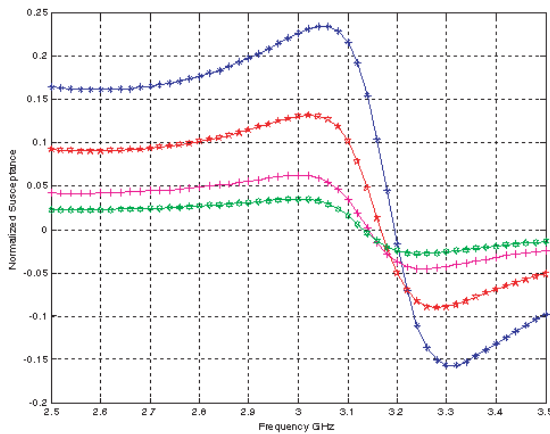

Figure 5: Normalized susceptance variation of the conventional slot versus slot offset and frequency for slot length equal to 45 mm.
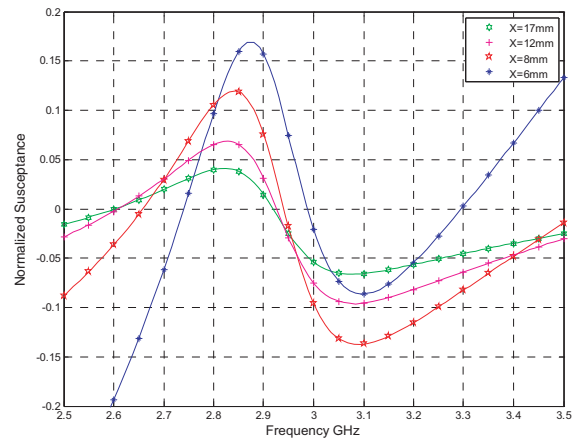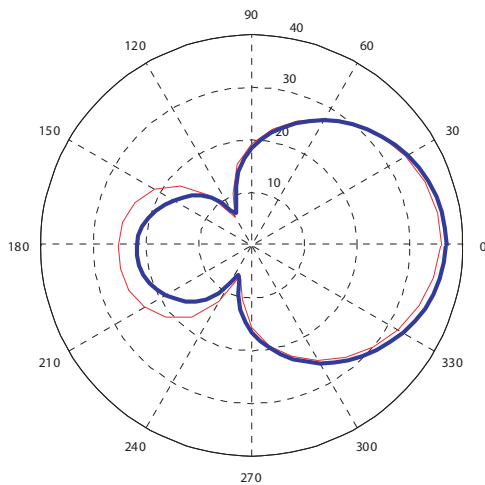

Figure 6: Normalized susceptance variation of the proposed longitudinal slot versus parabolic shaped depth and frequency for slot length equal to 45 mm.

Figure 7: E-plane radiation pattern of the proposed (solid line) and conventional slots (thick solid line).
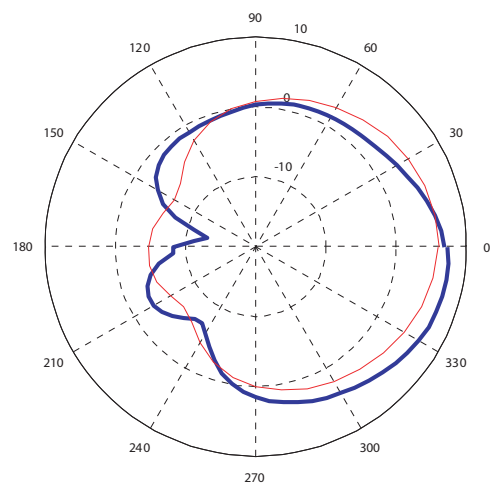
Figure 8: H-plane radiation pattern of the proposed (solid line) and conventional slots (thick solid line).

For better comparison, the radiation pattern of the proposed and conventional slot in E- and H-plane are shown in Figs. 7 and 8, respectively.

It is confirmed from the figures that the radiation pattern of the proposed slot has better symmetry in comparison to the conventional slot, while the front to back ratio of the proposed slot is reduced. That is, due to the reduction of the ground plane width of the slot (The width of the waveguide).

The simulation results show the effectiveness of the proposed slot and its ability for replacing the conventional longitudinal slot.

4. CONCLUSIONS

A modified configuration has been proposed for replacing conventional longitudinal slot. The slot was cut exactly along the centerline of a typical single ridge waveguide. The parabolic shaped was added to the ridge part of the ridge waveguide just under the slot position. The density of electromagnetic field was deviated from the centerline by the parabolic shaped part of the ridge waveguide as some surface current was interrupted by the slot and the radiation occurs. Comparison of the simulated parameters of the conventional and proposed topology indicates the efficacy of the proposed structure and its capability for replacing the conventional longitudinal slot.

REFERENCES

1. Kurtz, L. A. and J. S. Yee, "Second-order beams of two dimensional slot arrays," *IRE Trans.*, Vol. AP-5, 356–36, 1957.

2. Forooraghi, K. and P.-S. Kildal, "Transverse radiation pattern of a slotted waveguide array radiating between finite height baffles in terms of a spectrum of two-dimensional solutions," *Proc. Inst. Elec. Eng., Pt. H, Microwaves Antennas and Propagation*, Vol. 140, No. 1, 52–58, Feb. 1993.

3. Guenberg, H., "Second-order beams of slotted waveguide arrays," *Can. J. Phys.*, Vol. 31, 55–69, 1953.

4. Dudley, D. G., "An iris-excited slot radiator in the narrow wail of a rectangular waveguide," *IRE Trans.*, Vol. AP-9, 361–364, 1961.

5. Rengarajan, S. R., "Mutual coupling between waveguide-fed transverse broad wall slots radiating between baffles," *Electromagnetics*, Vol. 17, No. 5, 421–435, 1997.

6. Kim, D. Y. and R. S. Elliott, "A design procedure for slot arrays fed by single-ridge waveguide," *IEEE Trans. Antennas Propag.*, Vol. 36, No. 11, 1531–1536, Nov. 1988.

7. Elliott, R. S., "An improved design procedure for small arrays of shunt slots," *IEEE Trans. Antennas Propagat.*, Vol. 31, No. 1, 48–53, Jan. 1983.

Planar Antenna Array Mutual Coupling Identification: A Direct Method Applied to Quasi-Yagi Elements

C. E. Capovilla[1], **A. Tavora A. S.**[1], **S. E. Barbin**[2,3], **and L. C. Kretly**[1]

[1]Department of Microwaves and Optics, School of Electrical and Computer Engineering
University of Campinas, Brazil
[2]Department of Telecommunications and Control Engineering, Polytechnic School
University of Sao Paulo, Brazil
[3]Center for Information Technology Renato Archer, Campinas, Brazil

Abstract— Mutual coupling between adjacent elements is a critical issue for the design of antenna arrays. In this way, it is important to know the influence that each type of coupling mechanism represents in the total mutual coupling. With this objective for a quasi-Yagi antenna array, four prototypes with special details were built and tested. Surface waves and free space coupling were analyzed aiming at proposing a better optimization procedure for the structure. Concluding, FDTD simulation results for the currents distribution on the surface of the array are presented.

1. INTRODUCTION

Modern wireless communication systems demand antennas with small dimensions, reasonable gain over a wide band, and offering the possibility of multi-band operation.

Satisfying all these requirements, the quasi-Yagi antenna is a serious candidate for this kind of application. It was introduced as a class of antennas, which presents better characteristics than many others [1]. These characteristics include both, the compactness of resonant-type antennas and the broadband behavior of a travelling-wave radiator.

These attributes strongly suggest the application of quasi-Yagi elements in antenna arrays. A simple quasi-Yagi element operating alone in free space presents a wideband, high efficiency, moderate gain, among other characteristics [2]. Its radiation pattern shows an excellent stability over the entire operational frequency band.

These characteristics together denote it as an ideal element for applications in systems for which directional beams and reconfigurable radiation patterns are needed. All these favorable conditions show that quasi-Yagi antennas have great potential for applications on phased arrays, switched systems, and, consequently, for adaptive systems.

Using a variety of processing algorithms, usually managed by a DSP, an adaptive antenna array can adjust and update its radiation pattern to enhance the desired signal, eliminate or reduce interference, and collect correlated multipath signal [3].

The radiation pattern can be configured according to the environment in real time, if control algorithms as MUSIC (Multiple Signal Classification) or ESPRIT (Estimation of Signal Parameters via Rotational Invariance Techniques) are used.

The adaptive array is also necessary for capacity and range improvement [4], due to the need of finding users in a cell and directing the main beam to it. In this way, the same channel can be used more than a time in the same cell. Polarization diversity can be used to differentiate two very close users in a cell. Space diversity is fundamental to avoid signal fading due to the multipath, caused by reflections of the signal, mainly when the signal bandwidth is small compared to the coherence channel bandwidth.

During a planar antenna array design, it become necessary to analyze the mutual coupling effects between adjacent elements [5], thus determining the way each type of coupling influences the total mutual coupling.

This work is organized in five sections. Following this brief introduction, Section 2 presents the quasi-Yagi antenna and the assumptions made for building the array under investigation. Section 3 discusses experimental results of the mutual coupling mechanism analysis applied to two adjacent antennas. In Section 4, a FDTD simulation of the currents that are functions of the mutual coupling is presented. Finally, Section 5 presents the conclusions of the proposed analysis.

2. QUASI-YAGI ANTENNA ARRAY

The antenna array prototype was fabricated using FR4 with $\varepsilon_r = 4.4$ and 1.6 mm thickness. This material was chosen for the substrate due to its good electrical characteristics, easy acquisition, and low cost. The prototype was designed to operate at the ISM band 2.4–2.5 GHz (WLAN applications). A schematic of the quasi-Yagi antenna is shown in Fig. 1, where its main elements can be seen: the feeding line, MS-to-CPS balun, CPS line, reflector, driver, and director. A half wavelength delay line introduced in one of the arms of the balun is required to obtain the odd mode coupling [6].

The design and optimization of this antenna is reported in [7] with bandwidth results of 44% and 41%, for FDTD simulation and measurement, respectively. These results illustrate that the quasi-Yagi antenna has a bandwidth like a horn antenna, although having only a fraction of its weight and size.
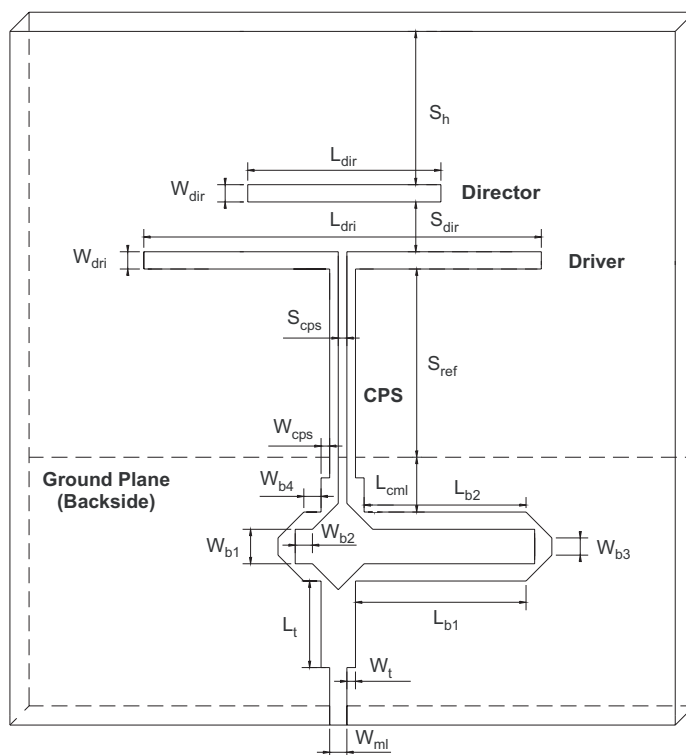
Figure 1: Schematic of the quasi-Yagi antenna.

The antenna final dimensions, after an interactive optimization process, are (unit: millimeters): $W_{ml} = W_{dri} = W_{dir} = 2.86$, $W_t = 1.27$, $W_{cps} = 1.35$, $W_{b1} = 5.40$, $W_{b2} = W_{b3} = W_{b4} = 2.70$, $L_{dri} = 62.43$, $L_{dir} = 30.00$, $L_{b1} = 26.75$, $L_{b2} = 25.40$, $L_{cml} = 5.35$, $L_t = 13.55$, $S_{dir} = 7.80$, $S_{ref} = 29.50$, $S_{cps} = 1.35$, and $S_h = 24.00$.

The antenna array with two elements was exhaustively simulated, changing the distance between the adjacent elements aiming at reaching an optimal performance. This parameter is critical for its influence in the mutual coupling and the level of secondary lobes [8]. Small spaces between the elements generate strong mutual coupling degenerating the array overall performance, while great spaces favor the appearance of secondary lobes in the radiation pattern [4].

A distance of $\lambda_0/2$ is considered ideal for adjacent elements in antenna arrays, as demonstrated in a previous work, for a patch array [9], as well as, for a PIFA (Planar Inverted F Antenna) array [10]. In this work, the $\lambda_0/2$ distance was adopted between the extremities of the adjacent drivers.

3. ANALYSIS OF THE MUTUAL COUPLING MECHANISM: A PRACTICAL METHODOLOGY

With the objective of finding which mechanisms contribute to the mutual coupling between adjacent quasi-Yagi elements, aiming at better optimization of array structures, several tests were conducted.

These tests used arrays of two identical antennas placed in line, side by side. As a measuring reference for the mutual coupling, S_{21} was used.

The possible coupling mechanisms are:

- Coupling by superficial waves excited by the feeding lines of the adjacent elements;
- Coupling by superficial waves excited directly by the antennas;
- Coupling through free space.

Four prototypes were made for the experiments, each one with two antennas. The first prototype is a conventional array, without any modification in the structure.

The second prototype, designed for analyzing the coupling of superficial waves excited by the feeding lines, has a cut in the ground plane between the elements, in order to suppress the waves. The ideal opening length was experimentally found to be 15 mm.

The third prototype was designed to test the mutual coupling due to superficial waves excited directly by the antennas. The coupling was eliminated by making a cut of 15 mm in the substrate between the adjacent antennas up to the height of the truncated ground plane.

Finally, the last prototype was designed to test both effects together and has openings in the ground plane and in the substrate, as shown in Fig. 2. The mutual coupling effects were observed using a ZVRE network analyzer from Rhode & Schwarz, producing the results in Fig. 3.
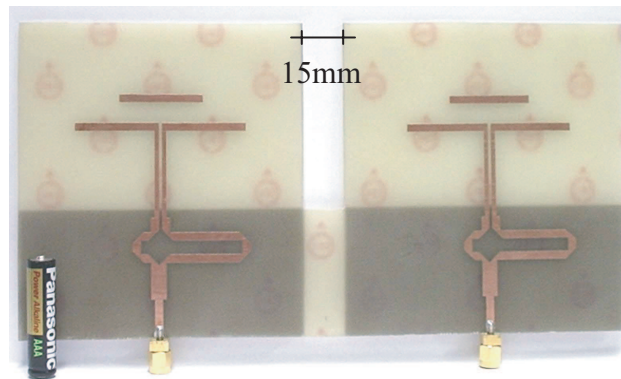
Figure 2: Array under test with both cuts, on ground plane and into the substrate between the antennas.
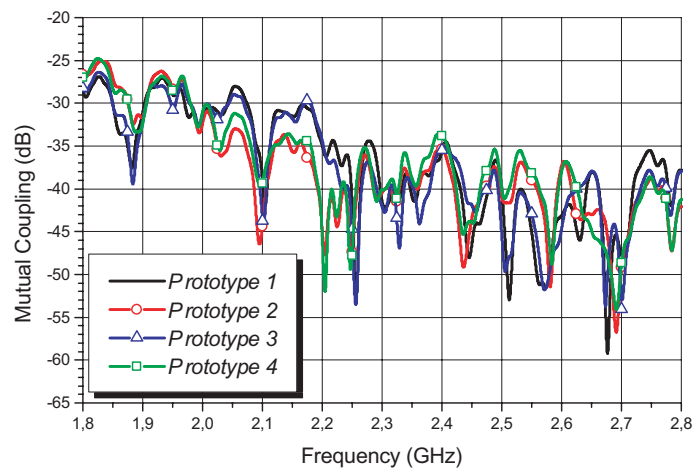

Figure 3: Measured mutual coupling of the arrays.

Analyzing these results at the operation band, only small differences between the measurements of the prototypes are observed. The first and the third prototype have an unchanged ground plane, and the difference between them is the cut on the substrate up to the truncated ground plane that the third prototype has. The behavior of the first and the third prototypes is the same during the whole operation band. This means that the superficial waves excited directly by the antennas do

not significantly influence the mutual coupling of the array. The analysis of the second and fourth prototype shows that they exhibit the same behavior, as expected, presenting only small differences when compared to the first and the third prototype.

Therefore, the conclusion is that mutual coupling happens mainly through the free space. The superficial waves coming directly from the antennas do not have any significant interference on the array, and the superficial waves coming from the feeding lines only result on a slight perturbation of the mutual coupling level.

In this way, no changes in the elements structure are necessary for the design of a quasi-Yagi antenna array, only their correct positioning and appropriate feeding lines are required.

4. CURRENTS ON THE SURFACE OF A QUASI-YAGI ANTENNA ARRAY

In order to determine the surface currents on the array numerically, an FDTD code was used. The QFDTD90 code by Realize Inc. was chosen, since it allows easy access to the surface currents for post-processing and analysis. The size of the FDTD grid in the H plane is approximately the same as in the E plane. The relatively large thickness of the substrate allows a significant amount of current to flow on the substrate as surface waves, increasing the mutual coupling. The steady-state current profile is shown in Fig. 4 for an array with two quasi-Yagi elements. This picture illustrates the surface currents due to two elements of the array, where each element is excited with currents with the same amplitude and phase.
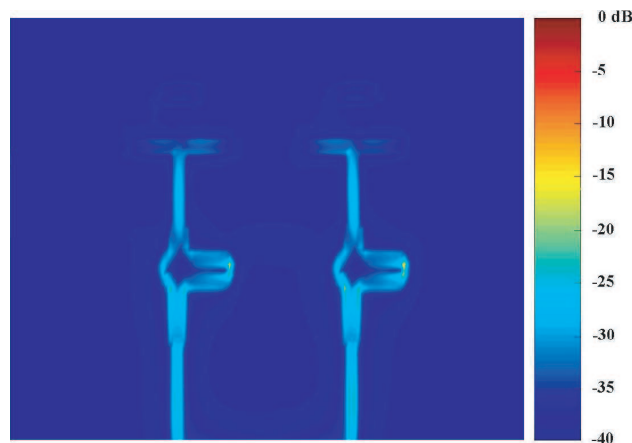

Figure 4: Current density distribution of the array.

The expected out-of-phase field distribution in both the coupled microstrip and the CPS feeding line is clearly seen. It is also interesting to observe that the director dipole is fairly strong excited, and in opposite phase to the drive dipole. While these results are more qualitatively rather than quantitative, they do provide additional insights into the electromagnetic phenomenon of the mutual coupling of the adjacent antennas in the array.

5. CONCLUSIONS

This work presented an experimental analysis of the mutual coupling mechanism on a quasi-Yagi antenna array. No change in the elements structure design is necessary for building an array prototype, but only the correct positioning of the elements and feeding is required. In addition, FDTD simulation results of the current distribution on the array surface FDTD were presented to complete the mutual coupling analysis.

REFERENCES

1. Deal, W. R., N. Kaneda, J. Sor, Y. Qian, and T. Itoh, "A new quasi-yagi antenna for planar active antenna arrays," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 48, No. 06, 910–418, June 2000.
2. Kaneda, N., W. R. Deal, Q. Yongxi, R. Waterhouse, and T. Itoh, "A broad-band planar quasi-yagi antenna," *IEEE Transactions on Antennas and Propagation*, Vol. 50, No. 08, 1158–1160, August 2002.
3. Liberti, J. C. and T. S. Rappaport, *Smart Antennas for Wireless Communications IS-95 and Third Generation CDMA Applications*, Prentice Hall PTR, 1999.

4. Petrus, P., "Novel adaptive array algorithms and their impact on cellular system capacity," Doctor of Philosophy in Electrical Engineering, Faculty of the Virginia Polytechnic Institute, March 1997.

5. Jedlicka, R., M. Poe, and K. Carver, "Measured mutual coupling between microstrip antennas," *IEEE Transactions on Antennas and Propagation*, Vol. 29, 147–149, 1981.

6. Song, H. J., M. E. Bialkowski, and P. Kabacik, "Parameter study of a broadband uniplanar quasi-yagi antenna," *13th International Conference on Microwaves, Radar and Wireless Communications*, Vol. 01, 166–169, 2000.

7. Kretly, L. C. and C. E. Capovilla, "Analysis of radiation patterns and broad-band characteristics of a novel quasi-yagi antenna array for wireless communications," *IEEE International Symposium on Wireless Personal Multimedia Communications*, Vol. 03, 440–443, October 2003.

8. Balanis, C. A., *Antenna Theory: Analysis and Design*, 2nd ed., John Wiley Sons, 1997.

9. Kretly, L. C., A. C. Sodre, and A. S. Tavora, "A hexagonal adaptive antenna array concept for wireless communication applications," *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, September 2002.

10. —, "Triangle PIFA antenna array prototype for wireless system applications," *IEEE International Telecommunications Symposium*, September 2002.

Study of High T_c Superconducting Microstrip Antenna

T. Fortaki[1], M. Amir[1], S. Benkouda[1], and A. Benghalia[2]

[1]Electronics Department, University of Batna, Algeria
[2]Electronics Department, University of Constantine, Algeria

Abstract— Resonant characteristics of microstrip antennas with superconducting films is presented. The analysis is based on a full electromagnetic wave model with London's equations and the two-fluid model. It is shown that the full-wave analysis presented here gives numerical results which are in excellent agreement with the measured data available in the literature. Results showing the effect of the temperature on the resonant frequency and half-power bandwidth of superconducting microstrip antenna are given. Variations of the resonant frequency with the high T_c superconducting film thickness are also presented.

1. INTRODUCTION

Superconducting passive microwave devices such as antennas, filters, transmission lines, and phase shifters have shown significant superiority over corresponding devices fabricated with normal conductors such as gold, silver, or copper due to the advantages of superconductors [1]. Advantages of using high T_c superconducting materials at high frequencies include [2]: 1) very small losses, which means reduction of attenuation and noise level; 2) very small dispersion up to frequencies of several tens of GHz; 3) smaller devises due to the lower losses, which leads to larger integration density; and 4) the propagation time can be greatly reduced because of the smaller size and the shorter interconnects.

In the literature, the studies concerning the resonance characteristics of microstrip antennas using perfectly conducting patches are abundant. However, few works have been done for the case of microstrip antennas using superconducting patches. The determination of the resonant frequencies of superconducting microstrip patch antennas was initially carried out by means of the magnetic wall cavity model [3]. Later on, these resonant frequencies were obtained by using the rigorous full-wave analysis [4]. To validate the theoretical analysis, the authors in [4] have compared their numerical results with the experimental data of Richard et al. [3]. This comparison has not been done in a convenient way for two reasons: the variation of the permittivity of the lanthanum aluminate substrate with the variation of the temperature, as indicated by the experiment of Richard et al. [3], has not been take into account by Silva et al. [4] and the effect of varying the temperature on the resonant frequency is insignificant.

In this paper, we present a theoretical and numerical analysis of the resonant frequencies of high T_c superconducting rectangular microstrip antennas which yields excellent agreement with the measured data of Richard et al. [3]. To include the effect of the superconductivity of the microstrip patch in the full-wave analysis, a surface complex impedance is considered. This impedance is determined by using London's equation and the two-fluid model of Gorter and Casimir [2]. Numerical results for the effect of the temperature on the resonant frequency and half-power bandwidth of superconducting microstrip antenna are given. Finally, the influence of the thickness of the high T_c superconducting film on the resonant frequency is also presented.

2. THEORY

The problem to be solved is illustrated in Figure 1. We have a rectangular superconducting patch of thickness e printed on a dielectric layer. The dielectric layer of thickness d is characterized by the free-space permeability μ_0 and the permittivity $\varepsilon_0\varepsilon_r$ (ε_0 is the free-space permittivity and the relative permittivity ε_r can be complex to account for dielectric loss). The superconducting patch is characterized by a critical temperature T_c, a zero-temperature penetration depth λ_0, and a normal state conductivity σ_n. Following a mathematical reasoning similar to that shown in [5, Equations (2)–(14)] for obtaining a relation among the surface electric field at the plane of the superconducting patch and the surface current on the patch in the spectral domain given by

$$\begin{bmatrix} \tilde{E}_x \\ \tilde{E}_y \end{bmatrix} = \begin{bmatrix} Q_{xx} & Q_{xy} \\ Q_{yx} & Q_{yy} \end{bmatrix} \cdot \begin{bmatrix} \tilde{J}_x \\ \tilde{J}_y \end{bmatrix} \tag{1}$$

where Q_{xx}, Q_{xy}, Q_{yx}, and Q_{yy} are the components of the spectral dyadic Green's function. They are given in [5, Equation (14)]. The surface electric fields at the plane of the superconducting patch can be written as a superposition of an electric field in the patch and another out of the patch, this yields

$$\begin{cases} \tilde{E}_x = \tilde{E}_x^i + \tilde{E}_x^o \\ \tilde{E}_y = \tilde{E}_y^i + \tilde{E}_y^o \end{cases} \tag{2}$$

The electric field in the superconducting patch is given by

$$\begin{cases} \tilde{E}_x^i = Z_s \, \tilde{J}_x \\ \tilde{E}_y^i = Z_s \, \tilde{J}_y \end{cases} \tag{3}$$

where Z_s is the surface impedance of the superconducting patch. When the thickness of the superconducting patch is less than three times the penetration depth λ at a temperature $T = 0\,\mathrm{K}$ (λ_0), the surface impedance can be approximated as in [4, Equation (28)]. Substituting Equations (2) and (3) in Equation (1) yields

$$\begin{bmatrix} \tilde{E}_x^o \\ \tilde{E}_y^o \end{bmatrix} = \begin{bmatrix} Q_{xx} - Z_s & Q_{xy} \\ Q_{yx} & Q_{yy} - Z_s \end{bmatrix} \cdot \begin{bmatrix} \tilde{J}_x \\ \tilde{J}_y \end{bmatrix} \tag{4}$$

The transverse electric fields out of the patch can be expressed via the inverse Fourier transform as follows:

$$E_x^o = \frac{1}{4\pi^2} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \left[(Q_{xx} - Z_s)\tilde{J}_x + Q_{xy}\tilde{J}_y \right] \exp\left(ik_x x + ik_y y\right) dk_x dk_y \tag{5a}$$

$$E_y^o = \frac{1}{4\pi^2} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \left[Q_{yx}\tilde{J}_x + (Q_{yy} - Z_s)\tilde{J}_y \right] \exp\left(ik_x x + ik_y y\right) dk_x dk_y \tag{5b}$$

Enforcement of the boundary condition requiring the transverse electric fields of (5a) and (5b) to vanish on the area of the superconductor patch, yields the sought integral equations

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \left[(Q_{xx} - Z_s)\, \tilde{J}_x + Q_{xy}\tilde{J}_y \right] \exp\left(ik_x x + ik_y y\right) dk_x dk_y = 0, \quad (x, y) \in \text{patch} \tag{6a}$$

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \left[Q_{yx}\tilde{J}_x + (Q_{yy} - Z_s)\, \tilde{J}_y \right] \exp\left(ik_x x + ik_y y\right) dk_x dk_y = 0, \quad (x, y) \in \text{patch} \tag{6b}$$

Now, the Galerkin moment method can be easily applied to Equations (6a) and (6b) to obtain the resonant frequencies and the bandwidths of the resonant modes of the high T_c superconducting microstrip antenna shown in Figure 1.
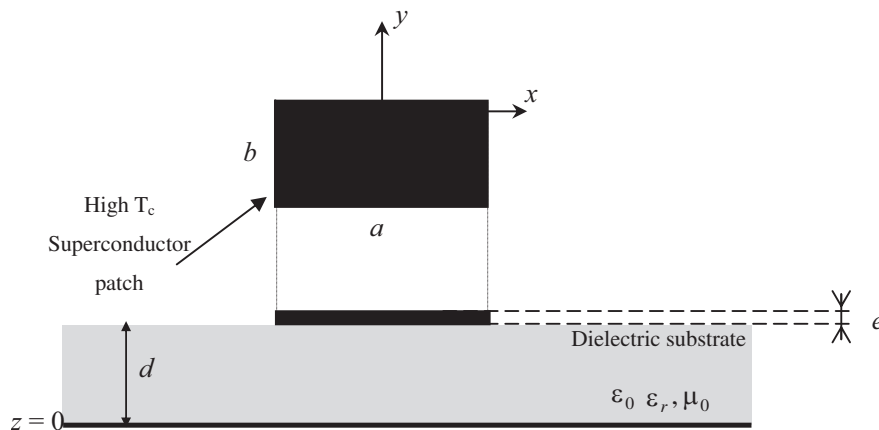

Figure 1: Geometry of a high T_c superconducting microstrip antenna.

3. NUMERICAL RESULTS AND DISCUSSION

Numerical results are obtained for the parameters used in the experiment of Richard et al. [3]. A rectangular patch with a thickness 350 nm, a resonant length 935 μm and a width of 1630 μm is printed on a lanthanum aluminate substrate of thickness 254 μm. The patch is fabricated using a YBCO superconducting thin film with a transition temperature $T_c = 89$ K, a zero-temperature penetration depth $\lambda_0 = 140$ nm, and a normal state conductivity $\sigma_n = 10^6$ S/m. The lanthanum aluminate substrate was chosen in the experiment of Richard et al. [3] despite its high permittivity because it allows the growth of high-quality (low surface resistance) YBCO superconducting films [3]. Unlike the theoretical analysis of Silva et al. [4], the variation of the permittivity of the lanthanum aluminate substrate with the variation of the temperature, as indicated by the experiment of Richard et al. [3], is taken into account in the present paper. Figure 2(a) shows the resonant frequency versus the operating temperature. The measured results reported in [3] are also included in this figure for comparison purpose. Excellent agreement between theory and experiment is observed. From the results of Figure 2(a), it can be observed that increasing the temperature will decrease the resonant frequency. This decrease is significant for temperatures near the transition temperature. In Figure 2(b), we study the influence of the temperature on the bandwidth of the superconducting antenna. It is observed that the effect of varying the temperature on the bandwidth is significant only for temperatures near the critical temperature. The behavior of the bandwidth shown in Figure 2(b) is different from that given in [4, Figure 3], but it is in accordance with the measured one illustrated in [3, Figure 12].

The influence of the thickness of the high T_c superconducting film on the operating frequency is investigated in Figure 3. The superconducting film characteristics are: $\lambda_0 = 140$ nm, $\sigma_n = 10^6$ S/m, and $T_c = 89$ K. For the microstrip antenna, the following parameters are used: $a = 1630$ μm,
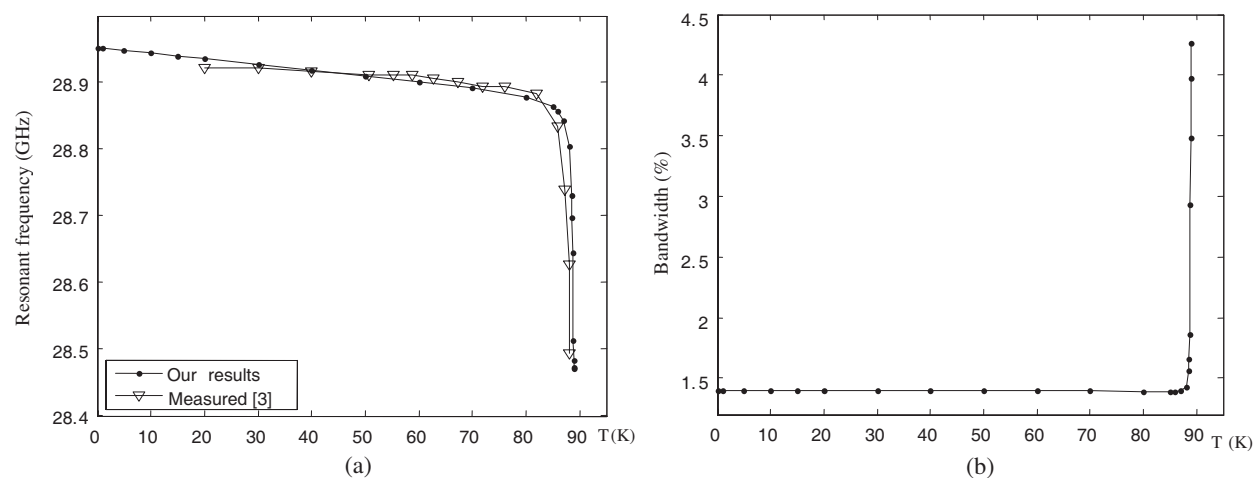

Figure 2: Resonant frequency and bandwidth of a high T_c superconducting antenna versus temperature.
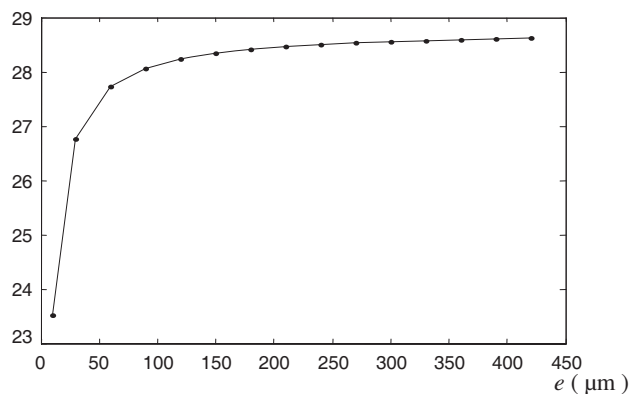

Figure 3: Resonant frequency versus the thickness of the superconducting film.

$b = 935\,\mu\text{m}$, $d = 254\,\mu\text{m}$, and $\varepsilon_r = 23.81$. The operating temperature is $T = 50\,\text{K}$. It is observed that when the thickness of superconducting film grows, the resonant frequency increases. Note that the effect of the thickness of the superconducting film is more pronounced for small values of e ($e < \lambda_0$). When e exceeds λ_0, increasing the superconducting film thickness will increase slowly the resonant frequency. Extreme care should be taken when designing an antenna with thin superconducting film; since small uncertainty when the patch is fabricated can result in an important shift of the resonant frequency.

4. CONCLUSIONS

This paper has presented a rigorous full-wave analysis of microstrip antennas with superconducting films. To include the effect of the superconductivity of the microstrip patch in the full-wave analysis, a surface complex impedance has been considered. This impedance has been determined using London's equation and the two fluid model. Our numerical results for the resonant frequency are in excellent agreement with the experimental data available in the literature. Numerical results for the effect of the temperature on the resonant frequency and half-power bandwidth have been presented. Variations of the resonant frequency versus the thickness of the superconducting film have been also presented.

REFERENCES

1. Sekiya, N., A. Kubota, A. Kondo, S. Hirano, A. Saito, and S. Ohshima, "Broadband superconducting microstrip patch antenna using additional gap-coupled resonators," *Physica C*, Vol. 445–448, 994–997, 2006.
2. El-Ghazaly, S. M., R. B. Hammond, and T. Itoh, "Analysis of superconducting microwave structures: application to microstrip lines," *IEEE Trans. Microwave Theory Tech.*, Vol. 40, No. 3, 499–508, 1992.
3. Richard, M. A., K. B. Bhasin, and P. C. Claspy, "Superconducting microstrip antennasd: An experimental comparison of two feeding methods," *IEEE Trans. Antennas Propagat.*, Vol. 41, No. 7, 967–974, 1993.
4. Silva, S. G., A. G. d'Assuncao, and J. R. S. Oliveira, "Analysis of high T_c superconducting microstrip antennas and arrays," *Proceedings of SBMO/IEEE MTT-S IMOC*, 243–246, 1999.
5. Fortaki, T. and A. Benghalia, "Efficient analysis of the far field pattern of rectangular microstrip patch using the stationary phase method," *Proceedings of 16th International Conference on Microelectronics*, 278–281, Tunis, Tunisia, 2004.

Directivity Enhancement of Microstrip Patch Antennas Using a Dieletric Superstrate

Yanfei Li[1], Raj Mittra[2], Guizhen Lu[1], and Wenhua Yu[1, 2]

[1]Information Engineering School, Communication University of China, Beijing 100024, China
[2]Electromagnetic Communication Laboratory, 319 EE East, The Pennsylvania State University
University Park, PA16802, USA

Abstract— In this paper, we present a systematic design methodology for enhancing the directivity of microstrip patch antennas (MPAs), by using it as an exciter of a Fabry-Perot (FP) [6] resonant cavity. The design procedure involves three steps: (i) determing the resonance frequency of the empty cavity, with the MPA removed, when operating in the receive mode; (ii) designing the patch antenna to operate at the same frequency in the presence of the superstrates; (iii) finally, we operate the composite system in the transmit mode to investigate its aperture distribution that provides us a clue for enhancing the directivity. We show that the directivity can be enhanced by more than 8 dB, over that of the MPA without the superstrate, by using either planar or curved superstrates as covers for the MPA.

1. INTRODUCTION

Resonator antennas based on Fabry-Perot (FP) cavities are low-profile directive radiators that are suitable for high-gain applications. In recent years the design of highly directive cavity-type antennas has attracted many researchers in the field of antenna technology [1, 2, 7]. In a previous work, Shen and Vandenbosch [3] have used the electric field distribution on top surface above a superstrate to determine its shape and size that minimize its edge diffraction. However, neither the issue of directivity variation with frequency is considered in the above work, nor is the phase variation in the aperture, which obviously controls the directivity, is investigated. Although the literature on FP resonators is quite extensive (see e.g., [4]), most of the prior works deal with large structures, whose lateral dimensions are large in terms of the wavelengths at the operating frequency of the antenna. To the best of our knowledge, there is only one paper that deals with reduced-size FP antennas unilizing Bragg-type dielectric mirrors [5].

In this paper, we consider the problem of enhancing the directivity of a microstrip patch antenna covered by a dielectric superstrate (DS), whose size is $2\lambda_0 \times 2\lambda_0$. We begin by studying the receive-mode characteristics of the FP resonator, and determine the vertical dimension of the cavity, comprising of the ground plane, the substrate and the superstrate, such that the E-field is maximum at the desired frequency. Next, we compare the directivity of the MPA with and without the DS, and show that the directivity of the MPA/DS composite can be enhanced by more than 7 dB at the design frequency of 6 GHz. Following this, we study the magnitude and phase distributions of the E-field at 6 GHz in the observation plane located about $0.5\lambda_0$ above the DS. Finally, we investigate the possibility of using curved superstrates to render the phase distribution more uniform at 6 GHz than it is for the planar case, and show that the directivity can be further enhanced by about 1.5 dB over that of the planar DS design.

2. RECEIVE MODE

Since our operating frequency, at which we wish to achieve the maximum directivity is 6 GHz, we choose the height of the FP cavity to be around 0.5λ at that frequency. The geometry for the receive-mode operation is illustrated in Fig. 1, and the cavity parameters are listed in Table 1. We set the location of the observation point at the center of the cavity, and monitor the normalized E-field magnitude, which is shown in Fig. 2. We found that the cavity resonates at 6 GHz, at which frequency the maximum of the spatial distribution of the E-field reaches a maximum inside the cavity as a function of frequency. The simulations have shown that the frequency at which maximum directivity of the antenna/superstrate composite is realized when operating in the transmit mode inside the cavity coincides with the frequency at which the field attains a maximum in the receive mode. This knowledge is helpful in providing us an important design guideline for this type of antenna/superstrate composite.

Table 1: Parameters of receive mode structure ($f_0 = 6$ GHz).

Superstrate ($\varepsilon_{r2} = 10.2$, $h_2 = 4.8$ mm, $2\lambda_0 \times 2\lambda_0$)

Air gap, $h = 25$ mm
Observation point (in the center of the cavity, $h_3 = 0.5h$)

Substrate ($\varepsilon_{r1} = 4.4$, $h_1 = 2.5$ mm, $2\lambda_0 \times 2\lambda_0$)

Ground plane (PEC, $2\lambda_0 \times 2\lambda_0$)
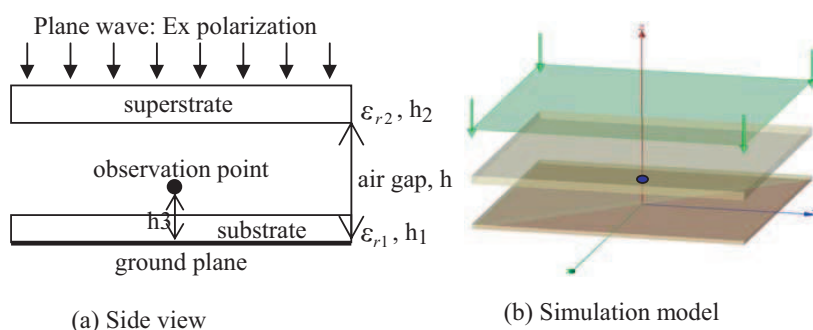
λ_0: the lambda of free space at 6 GHz

(a) Side view

(b) Simulation model

Figure 1: Structure of receive mode.
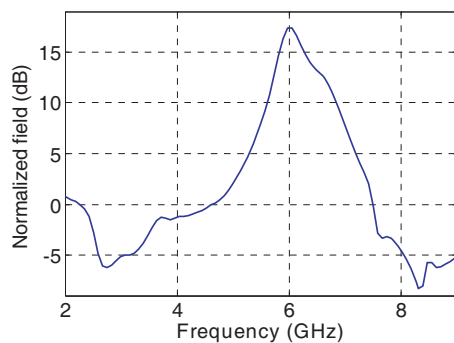
Figure 2: Distribution of normalize E-field magnitude.
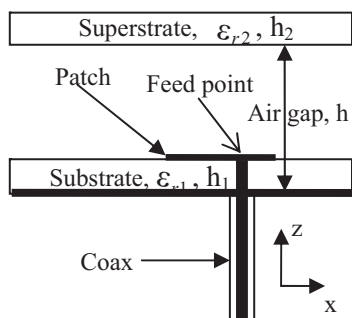
Figure 3: Side view of MPA with plane DS.
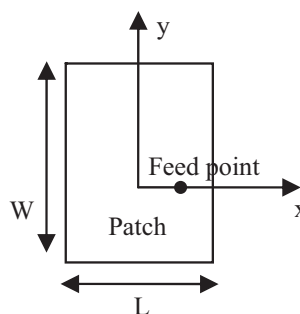
Figure 4: Top view of patch.

3. DIRECTIVITY ENHANCEMENT USING A PLANAR AND CURVED DIELECTRIC SUPERSTRATE (DS)

The MPA covered by a planar DS structure is shown in Fig. 3, and the top view of the patch, which operates at 6 GHz, is shown in Fig. 4. The parameters of this MPA/DS composite are presented in Table 2. We use GEMS, which is 3D high performance parallel EM simulation software [8], to simulate this structure. The variation of the directivity as well as of S_{11} with frequency is shown in Figs. 5(a) and (b), respectively. In Fig. 5(a), we compare the performance of the MPA with and without the DS. The results show that directivity of patch with a planar DS can reach 13.92 dB at 6 GHz, which is 7 dB above that of the MPA without a superstrate. The E-field distributions at 6 GHz on a plane, which is $0.5\lambda_0$ above the DS, are shown in Fig. 6. We observe that neither the magnitude nor the phase distribution on this plane is very uniform. Next, to improve the phase behavior and to realize a more uniform distribution, we replace the planar DS with a curved one whose height h is 26 mm (see Fig. 7), while leaving the other parameters unchanged. The curved superstrate includes a thin planar dielectric layer at the bottom, whose thickness is 1 mm. Its upper surface is spherical, whose radius h_r is $5.6\lambda_0$, and the height of the arch h_c is $0.09\lambda_0$. The schematic
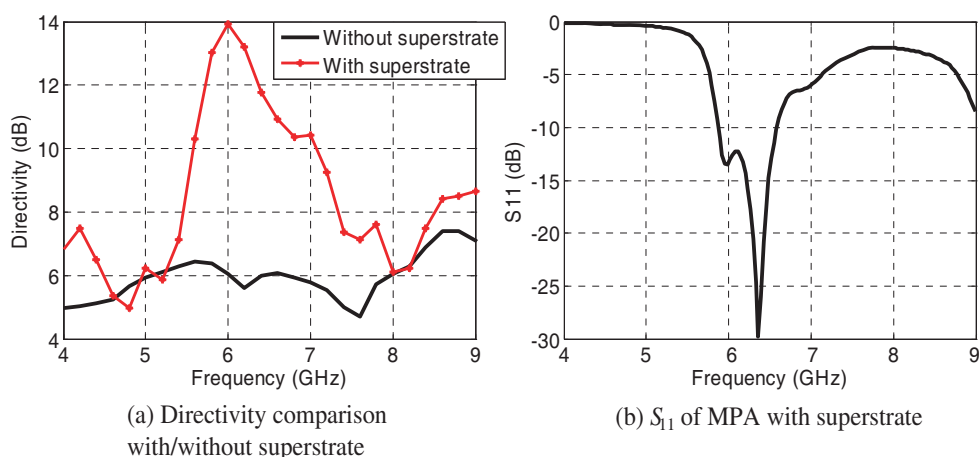

(a) Directivity comparison with/without superstrate

(b) S_{11} of MPA with superstrate

Figure 5: Simulation result with/without plane DS.

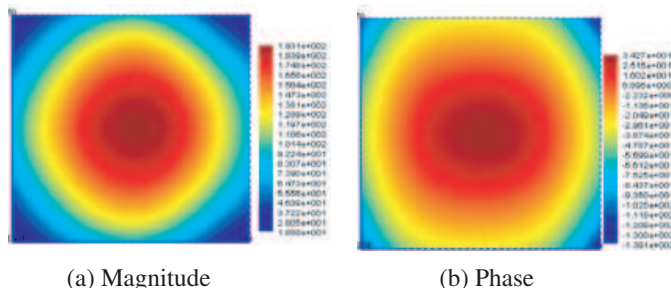(a) Magnitude   (b) Phase

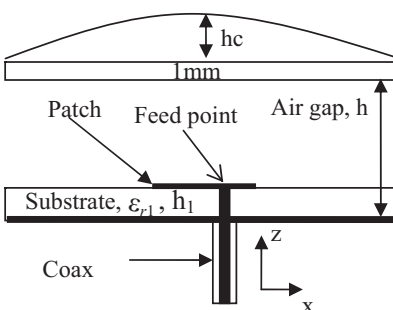Figure 6: E-field distribution of the plane $0.5\lambda_0$ above plane DS at 6 GHz.

Figure 7: Side view of MPA with curve DS.

diagram of MPA with the curved superstrate is shown in Fig. 7. The relative permittivity of both the thin dielectric layer as well as that of the spherical cap is 10.2.

The magnitude and phase distributions along the lines (1) and (2) are shown in Figs. 8 and 9, respectively, and the location of the observation plane is shown in Fig. 10. We observe that the phase distribution is more uniform for the curved superstrate than it is for its planar counterpart, though the same cannot be said for the magnitude distribution. We note, however, that we realize

Table 2: Parameters of MPA with plane DS ($f_0 = 6$ GHz).

Superstrate ($\varepsilon_{r2} = 10.2$, $h_2 = 4.8$ mm, $2\lambda_0 \times 2\lambda_0$)

Air gap, $h = 25$ m
Patch: $L = 9.7$ mm, $W = 14$ mm

Coax feeds at $x = 0.25L + 0.03\lambda_0$, $y = 0$, $z = h$

Substrate ($\varepsilon_{r1} = 4.4$, $h_1 = 2.5$ mm, $2\lambda_0 \times 2\lambda_0$)

Ground plane (PEC, $2\lambda_0 \times 2\lambda_0$)
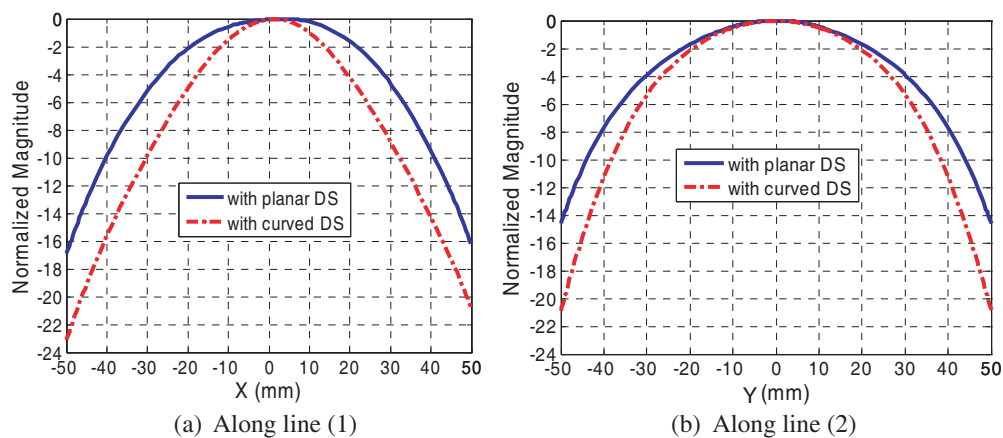
λ_0: the lambda of free space at 6 GHz

(a) Along line (1)   (b) Along line (2)

Figure 8: E-field magnitue distribution along line (1) and line (2).

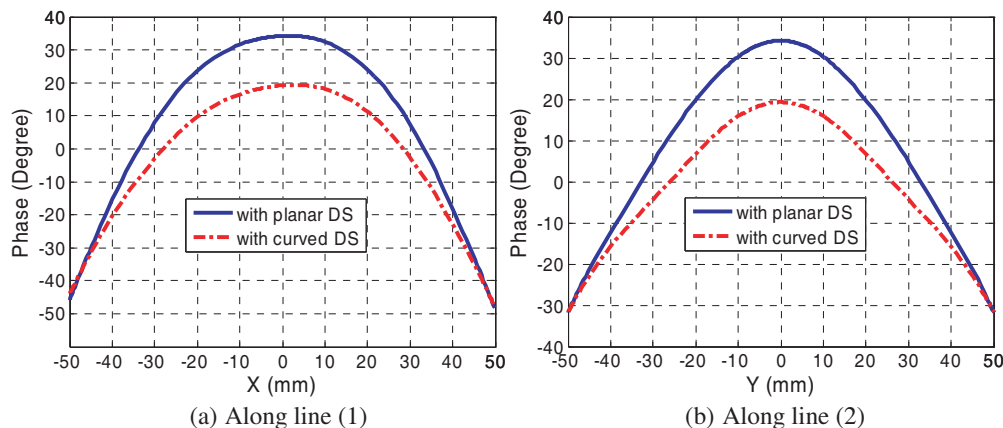(a) Along line (1)   (b) Along line (2)

Figure 9: E-field phase distribution along line (1) and line (2).

a higher directivity using a curved superstrate, as may be seen from Fig. 11. This leads us to conclude that the phase characteristic of the aperture distribution has a more dominant effect on the directivity than the amplitude behavior in this case.
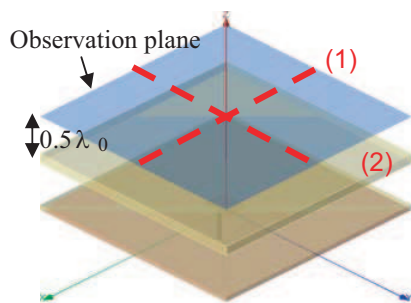

Figure 10: Position illustration of observation plane and structure. (1): line alone x in the center of observation plane. (2): line alone y in the center of observation plane.
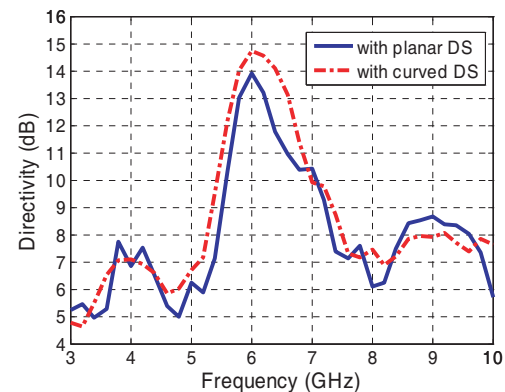

Figure 11: Directivity vary with frequency.

4. CONCLUSIONS

In this paper, microstrip patch antennas with planar and curved dielectric superstrates have been studied. The main impetus for studying this antenna structure with a dielectric superstrate was the desire to realize increased directivity without using complex structures such as DNG or EBG materials for the superstrate. The directivity characteristics of the composite structure have been investigated by using GEMS, which has been found to be a useful tool for designing antennas of this type. Finally, it was found that both the directivity level as well as the bandwidth could be further enhanced by using superstrates with curved upper surfaces, rather than a planar one.

REFERENCES

1. Lee, Y. J., J. Yeo, R. Mittra, and W. S. Park, "Design of a high-directivity electromagnetic band gap (EBG) resonator using a frequency selective surface (FSS) superstrate," *Microw. Opt. Tech. Lett.*, Vol. 43, No. 6, 462–467, Dec. 2004.
2. Lee, Y. J., J. Yeo, R. Mittra, and W. S. Park, "Application of electromagnetic bandgap (EBG) superstrates with controllable defects for a class of patch antennas as spatial angular filter," *IEEE Trans. Antennas Propagat.*, Vol. 53, No. 1, 224–235, Jan. 2005.
3. Shen, X.-H. and G. A. E. Vandenbosch, "Aperture field analysis of gain enhancement method for microstrip antennas," *Proc. 10th IEE Int. Conf. Antennas Propagat.*, Vol. 1, 186–189, Apr. 14–17, 1997.
4. Sauleau, R., "Fabry Perot resonators," *Encyclopedia of RF and Microwave Engineering*, Vol. 2, 1381–1401, Ed. K. Chang, John Wiley & Sons, Inc., May 2005.
5. Kishk, A. A., "One dimensional electromagnetic bandgap for directivity enhancement of waveguide antennas," *Microw. Opt. Tech. Lett.*, Vol. 47, No. 5, 430–434, Dec. 2005.
6. Clarke, R. N. and C. B. Rosenberg, "Fabry-Perot and open resonators at microwave and millimeter wave frequencies, 2–300 GHz," *J. Phys. E: Sci. Instrum.*, Vol. 15, 9–24, Jan. 1982.
7. Young, J. L., Y. Junho, R. Mittra, and S. P. Wee, "Design of a frequency selective surface (FSS) type superstrate for dual-band directivity enhancement of microstrip patch antennas," *IEEE Antennas and Propagation Society International Symposium*, Vol. 3A, 2–5, Washington, USA, Jul. 3–8, 2005.
8. GEMS, 3-D High Performance parallel EM simulation Software, www.2comu.com.

The Calculation of Back Scattering Field of Unmanned Air Vehicle

N. Altın[1] and E. Yazgan[2]

[1]Turkish Aerospace Industries, Inc., Ankara, Turkey
[2]Electrical & Electronics Engineering Department, Hacettepe University, Turkey

Abstract— In this paper, GTD, GO and PO is employed in calculating the back scattering field of Unmanned Air Vehicle (UAV) in the 1 GHz. The calculated result is compared by FEKO analysis program result. When target's electric size is two times greater than its wavelength, higher order diffraction field is ignored. The first order diffracted field will be sufficient to calculate the target's RCS.

1. INTRODUCTION

The problem of calculating the RCS of a complex target can be reduced, in many cases, to a set of problems involving calculation of the RCS of relatively simple shapes, with which various parts of the target are approximated, and then results are vectorially summed for the solution [3].

Asymptotic high frequency techniques (GTD/UTD, PO/PTD) remain essential for the solution of scattering problems involving large objects of arbitrary shapes like airplanes, helicopters, missiles, tanks and ships, at radar frequencies. But at present their implementation on a computer is rapidly changing due to the used modellings. Especially, the shapes to which high frequency techniques are applied are of growing complexity and the need for general computer codes which manage automatically the geometrical modeling and ray searching is more and more pronounced [1].

Analytic determination of scattering from complex targets such as an aircraft is extremely difficult, if not impossible. Some computer codes capable of providing numerical solutions are available [5–8]. On the other hand depending on the directions of incidence and observation, high frequency radar returns from a complex target may be thought of as composed of contributions from a number of scattering centers [2]. A complex target can be assumed to consist of a number of simple shapes whose scattering properties are known and based on this assumption a number of models have been developed to determine the RCS of complex targets.

In this paper, GTD, GO and PO are employed in calculating the back scattering field of unmanned air vehicle (UAV) for 1 GHz. Afterwards the calculated result is compared by FEKO analysis program result.

2. ANALYSIS AND CALCULATION

Figure 1 shows the outline geometry of UAV model for the modeling purpose. In order to decompose into basic geometrical shapes. This geometry can be assumed to consist of structures like in Fig. 2. It is made up of metal material. Incidence and receiving are all plane waves of horizontal and vertical polarizations.

The following steps describe the procedure to determine the RCS of complex structures such as satellites, aircraft and missile [2].

1. The complex shape is divided into a number of segment simple shapes. Geometrically simple shapes are now chosen in order to approximate the component shapes.
2. The RCS is calculated for each of simple shape.
3. RCS's of individual segments are combined to obtain the RCS of the complete shape.

2.1. Analysis

The GTD backscattering RCS of the UAV is determined by using analytical expressions. As shown in Fig. 1, body of UAV consists of two cone frustum and cylinder and wings and tails of UAV consist of plane.

The expressions for RCS at angles $\theta = 0$, $(\pi/2 - \alpha_1)$, $\pi/2$, $(\pi/2 - \alpha_2)$, π, $0 < \theta < \theta_1$, $\theta_2 < \theta < \pi$ are obtained by the PO expressions. For angles other than these, RCS is obtained by the GTD/UTD and superposition of contributions from individual scattering centers determined by the range of aspect angles of interest. The contribution to the scattered field by each scattering center is obtained by multiplying the incident field at position of the scattering center by the diffraction coefficient and the associated phase contributions.
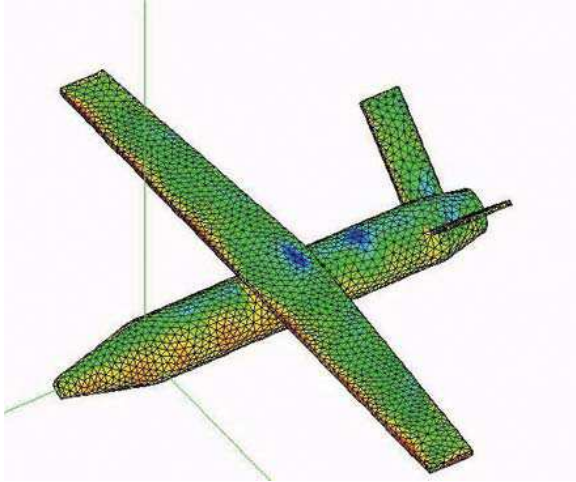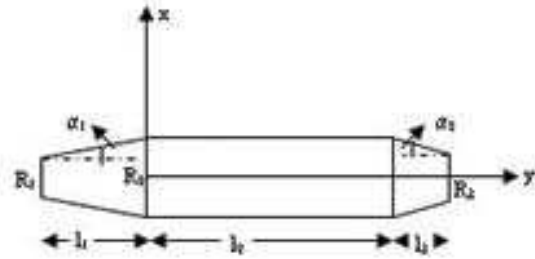
Figure 1: Geometry of UAV model.

Figure 2: Modeling of UAV body with basic geometrical shapes.

Commonly diffraction field can be described as

$$\vec{E}^d(\vec{r}) = \vec{E}^i(Q) \cdot \overline{\overline{D}}(\phi, \phi') \cdot A(r) \cdot e^{-jk \cdot \vec{r}} \tag{1}$$

where $\vec{r}$ is radius vector from diffraction point to field point, $A(\vec{r})$ is attenuation coefficient and $\overline{\overline{D}}$ is dyadic diffraction coefficient.

$$A(\vec{r}) \approx \sqrt{\frac{\rho_c}{r \cdot (\rho_c + r)}} \tag{2}$$

$$\frac{1}{\rho_c} = \frac{1}{\rho_e} - \frac{\hat{n}_e \cdot (\hat{s}' - \hat{s})}{a \cdot \sin^2 \beta_0} \tag{3}$$

All formulas above are built in ray-based coordinates. Dyadic diffraction coefficient can be described in ray based coordinates as

$$\overline{\overline{D}}(\phi, \phi'; \beta_0) = -\hat{\beta}'_0 \hat{\beta}_0 D_s - \hat{\phi}' \hat{\phi} D_h \tag{4}$$

where D_s and D_h are scalar diffraction coefficient of soft and hard boundary condition. The diffraction coefficients are given by

$$D_{s,h}(\varphi, n) = \frac{e^{i\pi/4} \sin(\pi/n)}{n\sqrt{2\pi k_0}} \left[\left(\cos\frac{\pi}{n} - 1\right)^{-1} \mp \left(\cos\frac{\pi}{n} - \frac{2\varphi}{n}\right)^{-1} \right] \tag{5}$$

where n is the $3/2 - \alpha_{1,2}/\pi$, the phase of diffracted ray, referred to a reference plane φ is given by $\mp 2k_0 r \sin\theta + 2k_0 l \cos\theta$.

2.2. Synthetically Processing Data

For n scattering sources, RCS produced by the combination of them is

$$\sigma = \left| \sum_{m=1}^{n} \sqrt{\sigma_m} e^{j2kR_m} \right|^2 \tag{6}$$

where σ_m is the RCS of every scattering source; R_m is the distance from the m scattering center to the radar. Considering the relations among phases, the method of relative phase is used to plus

every scattering field.

$$\theta = 0$$
$$\sigma(0) = \pi k_0^2 r_1^4 \tag{7}$$

$$0 < \theta < \theta_1$$
$$\sigma(\theta) = \pi r_1^2 \left[\frac{J_1(2k_0 r_1 \sin\theta)}{\tan\theta} \right]^2 \tag{8}$$

$$\theta_1 < \theta < \alpha_2$$
$$\sigma_{s,h} = 4\pi \left| D_{s,h} \left(\frac{\pi}{2} + \theta, n_1 \right) \rho(r_1)^{1/2} \exp(-i\psi_1) \right.$$
$$\left. + D_{s,h} \left(\frac{\pi}{2} - \theta, n_1 \right) \rho(-r_1)^{1/2} \exp(i\psi_1) + D_{s,h}(\theta, n_2)\rho(r_3)^{1/2} \exp(i\psi_2) \right|^2 \tag{9}$$

RCS value in the other angle values is calculated alike.

2.3. Numerical Result

Basic dimensions of the sample Unmanned Air Vehicle (UAV) are given below. UAV body length is 8 m, UAV wingspan is 16 m and wing thickness is 11 cm. As shown in Fig. 2, back vertical tail is located with 45°. Back tail width is 0.7 m, length is 2.8 m and thickness is 8 cm. Operating frequency is 1 GHz. RCS value is examined for hard and soft polarization. As shown in Fig. 1, the analysis of UAV's RCS using FEKO progamme takes approximately 4 hours.
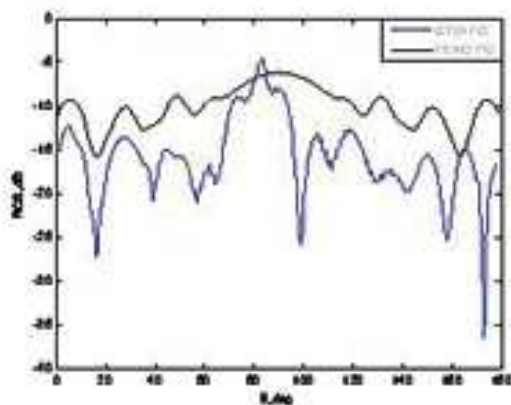

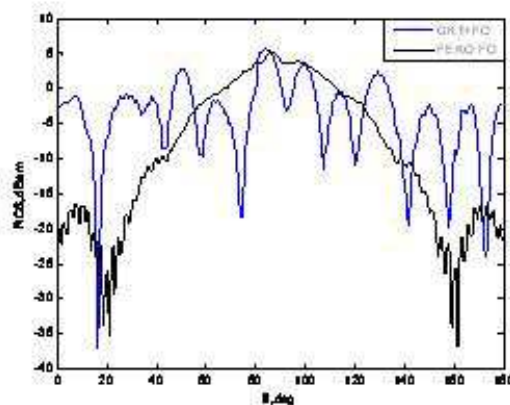Figure 3: RCS value of UAV for soft polarization.

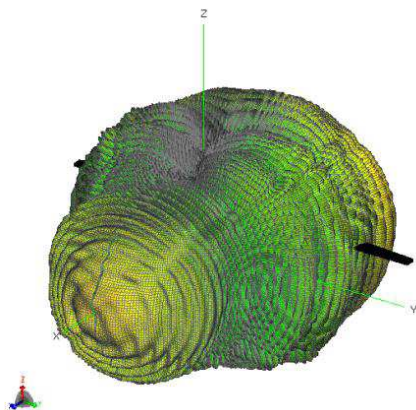Figure 4: RCS value of UAV for hard polarization.

Figure 5: 3D sketch of RCS value UAV for hard polarization with FEKO program.
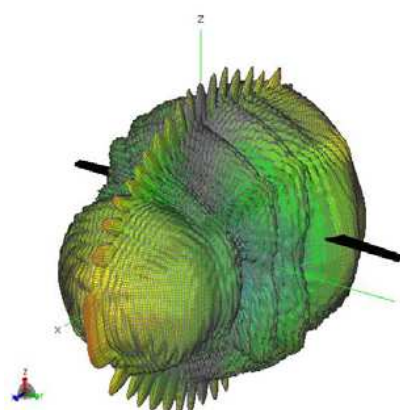

Figure 6: 3D sketch of RCS value UAV for soft polarization with FEKO program.

Back scattering field of unmanned air vehicle (UAV) is obtained with GTD and PO. The calculated result is compared by FEKO analysis program result.

PO method is used in FEKO program. Because of this, the shadow and diffraction effects have not been taken into account. This is the reason of the difference between the calculated result and the FEKO result. The result of calculation and result of FEKO are showed in Fig. 3 and Fig. 4. Since the base of FEKO program depends on moment method, its performance is not satisfactory for real platform analysis [4].

Figure 5 and Fig. 6 show 3D graph of UAV RCS value obtained with FEKO.

3. CONCLUSIONS

In this paper, the calculated result is compared with the results of FEKO for 1 GHz due to hard and soft polarization. But according to the results of calculation, when the electric dimension of the objects are greater than double wavelength, the influence of the second diffraction field and higher can be ignored. It is proved that the high-frequency calculation method can be applied with reasonable effort and the results efficient in the performance for radar cross section calculations of air vehicles.

REFERENCES
1. Crispin, J. W. and K. M. Siegel, *Methods of Radar Cross Section Analysis*, Academic Press, New York and London, 1968.
2. Bhattacharyya, A. K. and D. L. Sengupta, *Radar Cross Section Analysis and Control*, Artech House, Boston, London, 1991.
3. Youssef, N. N., "Radar cross section of complex," *Proceedings of the IEEE*, Vol. 77, No. 5, May 1989.
4. The FEKO website. [Online]. Available: http://www.feko.info/.
5. Burke, G. J. and A. J. Poggio, "Numerical electromagnetic code (NEC)-method of moments," *Naval Ocean Syst. Cont. Tech. Doc.*, 116, AFWL-TR-76-320, Jul. 1977.
6. Newman, E. H., "A users manual for electromagnetic surface patch code (ESP)," *OSUESL*, Dept. Elect. Eng. Rep., 713602-1, prepared under Contract DAAG 39-81-K-0020 for Dept. Army, U. S. Army Research Office, Research Triangle Park, NC, Jul. 1981.
7. Newman, E. H. and P. Alexandropoulous, "A users manual for electromagnetic surface patch code: Version II-polygonal plates and wires," *OSUESL*, Dept. Elect. Eng. Rep., 713602-1, prepared under Contract N0016-18-C-0049 for the Dept. of Navy, Office of Naval Research, Arlington, VA, Sept. 1983.
8. Crispin, Jr., J. W. and A. L. Moffatt, "Radar cross section estimating for complex targets," *Proc. IEEE*, Vol. 53, Aug. 1965.

Estimation Error of Topographic Phase Based on RVoG Model Using POLinSAR Data

L. Bai[1, 2, 3]**, W. Hong**[1, 2]**, and F. Cao**[1, 2]

[1]National Key Laboratory of Microwave Imaging Technology, China
[2]Institute of Electronics, Chinese Academy of Sciences, China
[3]Graduate University of Chinese Academy of Sciences, China

Abstract— Topographic phase is a key parameter in the Random Volume over Ground forest model. Its estimation error influences the estimation accuracy of other forest parameters. Therefore, the error analysis and error control in topographic is necessary to obtain the accurate forest parameters. This paper analyzes the main error sources in topographic phase and provides a method to estimate the possible error range. To examine the proposed method, the error analysis is applied on the results of ESAR L band forest data.

1. INTRODUCTION

The estimation of biophysical parameters of forest from polarmetric and interferometric SAR data has been demonstrated recently [1]. The inversion is based on Random Volume over Ground (RVoG) model, which relates forest biophysical parameters with the observables. This model has been widely applied to the deciduous forest, conifer forest and rain forest [2]. Based the model, many techniques are proposed to inverse forest height and underlying ground height from the POLinSAR coherences [3].

Up to now, forest height error due to bad estimation of the topographic phase has been analyzed [4]. However, the estimation error of topographic phase has not been discussed. Since its errors would be passed to the results of other forest parameters, it is necessary to analyze and control the error.

This paper analyzes and estimates the estimation error according to the geometrical characteristics of coherence distribution. In the RVoG model, the topographic phase can be estimated from the intersections of the straight line, fitted by certain group of coherence samples, and the unit circle. According to the realistic coherence distribution, the factors leading to the estimation errors in topographic phase are analyzed. Using the geometrical characteristics of coherence distribution, this paper presents a method to estimate the possible estimation error. Experiments are applied to examine the effectiveness of the proposed error analysis.

This paper describes the RVoG model and its topographic phase inversion in Section 2. Three stages to analyze errors in topographic phase are presented in Section 3. The method to estimate possible error range in each stage is proposed, too. To examine the error analysis method, experiments are applied to the Traunstein forest data acquired by the L-band ESAR system. As the results in Section 4 shown, the proposed analysis method is effective. Finally, the applicability of the proposed method is drawn in Section 5.

2. RVOG MODEL AND TOPOGRAPHIC PHASE

RVoG model is a layered physical model for the forest. It assumes the forest as a homogeneous layer of randomly oriented scatterers covering on the ground. As shown in Figure 1(a), the thickness of the layer or the forest height is h_v and the scattering amplitude per unit volume scatterers is m_v. The ground locates on the altitude z_0 and its scattering amplitude is denoted as m_G. When SNR decorrelation and temporal decorrelation are negligible, the interferometric coherence in the forest can be expressed as:

$$\tilde{\gamma}(\vec{\omega}) \equiv \frac{\vec{\omega}^H \Omega_{12} \vec{\omega}}{\sqrt{(\vec{\omega}^H \mathbf{T}_{11} \vec{\omega})(\vec{\omega}^H \mathbf{T}_{22} \vec{\omega})}} = e^{j\phi_0} \frac{\tilde{\gamma}_V + m(\vec{\omega})}{1 + m(\vec{\omega})} \tag{1}$$

where superscript H represents the conjugate and transposition operation. Vector $\vec{\omega}$ is related to the polarization states. Matrix $\mathbf{T}_{11}$, $\mathbf{T}_{22}$ and Ω_{12} are the outer products of the POLinSAR data:

$$\Omega_{12} = \left\langle \vec{k}_1^H \vec{k}_2 \right\rangle, \quad \mathbf{T}_{11} = \left\langle \vec{k}_1^H \vec{k}_1 \right\rangle, \quad \mathbf{T}_{22} = \left\langle \vec{k}_2^H \vec{k}_2 \right\rangle$$

where

$$\vec{k}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} S_{1HH} + S_{1VV} \\ S_{1HH} - S_{1VV} \\ 2S_{1HV} \end{bmatrix} \quad \text{and} \quad \vec{k}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} S_{2HH} + S_{2VV} \\ S_{2HH} - S_{2VV} \\ 2S_{2HV} \end{bmatrix} \tag{2}$$

In (1), Symbol ϕ_0 is the interferometric phase indicating the height of the ground. It is called topographic phase in this paper. $\tilde{\gamma}_V$ represents the complex coherence for the layer. It is closely related to the forest height and forest extinction coefficient σ. m denotes the effective ground-to-volume amplitude ratio and it depends on the polarization state $\vec{\omega}$. Therefore, underlying topography can be estimated from the RVoG model [3].

The coherence in (1) can be rewritten as:

$$\tilde{\gamma}(\vec{\omega}) = e^{j\phi_0} \left[L(\vec{\omega})(1 - \tilde{\gamma}_V) + \tilde{\gamma}_V \right] \quad \text{where} \quad L(\vec{\omega}) = \frac{m(\vec{\omega})}{1 + m(\vec{\omega})} \tag{3}$$

$L(\vec{\omega})$ is the ratio and falls in $[0, 1)$. Figure 1(b) shows the distribution of complex coherences. They form a solid blue segment of the red dash straight line. Moreover, the complex topographic phase $e^{j\phi_0}$ is one intersection of the red line with the unit circle. Thus, the straight line and topographic phase can be estimated from any set of coherence samples, such as complex coherence in HH, HV or VV polarization state. In realistic, however, the coherence distribution in forest does not accord to a straight line strictly. To obtain the accurate topographic phase, many inverse techniques proposed to choose the reliable coherences sample. The coherences given by coherence optimization are widely used because they have the possible highest correlation [3].
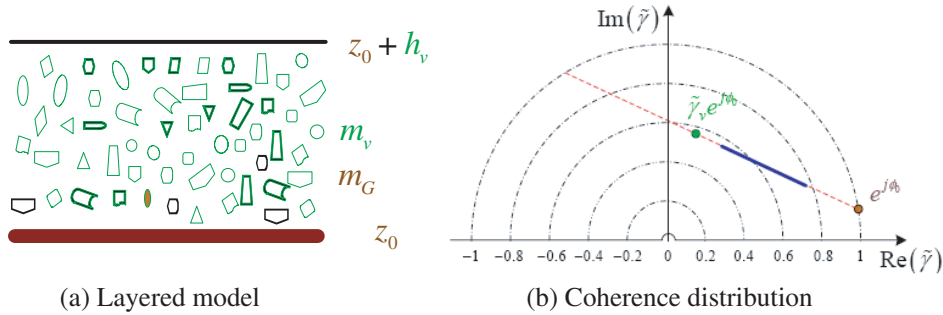

(a) Layered model

(b) Coherence distribution

Figure 1: Schematic description for RVoG model.

3. ERROR ANALYSIS OF TOPOGRAPHIC PHASE

In real forest, the canopy scatterers are partly oriented and the volume scattering varies with the polarization states. The forest observations do not satisfy the assumption of RVoG model. The coherences in forest do not distribute along the straight line but disperse in a region. The shape of coherence region may be an ellipse or a triangle. The choices of coherence samples influence the estimation for the straight lines and topographic phase. Therefore, the error in topographic phase is analyzed based on geometrical parameter of the coherence region.

This paper analyzes topographic phase estimation $\hat{\phi}_0$ given by the optimum coherences. The whole analysis can be divided into three stages. Stage 1 analyzes the error when the coherence distribution has preferred orientation. Stage 2 gives the fine error analysis when the coherence distribution satisfies RVoG model. Stage 3 analyzes the error caused by SNR decorrealtion when the estimate suffers from much less volume decorrelation.

3.1. Error Analysis in Stage 1

The coherence region is defined as the area within which all possible complex coherences disperse [5]. To describe its shape, a parameter named narrowness ratio is introduced. It is defined as the distance ratio:

$$Na = \frac{\max\left\{ |\tilde{\gamma}_3 - \tilde{\gamma}_4|, (\tilde{\gamma}_1 - \tilde{\gamma}_2)(\tilde{\gamma}_3 - \tilde{\gamma}_4)^* = -(\tilde{\gamma}_1 - \tilde{\gamma}_2)^*(\tilde{\gamma}_3 - \tilde{\gamma}_4), \tilde{\gamma}_3 \subset R, \tilde{\gamma}_4 \subset R \right\}}{\max\left\{ |\tilde{\gamma}_1 - \tilde{\gamma}_2|, \tilde{\gamma}_1 \subset R, \tilde{\gamma}_2 \subset R \right\}} \tag{4}$$

where symbol R consists of the coherence on the border of coherence region. The red ring in Figure 2 forms the R and they are estimated as [5] did. The denominator in (6) is the length of longest

axis in the coherence region. It is shown as the long solid line in Figure 2. The numerator is the maximum length of the axis orthogonal to the longest axis. The narrowness ratio falls in the range $[0, 1)$. The smaller the narrowness ratio is, the closer RVoG model is to the forest observations. The less possible estimation error occurs in topographic phase. Hence, $1 - Na$ can be regarded as the error risk indicator.

When the narrowness ratio is lower than 0.5, coherence region has obvious preferable direction and the estimation error can be analyzed according to RVoG model. According to the model, the line segment related to the topographic phase gives the possible largest distance. Hence, the longest axis gives the expected topographic phase $\hat{\phi}_0$ shown as brown circle in Figure 2. Considering the reliability of RVoG model, the possible error in $\hat{\phi}_0$ can be estimated from:

$$\phi_{e1} = \frac{1}{1 - Na} \left| \arg \left(e^{j\hat{\phi}_0} e^{-j\bar{\phi}_0} \right) \right| \tag{5}$$

Symbol arg() is to calculate the argument. When the RVoG model does not fit to the data, the estimation error $\hat{\phi}_0 - \bar{\phi}_0$ is amplified to large possible error. In this situation, no further error analysis can be done according to the RVoG model. When RVoG model is reliable, the possible error estimated in Stage 1 is very close to $\hat{\phi}_0 - \bar{\phi}_0$ and detailed analysis in Stage 2 can be applied.

3.2. Error Analysis in Stage 2

Besides the shape of the coherence region, the RVoG model can be examined according to the interferometric phase. Deduced from (1), the interferometric phase changes monotonically. The preferred segment of the coherence region gives the maximum phase difference. The coherence samples with the maximum phase gradient can be used to estimate the expected topographic phase $\underline{\phi}_0$. It is shown as the blue circle in Figure 3.

Since $\underline{\phi}_0$ and $\bar{\phi}_0$ are estimations given by the RVoG model, their differences serves as a goodness indicator for the model and the expected ones. The small difference between $\underline{\phi}_0$ and $\bar{\phi}_0$ indicates the expected topographic phase is reliable. Stage 2 analyzes the possible error when the phase difference between $\underline{\phi}_0$ and $\bar{\phi}_0$ is less than 0.3 rad. In this situation, the deviation of estimated topographic phase from the expected ones can be taken as the estimation error:

$$\phi_{e2} = \max \left\{ \left| \arg \left(e^{j\hat{\phi}_0} e^{-j\bar{\phi}_0} \right) \right|, \left| \arg \left(e^{j\hat{\phi}_0} e^{-j\underline{\phi}_0} \right) \right| \right\} \tag{6}$$

When the phase difference is over the threshold 0.3 rad, it is hard to identify whether $\underline{\phi}_0$ or $\bar{\phi}_0$ is the reliable. Maybe both $\underline{\phi}_0$ and $\bar{\phi}_0$ suffer from the estimation errors. That is why the fine error can not be estimated.

3.3. Error Analysis in Stage 3

In RVoG model, the errors caused by the SNR decorrelation are neglected. However, the SNR decorrelation would dominate the phase error when the volume decorrelation effects are decreased. It is to say the phase difference between $\underline{\phi}_0$ and $\bar{\phi}_0$ need to be less than 0.1 rad. The Stage 3 is to analyze the estimation error caused by the SNR decorrelation.

Taking the SNR decorrelation γ_{SNR} into account, the forest coherence can be expressed as:

$$\tilde{\gamma}(\vec{\omega}) = \gamma_{\mathrm{SNR}} \cdot e^{j\phi_0} \left[L(\vec{\omega})(1 - \tilde{\gamma}_V) + \tilde{\gamma}_V \right] \tag{7}$$

Figure 4 shows the effects of γ_{SNR} on the coherence distribution. The observed segment marked by the brown solid line shrinks from the real green solid line. Therefore, the expect topographic pahse $\bar{\phi}_0$ is deviated from the real topographic phase $\bar{\bar{\phi}}_0$. To get the phase information ϕ_0, the SNR decorrelation γ_{SNR} should be estimated. When γ_{SNR} is known, the observed line segment can be shifted to the real segment. The intersection of real segment and the unit circle is the real topographic phase. Deduced from the coherence sample with the optimum coherence:

$$|\tilde{\gamma}_{1opt}| = |\gamma_{\mathrm{SNR}}| \cdot |\tilde{\gamma}(\vec{\omega}_{1opt})| \overset{|\tilde{\gamma}(\vec{\omega}_{1opt})| \leq 1}{\Longrightarrow} |\gamma_{\mathrm{SNR}}| \geq |\tilde{\gamma}_{1opt}| \tag{8}$$

To estimate the maximum possible estimation error, the largest effect caused by $\gamma_{\mathrm{SNR}} = |\tilde{\gamma}_{1opt}|$ is analyzed. The endpoints $\tilde{\gamma}_1$ and $\tilde{\gamma}_2$ of the longest axis in the observed coherence region are

shifted to $\tilde{\gamma}_1/\gamma_{\mathrm{SNR}}$ and $\tilde{\gamma}_2/\gamma_{\mathrm{SNR}}$ respectively. The possible topographic phase $\bar{\bar{\phi}}_0$, farthest from $\hat{\phi}_0$, is obtained. The possible estimation error, accounting for γ_{SNR}, can be calculated from:

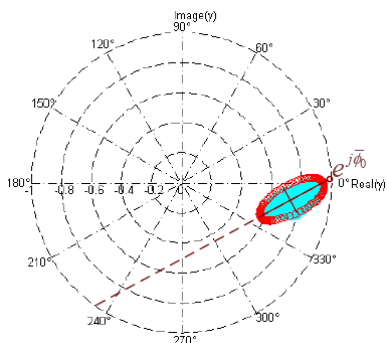$$\phi_{e3} = \left| \arg\left(e^{j\hat{\phi}_0} e^{-j\bar{\bar{\phi}}_0} \right) \right| \tag{9}$$
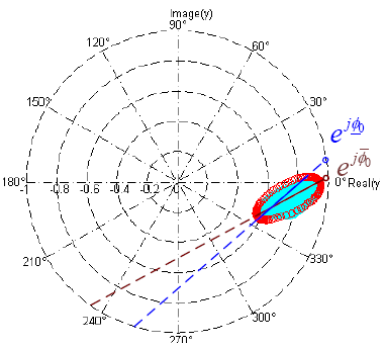

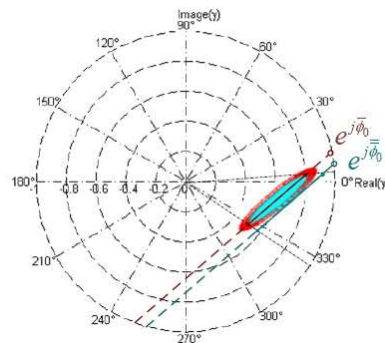Figure 2: Analysis in Stage 1.   Figure 3: Analysis in Stage 2.   Figure 4: Analysis in Stage 3.

4. EXPERIMENTS AND RESULTS

The proposed method is applied on the experimental results to demonstrate its effectiveness. The data were acquired by L band ESAR system in 2003 and they covered Traunstein Forest sites in Germany. The topographic phase estimations $\hat{\phi}_0$ given by the three optimum coherences are analyzed. Using the X band InSAR DEM results of the bare ground near the forest (10-pixel distance), the reference topographic phase ϕ_0 in the forest can be calculated. Compared with the deviation $|\arg(e^{j\hat{\phi}_0} e^{-j\phi_0})|$, the estimation error given by the three stages is assessed:

1. Calculate R and Na. If $Na < 0.5$, the coarse error can be estimated from (5) and go to Stage 2; or else, begin the analysis for the next result.
2. Calculate phase difference between $\underline{\phi}_0$ and $\bar{\phi}_0$. If the phase difference is less than 0.3 rad, the error can be replaced with the error defined in (6) and step in Stage 3; or else, go back to Stage 1 to analyze next result.
3. If the phase difference is less than 0.1 rad, the error can be replaced with the result of (9); or else, go back to Stage 1 and analyze the next result.

Figures 5–7 compare the error estimations and the measured error in three stages respectively. From the distribution of the red samples, most error estimation in Stages 1 and 2 are larger than the measured ones. They can provide the coarse range of the estimation error. The estimated error in stage three gives more fine estimations because more percentages of estimated errors are close to the measured ones. It is potent to analyze the accuracy of topography estimations.
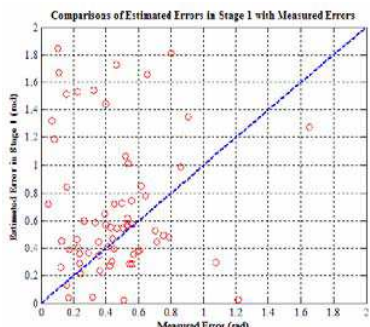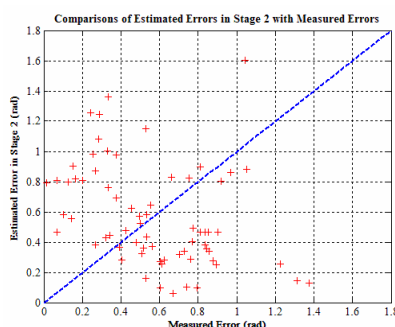

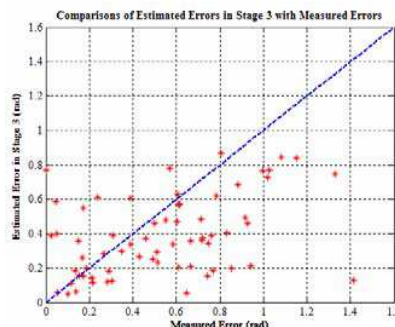Figure 5: Analysis in Stage 1.   Figure 6: Analysis in Stage 2.   Figure 7: Analysis in Stage 3.

5. CONCLUSIONS

This paper analyzes the error in topographic phase estimated from RVoG model. According the goodness of RVoG model, the possible error can be estimated in three different stages. The latter stage can gives finer error estimation than that given in the front stage, but applicability of the latter one is not as good as the front one.

ACKNOWLEDGMENT

REFERENCES

1. Treuhaft, R. N. and P. R. Siquira, "Vertical structure of vegetated land surfaces from interferometric and polarimetric radar," *Radio Science*, Vol. 35, No. 1, 141–177, 2000.
2. Cloude, S. R., K. P. Papathanassiou, et al., "The Glen Affric project: Forest mapping using polarimetric radar interferometry," *Proceedings of IGARSS*, 1642–1644, Sydney, Australia, July 2001.
3. Cloude, S. R. and K. P. Papathanassiou, "Three-stage inversion process for polarimetric SAR interferometry," *IEE Proceedings: Radar, Sonar and Navigation*, Vol. 150, No. 1, 125–134, 2003.
4. Tobias, M., K. Florian, et al., "Forest and the random volume over ground-nature and effect of 3 possible error types," *Proceedings of EUSAR Conference*, Dresden, Germany, May 2006.
5. Flynn, T., M. Tabb, and R. Carande, "Coherence region shape extraction for vegetation parameter estimation in polarimetric SAR interferometry," *Proceedings of IGARSS*, 2596–2598, Toronto, Canada, July 2002.

Ferrite Image Lines Studies by Transverse Operator Method

H. Sakli[1], H. Benzina[2], T. Aguili[1], and J. W. Tao[3]

[1]SYS'COM Laboratory, Ecole Nationale d'Ingénieurs de Tunis, Le Belvédère, Tunis 1002, Tunisia
[2]Ecole Nationale d'Ingénieurs de Gabès, Rue Omar Ibn El Khattab, Gabès 6029, Tunisia
[3]Groupe de Recherche en Electromagnétisme (GRE), ENSEEIHT-INPT, Toulouse 31071, France

Abstract— In this paper, a rigorous study of the transverse operator method (TOM) formulation is presented by an inhomogeneous rectangular structure of ferrite with transverse anisotropy which will be followed by the application of the Galerkin method.

Our study is essentially focused on the determination of the propagation constant in a ferrite image line magnetized longitudinally by a static magnetic field. The transverse operator method has been applied to a generalized multidielectric waveguide structure. Our results were obtained for the dielectric image line and compared with others approaches and good agreement with well kow analysis was obtained.

1. INTRODUCTION

The formulation based on the study of the transverse components of the electric and magnetic fields has been applied to dielectric waveguide problem by K. Ogusu [1]. The TOM was developed and applied to a number of multidielectric waveguides [2, 3].

In this paper, the formulation of TOM will be presented for the ferrite image line. Our results are in good agreement with those previously published [4]. The transverse operator which is another formulation of the Maxwell equations is here used to define the rectangular waveguide characteristics charged of ferrites. This method consists on eliminating the longitudinal components and resolving the propagation equation by developing the transverse fields in series of modes of a closed structure (empty metal guide).

The elimination of the longitudinal fields in the Maxwell equations lets appear an operator L named the transverse operator. The resolution of the propagation equation by the Galerkin method leads to an eigenvalues equations.

2. THEORY

At the microwaves frequencies, the ferrites are characterized by a tensorial permeability which represents their induced anisotropy under a magnetic field. The permeability of the ferrite magnetised longitudinally is described by the Polder tensor:

$$\overline{\overline{\mu}} = \mu_0 \cdot \begin{bmatrix} \overline{\overline{\mu}}_{tt} & \mu_{tz} \\ \mu_{zt} & \mu_{zz} \end{bmatrix} = \mu_0 \cdot \begin{bmatrix} \mu & -j\kappa & O \\ j\kappa & \mu & O \\ 0 & 0 & \mu_{rz} \end{bmatrix} = \mu_0 \cdot \overline{\overline{\mu}}_{rf}, \tag{1}$$

where μ, κ and μ_{rz} are real quantities. For a partial magnetization of the ferrite, Green [5] and Schlomann [6] give the empirical expressions of μ, κ and μ_{rz}. When the magnetization is equal to zero, $\kappa = 0$ and $\mu = \mu_{rz} = 1$. The ferrite becomes then an isotropic dielectric.

In this paper, we will study, using the transverse operator method, the propagation in the ferrite image guide represented by the Fig. 1. An eigenvalues equation can be obtained.

In the ferrite medium, the Maxwell equations are expressed as follows

$$rot \, \vec{E} = -j\omega\mu_0\overline{\overline{\mu}}_r \cdot \vec{H} \tag{2}$$

$$rot \, \vec{H} = j\omega\varepsilon_0\varepsilon_r \cdot \vec{E} \tag{3}$$

By considering the propagation in the Oz direction and by eliminating the longitudinal components of the electromagnetic field, we obtain the following expression [2, 3]

$$\hat{L}_t\Phi(x, y) = j\eta\partial_z\Phi(x, y) \tag{4}$$

$\hat{L}_t$ is the transverse operator defined by

$$\hat{L}_t = \begin{bmatrix} k_0 \varepsilon_r - 1/k_0 \partial_t \left[1/\mu_{rz} \partial_t^+ \right] & 0 \\ 0 & k_0 \overline{\overline{\mu}}_{tt} - 1/k_0 \partial_t \left[1/\varepsilon_r \partial_t^+ \right] \end{bmatrix} \tag{5}$$

With

$$\partial_t^+ = \begin{bmatrix} -\partial_y & \partial_x \end{bmatrix}; \qquad \partial_t = \begin{bmatrix} \partial_y \\ -\partial_x \end{bmatrix}; \qquad \eta = \begin{bmatrix} 0 & \eta_0 \\ \eta_0 & 0 \end{bmatrix}; \qquad \eta_0 = \begin{bmatrix} 0 & -j \\ j & 0 \end{bmatrix};$$

$$k_0 = \omega \sqrt{\varepsilon_0 \mu_0}; \qquad \Phi = \begin{bmatrix} E_t & H_t' \end{bmatrix}^t; \qquad H' = j z_0 \cdot H \tag{6}$$

Φ denotes the transverse field vector, k_0 and $z_0 = \sqrt{\mu_0/\varepsilon_0}$ are respectively the propagation constant and the characteristic impedance of free-space. ∂_t is the transverse curl operator with ∂_t^+ its adjoint operator.

By using the transverse anisotropy, we obtain a new formulation with two transversal magnetic components or two transversal electric components. For the first case, we have [2]

$$\hat{L}' H_t'(x, y) = k_z^2 H_t'(x, y) \tag{7}$$

with:

$$\hat{L}' = k_0^2 \varepsilon_r \overline{\overline{\mu}}_{tt} - \eta_0 \partial_t \frac{1}{\varepsilon_r} \partial_t^+ \eta_0 \varepsilon_r - \eta_0 \overline{\overline{\mu}}_{tt} \eta_0 \partial_t \frac{1}{\mu_{rz}} \partial_t^+ \tag{8}$$

We describe the discontinuity of the permittivity and the permeability in the guide by the Heaviside function U and the Dirac impulsion δ.
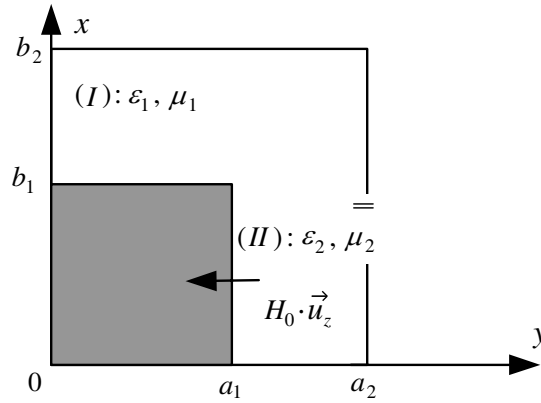

Figure 1: Cross-section of ferrite image guide.

The expression of ε_r and μ_{rz} in different mediums can be written as follows:

$$\varepsilon_r = 1 + (\varepsilon_{r2} - 1) \cdot U(X) \cdot U(Y) \tag{9}$$
$$\mu_{rz} = 1 + (\mu_{rz2} - 1) \cdot U(X) \cdot U(Y) \tag{10}$$

where:

$$U(X) = U(x - a_1) - U(x - a_2) \tag{11}$$
$$U(Y) = U(y) - U(y - b_1) \tag{12}$$
$$\delta(X) = \partial_x U(X) \tag{13}$$
$$\delta(Y) = \partial_y U(Y) \tag{14}$$
$$\frac{\partial_x \varepsilon_r}{\varepsilon_r} = 2 \frac{\varepsilon_{r2} - 1}{\varepsilon_{r2} + 1} \cdot \delta(X) \cdot U(Y) \tag{15}$$
$$\partial_x \left(\frac{1}{\mu_{rz2}} \right) = -2 \frac{\mu_{rz2} - 1}{(\mu_{rz2})^2 + 1} \cdot \delta(X) \cdot U(Y) \tag{16}$$

The decomposition of the field $\vec{H}'_t$ on a complete basis (trigonometric functions) permits to obtain a system with eigenvalues which can be written as follows:

$$G \cdot H'_t = k_z^2 \cdot H'_t \tag{17}$$

G is a $(2N \times 2N)$ matrix with $N = m \cdot n$: number of modes; m and n are natural integers, verifying: $(m, n) \neq (0, 0)$. The eigenvalues and the eigenvectors of H' are respectively the propagation constants k_z and the development coefficients of the guide field.

3. RESULTAT DE SIMULATION AND DISCUSSION

The structure of the Fig. 1 that we have studied is a rectangular guide formed by the air and a ferrite medium, named image guide. The parameters of this guide are given by the Table 1.

Table 1: Guide parameters of Fig. 1.

a_1/b_1	a_2/a_1	ε_{r1}	ε_{r2}
1	3	1	13.1

We obtain the following dispersion curve (Fig. 2). $B = \frac{(k_z/k_0)^2}{\varepsilon_{r2}}$ represents the normalized propagation constant and $V = 2a_1 k_0 \sqrt{\varepsilon_{r2}}$ is the normalized frequency.
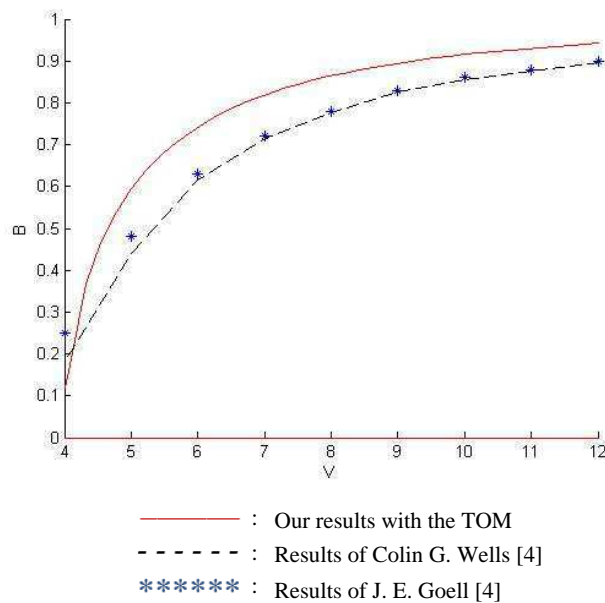

——————— : Our results with the TOM

- - - - - - : Results of Colin G. Wells [4]

****** : Results of J. E. Goell [4]

Figure 2: Normalized propagation constant of ferrite image line compared with Reference [4]; ($H_0 = 0\,\mathrm{KA/m}$).

In Fig. 2, our results are compared to those obtained by a mode-matching technique [4]. Good agreement can be observed for the normalized phase constant. The proposed method gives better numerical convergence than those approaches (The convergence stability of the propagation constant is obtained when reaching the value $N = 5$ modes) because of the appropriate choice of eigenfunctions. Other ferrite waveguides with arbitrary cross-sections can be studied using this procedure with no modifications.

The main advantage of the TOM is that the number of eigenmodes used in the analysis does not depend on the multidielectric structure because only one complete basis, that of the associated stratified dielectric waveguide, is used. This will not be the same with other methods such as mode matching techniques, finite difference and finite element methods.

4. CONCLUSION

In this paper, we have extended the transverse operator method to the case of ferrites. Our design of the permittivity and permeability discontinuities and our choice of the basic trigonometric functions

give better results of the propagation constant compared with other methods in the literature. Moreover, the convergence of the proposed method is very fast.

REFERENCES

1. Ogusu, K., "Numerical analysis of the rectangular dielectric waveguide and its modifications," *IEEE Trans.-MTT*, Vol. 25, 874–885, Nov. 1977.

2. Tao, J. W., J. Atechian, R. Ratovondrahanta, and H. Baudrand, "Transverse operator study of a large class of multidielectric waveguides," *Proc. of IEE*, Vol. 137, No. 5, Part H, 311–317, Oct. 1990.

3. Andriamanjato, R., B. Chan, J. W. Tao, and H. Baudrand, "Fullwave analysis of dielectric waveguide ferrite phase shifter with longitudinal magnetization," *Electronic Lett.*, Vol. 28, No. 20, 1907–1908, Sept. 1992.

4. Wells, C. G. and J. A. R. Ball, "Mode-matching analysis of a shielded rectangular dielectric-rod waveguide," *IEEE Trans.-MTT*, Vol. 53, No. 10, 3169–3177, Oct. 2005.

5. Green, J. J. and F. Sandy, "Microwave characterization of partially magnetized ferrites," *IEEE Trans.-MTT*, Vol. 22, 641–645, June 1974.

6. Schloemann, E., "Microwave behavior of partially magnetized ferrites," *J. Appl. Phys.*, Vol. 1, 1204, June 1970.

A Localization Scheme Using Bi-directional Metrics Joint Estimation

C. K. Seow, S. Y. Tan, and **S. W. Chen**

Nanyang Technological University, Singapore

Abstract— This paper presents a novel localization scheme that leverages on the joint estimation of the Angle of Arrival (AOA) and the Time of Arrival (TOA) at the mobile and reference devices. A two-step least squares (LS) location estimator that utilizes the TOA and AOA information at both the reference and mobile devices is developed. Numerical simulation results in a typical environment demonstrate that the proposed joint localization scheme which uses only one Line of Sight (LOS) reference device, provide much better location accuracy compared to the conventional unidirectional TOA and TOA/AOA localization schemes, which use disjoint metric estimation and three reference devices.

1. INTRODUCTION

Usually, the performance of the conventional localization schemes [1, 2] will degrade when the channel environment consists of a multipath-rich medium with numerous scattering objects (scatterers). One or more reference devices (RDs) may not be in the LOS range of the mobile device (MD), thus rendering the geometrical relationship derived for LOS localization erroneous. Various Non Line of Sight (NLOS) mitigation techniques have emerged [1, 3] to overcome this problem. However, these NLOS mitigation techniques will only perform satisfactorily if the number of LOS RDs is greater than the number of NLOS RDs, which may not be feasible in multipath environments. Furthermore, the measured NLOS AOA may be mistaken as LOS AOA for scheme [2] that obtains TOA and AOA measurement separately.

In this paper, we propose a novel localization scheme to overcome above limitations. The AOA and TOA are estimated jointly at both the RD and MD (referred to as bi-directional metrics estimation). This can be done as both ends are equipped with antenna arrays for carrying out AOA estimation [4, 5]. The advantages of using bi-directional estimation for localization are twofold. Firstly, we can devise a NLOS path rejection mechanism to infer whether each measured path is LOS or NLOS. The underlying mechanism is the LOS angle relationship that exists between each RD and MD pair. Secondly, bi-directional estimation allows an enhancement technique for readjusting and improving the accuracy of the AOA measurement values at both the RD and MD to be designed. Through these readjustments, the accuracy of identifying the MD's location is improved.

With bi-directional measurement information coupled with the NLOS path detection and AOA adjustment techniques, our proposed localization algorithm that uses only one LOS RD generally performs better than conventional unidirectional localization scheme with three RDs.

2. CONCEPT AND FORMULATION

As depicted in left plane of Fig. 1, RD_j with known coordinates (x_j, y_j) has a measured AOA α_j and TOA t_j, while MD with unknown coordinates (x, y) has a measured AOA β_j and TOA τ_j. It should be noted that MD does not know the position of RD_j. Therefore, RD'_j in Fig. 1 depicts the assumed but incorrect position of RD_j due to the values of τ_j and β_j measured at MD. Similarly, MD' denotes the assumed but incorrect position of MD due to the inaccuracy of the TOA and AOA (t_j and α_j) measurements at RD_j.

We infer that a measured path between RD_j and MD is an LOS path if the following devised geometrical relationship is satisfied

$$|\alpha_j - \beta_j| \leq 180° \pm \sigma_j \tag{1}$$

where $\sigma_j = \max\left(3\sigma_{\alpha_j}, 3\sigma_{\beta_j}\right)$. σ_{α_j}, σ_{β_j} are the standard deviations of the AOA measurement noise at the RD_j and MD respectively. Measured path with an angle difference greater than σ_j will be assumed as a NLOS path and rejected. Next, we devise the two step joint LS localization scheme. The measured angle and time parameters at RD_j and MD can be formulated as

$$l_j = ct_j = l_j^o + n_{l_j}, \quad r_j = c\tau_j = r_j^o + n_{r_j}, \quad \alpha_j = \alpha_j^o + n_{\alpha_j}, \quad \beta_j = \beta_j^o + n_{\beta_j} \tag{2}$$
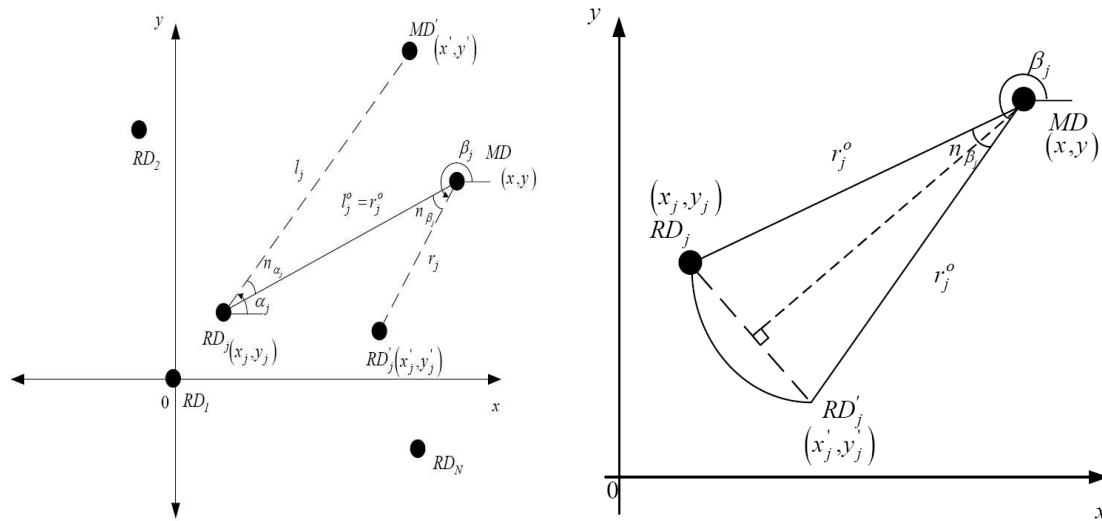
Figure 1: (Left) Localization in a 2D plane with one reference device. (Right) Geometrical relationship between the devices.

where $j = 1 \cdots N$ with N being the number of LOS RDs and c as the speed of propagation. l_j^o and $\alpha_j^o(r_j^o$ and $\beta_j^o)$ represent the noise free TOA (in terms of distance) and AOA at the $RD_j(MD)$ respectively. n_{l_j}, n_{r_j} and n_{α_j}, n_{β_j} are the TOA and AOA Gaussian measurement noises at RD_j and MD respectively. For TOA in (2), it can be expressed as devices' coordinates using the relationship $l_j^o = r_j^o = \sqrt{(x - x_j)^2 + (y - y_j)^2}$ and derived as

$$l_j^2 - K_j = -2x_j x - 2y_j y + R + 2l_j^o n_{l_j} + n_{l_j}^2, \quad r_j^2 - K_j = -2x_j x - 2y_j y + R + 2r_j^o n_{r_j} + n_{r_j}^2 \quad (3)$$

where $K_j = x_j^2 + y_j^2$ and $R = x^2 + y^2$. For the AOA α_j in (2), it can be expressed as

$$\alpha_j = \tan^{-1} \left(\frac{y - y_j}{x - x_j} \right) \Rightarrow x_j \sin \alpha_j - y_j \cos \alpha_j = x \sin \alpha_j - y \cos \alpha_j \quad (4)$$

To have non-singular matrices, β_j is formulated using a different approach as shown in right plane of Fig. 1. The chord of the circle (represented by the line RD_j and RD_j') can be used to derive the relationship between the devices as follows:

$$\left(x_j' - x_j\right)^2 + \left(y_j - y_j'\right)^2 = \left(2r_j^o \sin \frac{n_{\beta_j}}{2}\right)^2, \quad x_j' = x - r_j^o \cos(\beta_j - \pi), \quad y_j' = y - r_j^o \sin(\beta_j - \pi) \quad (5)$$

(5) can be rearranged and expressed as

$$K_j + r_j^{o2} - 2r_j^o (x_j \cos \beta_j + y_j \sin \beta_j) \approx 2 \left(x_j - r_j^o \cos \beta_j\right) x + 2 \left(y_j - r_j^o \sin \beta_j\right) y - R + r_j^{o2} n_{\beta_j}^2 \quad (6)$$

assuming for a $3\sigma_{\beta_j}$ confidence interval with $\sigma_{\beta_j} \leq 20°$, $\sin \left(\frac{n_{\beta_j}}{2}\right) \approx \frac{n_{\beta_j}}{2}$.

The accuracy of α_j and β_j in (4) and (6) can be further improved by the straight-line angle relationship in the LOS path as shown in left plane of Fig. 2. This is an advantage which cannot be realized in unidirectional estimation. In Fig. 2, α_j' and β_j' have the following relationships with α_j and β_j

$$\alpha_j' = \alpha_j, \quad \beta_j' = 360° - \beta_j \quad (7)$$

In ideal conditions with no noise perturbation, the following relationship will always hold

$$\alpha_j' + \beta_j' = \alpha_j^{o'} + \beta_j^{o'} = 180° \quad (8)$$

as α_j and β_j are equal to α_j^o and β_j^o respectively. With measurement noise, (8) becomes an inequality. However, the equality relationship can be reestablished through the following proposed criteria:

$$e_j = \alpha_j' + \beta_j' - 180°, \quad \alpha_j'' = \alpha_j' - e_j f_{\alpha_j}, \quad \beta_j'' = \beta_j' - e_j f_{\beta_j} \quad (9)$$
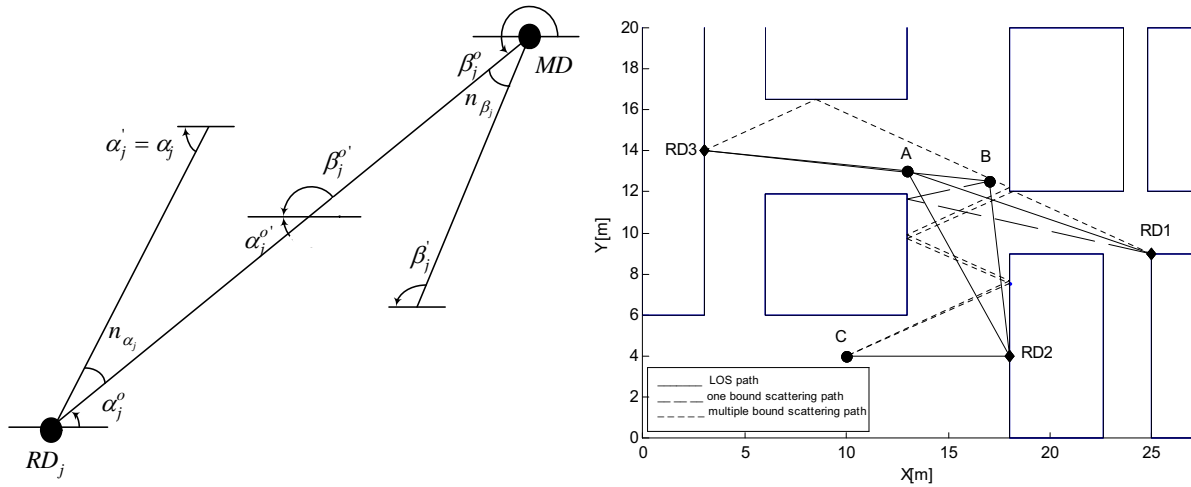
Figure 2: (Left) Geometrical relationship between the angle of arrival for a LOS path. (Right) Geometrical layout for the experiment in Nanyang Technological University, EEE S1, level B3 (S1-B3).

where α_j'' and β_j'' are the new AOAs at the RDs and MD respectively after the adjustment. e_j is defined as the noise angle difference for RD_j and f_{α_j}, f_{β_j} are the error weighting factors given as

$$f_{\alpha_j} = \frac{\sigma_{\alpha_j}}{\sigma_{\alpha_j} + \sigma_{\beta_j}}, \quad f_{\beta_j} = \frac{\sigma_{\beta_j}}{\sigma_{\alpha_j} + \sigma_{\beta_j}} \tag{10}$$

With unknown MD coordinate $\mathbf{Z_a} = [x, y, R]^T$, the new AOAs are put into (4), (6) and altogether with (3), the joint localization matrices can be formulated as $\mathbf{h} = \mathbf{G_a Z_a} + \boldsymbol{\psi}_a$ with $\mathbf{h} = \left[\mathbf{h}_{d1}^T, \cdots, \mathbf{h}_{dN}^T, \mathbf{h}_{\theta 1}^T, \cdots, \mathbf{h}_{\theta N}^T\right]^T$,

$$\mathbf{G_a} = [\mathbf{G}_{d1}, \mathbf{G}_{d1}, \cdots, \mathbf{G}_{dN}, \mathbf{G}_{dN}, \mathbf{G}_{\alpha 1}, \mathbf{G}_{\beta 1}, \cdots, \mathbf{G}_{\alpha N}, \mathbf{G}_{\beta N}]^T,$$

$$\mathbf{h}_{dj} = \left[l_j^2 - K_j, r_j^2 - K_j\right]^T, \mathbf{G}_{dj} = [-2x_j, -2y_j, 1]^T,$$

$$\mathbf{G}_{\alpha j} = \left[\sin \alpha_j^o, -\cos \alpha_j^o, 0\right]^T, \mathbf{h}_{\theta j} = \left[x_j \sin \alpha_j^o - y_j \cos \alpha_j^o, K_j + r_j^{o2} - 2r_j^o \left(x_j \cos \beta_j^o + y_j \sin \beta_j^o\right)\right]^T,$$

$$\mathbf{G}_{\beta j} = \left[2\left(x_j - r_j^o \cos \beta_j^o\right), 2\left(y_j - r_j^o \sin \beta_j^o\right), -1\right]^T, \boldsymbol{\psi}_a = 2\mathbf{Bn}_1 + \mathbf{Cn}_2 + \mathbf{Dn}_3 + \mathbf{En}_4 + \mathbf{Fn}_5 + \mathbf{Gn}_6$$

$$\mathbf{B} = diag\left\{l_1^o, r_1^o, \cdots, l_N^o, r_N^o, \frac{1}{2}\left(1 - f_{\alpha_1}\right)\lambda_1, r_1^o\left(1 - f_{\beta_1}\right)\gamma_1, \cdots \frac{1}{2}\left(1 - f_{\alpha_N}\right)\lambda_N, r_N^o\left(1 - f_{\beta_N}\right)\gamma_N\right\}$$

$$\mathbf{C} = diag\left\{0, 0, \cdots, 0, 0, f_{\alpha_1}\lambda_1, 2f_{\beta_1}r_1^o\gamma_1, \cdots f_{\alpha_N}\lambda_N, 2f_{\beta_N}r_N^o\gamma_N\right\}$$

$$\mathbf{D} = diag\left\{0, 0, \cdots, 0, 0, f_{\alpha_1}\kappa_1 n_{\alpha_1}^{\max}, \left(2f_{\beta_1}r_1^o\upsilon_1 + r_1^{o2}\right)n_{\beta_1}^{\max}, \cdots f_{\alpha_N}\kappa_N n_{\alpha_N}^{\max}, \left(2f_{\beta_N}r_N^o\upsilon_N + r_N^{o2}\right)n_{\beta_N}^{\max}\right\}$$

$$\mathbf{E} = diag\left\{0, 0, \cdots, 0, 0, f_{\alpha_1}^2\lambda_1, 2f_{\beta_1}^2 r_1^o\gamma_1, \cdots f_{\alpha_N}^2\lambda_N, 2f_{\beta_N}^2 r_N^o\gamma_N\right\}$$

$$\mathbf{F} = diag\left\{0, 0, \cdots, 0, 0, \left(f_{\alpha_1}^2 - f_{\alpha_1}\right)\kappa_1, 2\left(f_{\beta_1}^2 - f_{\beta_1}\right)r_1^o\upsilon_1, \cdots \left(f_{\alpha_N}^2 - f_{\alpha_N}\right)\kappa_N, 2\left(f_{\beta_N}^2 - f_{\beta_N}\right)r_N^o\upsilon_N\right\}$$

$$\mathbf{G} = diag\left\{1, 1, \cdots, 1, 1, 0, 0 \cdots 0, 0\right\},$$

$$\lambda_j = \cos\alpha_j^o\left(x^o - x_j\right) + \sin\alpha_j^o\left(y^o - y_j\right), \gamma_j = \cos\beta_j^o\left(y_j - y^o\right) - \sin\beta_j^o\left(x_j - x^o\right)$$

$$\kappa_j = \sin\alpha_j^o\left(x^o - x_j\right) - \cos\alpha_j^o\left(y^o - y_j\right), \upsilon_j = \cos\beta_j^o\left(x_j - x^o\right) + \sin\beta_j^o\left(y_j - y^o\right) \quad j = 1 \cdots N$$

$$\mathbf{n}_1 = [n_{l_1}, n_{r_1}, \cdots, n_{l_N}, n_{r_N}, n_{\alpha_1}, n_{\beta_1}, \cdots n_{\alpha_N}, n_{\beta_N}]^T, \mathbf{n}_2 = [0, 0, \cdots, 0, 0, n_{\beta_1}, n_{\alpha_1}, \cdots n_{\beta_N}, n_{\alpha_N}]^T$$

$$\mathbf{n}_3 = [0, 0, \cdots, 0, 0, n_{\alpha_1}, n_{\beta_1}, \cdots n_{\alpha_N}, n_{\beta_N}]^T,$$

$$\mathbf{n}_4 = [0, 0, \cdots, 0, 0, n_{\alpha_1}^2 n_{\beta_1}, n_{\beta_1}^2 n_{\alpha_1}, \cdots n_{\alpha_N}^2 n_{\beta_N}, n_{\beta_N}^2 n_{\alpha_N}]^T$$

$$\mathbf{n}_5 = [0, 0, \cdots, 0, 0, n_{\alpha_1} n_{\beta_1}, n_{\alpha_1} n_{\beta_1}, \cdots n_{\alpha_N} n_{\beta_N}, n_{\alpha_N} n_{\beta_N}]^T, \mathbf{n}_6 = \left[n_{l_1}^2, n_{r_1}^2, \cdots n_{l_N}^2, n_{r_N}^2, 0, 0 \cdots 0, 0\right]^T$$

$n_{\alpha_j}^{\max}$ and $n_{\beta_j}^{\max}$ are defined as the upper bound of the AOA errors for n_{α_j} and n_{β_j} respectively and are equated to $3\sigma_{\alpha_j}$ and $3\sigma_{\beta_j}$ respectively in the experiment. Similarly as in [1], $\mathbf{n}_6$ is neglected. As the odd order moment of a Gaussian random variable is null (moment of $\mathbf{n}_4$), the above joint

estimator will be an unbiased estimator. The first-step LS solution will be given as

$$\mathbf{Z_a} = \left(\mathbf{G_a}^T\mathbf{\Psi}^{-1}\mathbf{G_a}\right)^{-1}\mathbf{G_a}^T\mathbf{\Psi}^{-1}\mathbf{h}, \quad \mathbf{\Psi} = E\left[\boldsymbol{\psi}_a\boldsymbol{\psi}_a^T\right] \tag{11}$$

To solve for $\mathbf{Z_a}$ in (11), we need the matrices $\mathbf{\Psi}$, $\mathbf{h}$ and $\mathbf{G_a}$ which require the unknown actual distances l_j^o, r_j^o and angles α_j^o, β_j^o for $j = 1 \cdots N$. Therefore, we first use the measured values l_j, r_j, α_j, β_j instead of the actual values, and set $\mathbf{\Psi} = \mathbf{I}_{4N \times 4N}$ to estimate an initial solution. The corresponding matrices $\mathbf{\Psi}$, $\mathbf{h}$ and $\mathbf{G_a}$ are recalculated with the new estimated l_j, r_j using this initial solution. α_j, β_j in the above matrices are also readjusted using (9). A better estimate of $\mathbf{Z_a}$ is obtained by using the new matrices $\mathbf{\Psi}$, $\mathbf{h}$ and $\mathbf{G_a}$. The process can be repeated until the solution converges. Our simulations show that one iteration is sufficient for obtaining a good estimation of $\mathbf{Z_a}$. The second-step LS approach for obtaining the MD coordinate $\mathbf{Z} = [x, y]^T$ will be similar to [1] where $\mathbf{Z} = \pm\sqrt{\mathbf{Z_p}}$ and

$$\mathbf{Z_p} = \left(\mathbf{G_p^T}\mathbf{\Psi_p}^{-1}\mathbf{G_p}\right)^{-1}\mathbf{G_p^T}\mathbf{\Psi_p}^{-1}\mathbf{h_p}, \ \mathbf{h_p} = \begin{bmatrix} [\mathbf{Z_a}]_1^2 \\ [\mathbf{Z_a}]_2^2 \\ [\mathbf{Z_a}]_3 \end{bmatrix}, \ \mathbf{G_p} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}, \ \mathbf{\Psi_p} = 4\mathbf{B_p}\mathrm{cov}\left(\mathbf{Z}_a\right)\mathbf{B_p},$$

$$\mathbf{B_p} = \mathrm{diag}\left\{x^o, y^o, 0.5\right\}, \ \mathrm{cov}\left(\mathbf{Z_a}\right) = \left(\mathbf{G_a^{oT}}\mathbf{\Psi}^{-1}\mathbf{G_a^o}\right)^{-1} \tag{12}$$

Similarly, (x^o, y^o) and $\mathbf{G_a^o}$ can be replaced by the (x, y) in $\mathbf{Z_a}$ and $\mathbf{G_a}$ respectively during estimation. This process is carried out once to render an accurate result.

3. RESULTS AND DISCUSSIONS

Figure 2 (right plane) depicts the environment at Nanyang Technological University, School of Electrical and Electronic Engineering (EEE), block S1, level B3 [6] that is used to evaluate the performance of our proposed localization scheme with the conventional localization schemes [1, 2]. Three reference devices were located at RD_1 (25 m, 9 m), RD_2 (18 m, 4 m), RD_3 (3 m, 14 m). For simplicity, the first dominant path is extracted from each RD. The paths' parameters (TOA and AOA) are obtained through the ray tracing methodology [6, 7] and subject to Gaussian random variable noise with zeros mean. 10,000 MD locations which are uniformly distributed in the environment are simulated to obtain the Average Location Error (ALE) performance comparison between our proposed localization scheme and the conventional TOA, TOA/AOA (modified from TDOA/AOA scheme to TOA/AOA scheme for stricter comparison as TOA/AOA scheme has better performance than TDOA/AOA scheme) localization schemes with three RDs [1, 2]. The conventional TOA localization scheme [1] has its own mitigation technique while the conventional TOA/AOA localization scheme [2] is coupled with the mitigation technique in [3]. The location
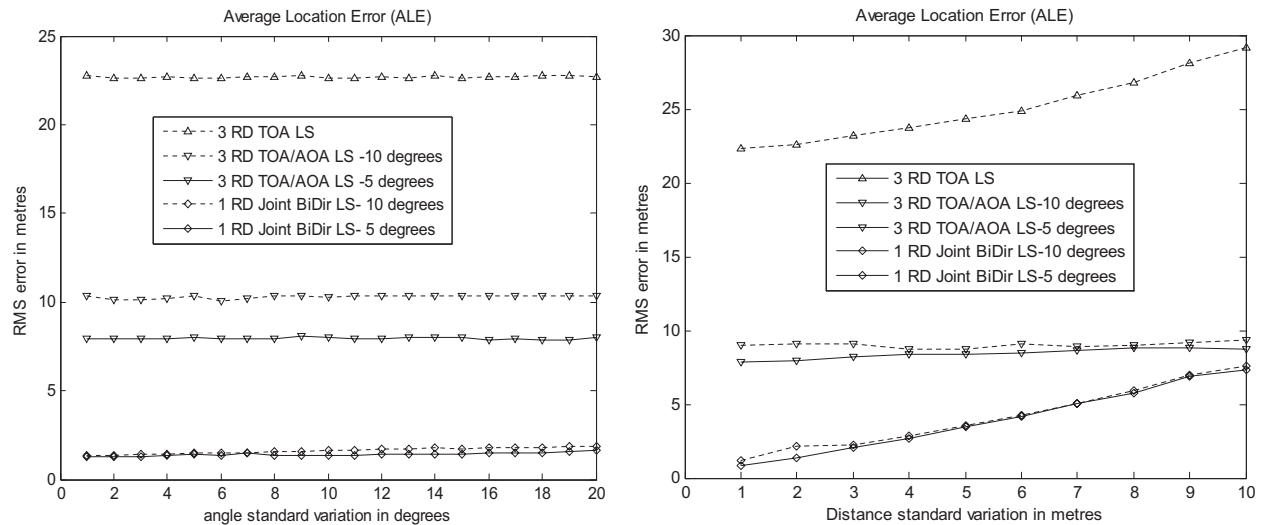

Figure 3: (Left) Average Location Error (ALE) when the distance standard deviation $\sigma_d = 2\,\mathrm{m}$, the AOA standard deviation $\sigma_\alpha = 5°$, $10°$ and σ_β varies. (Right) Average Location Error (ALE) when the AOA standard deviation $\sigma_\alpha = \sigma_\beta = 5°$, $10°$ and σ_d varies.

RMS error σ_{rms}, $\sqrt{E\left[(x-x^o)^2+(y-y^o)^2\right]}$ is computed based on 10,000 independent runs. $n_{\alpha_j}^{\max}$ and $n_{\beta_j}^{\max}$ were set to $3\sigma_{\alpha_j}$ and $3\sigma_{\beta_j}$ respectively.

Figure 3 (left plane) illustrates the various schemes' ALE performance as σ_β of our proposed joint LS scheme is varied. The standard deviations for the distance and the angle (σ_d and σ_α) are chosen to be the typical values of $2\,\mathrm{m}$ and ($5°$, $10°$) respectively. Our proposed joint LS scheme using one LOS RD outperforms the conventional schemes for all variations in σ_β. As σ_β escalates, the proposed joint LS scheme's ALE does not escalate substantively, because of the bi-directional angle relationship (9). Fig. 3 (right plane) depicts ALE performance as σ_d escalates while both $\sigma_\alpha = \sigma_\beta = 5°$, $10°$. As shown, the proposed joint LS scheme using one LOS RD outperforms the conventional scheme. It is worth noting that our proposed localization scheme is not sensitive to changes in f_{α_j}, f_{β_j} due to changes in σ_α, σ_β. This reflects the robustness of the AOA adjustment technique.

4. CONCLUSIONS

A novel approach for source localization using bi-directional joint estimation of TOA/AOA and only one LOS reference device has been proposed. The joint LS implementation coupled with AOA adjustment and NLOS path rejection techniques have been devised. The new approach is compared to two conventional localization schemes, the unidirectional TOA scheme and the disjoint unidirectional TOA/AOA scheme, in a typical environment. The ALE simulation results indicate that the joint LS, which uses only one LOS reference device, outperforms the conventional schemes for all cases of error variations with a much higher margin.

REFERENCES

1. Wang, X., Z. X. Wang, and B. O'Dea, "A TOA-based location algorithm reducing the errors due to Non-Line-of-Sight (NLOS) propagation," *IEEE Trans. Veh. Tech.*, Vol. 52, 112–116, Jan. 2003.
2. Cong, L. and W. H. Zhuang, "Hybrid TDOA/AOA mobile user location for wideband CDMA cellular systems," *IEEE Trans. Wireless. Commun.*, Vol. 1, 439–447, July 2002.
3. Cong, L. and W. H. Zhuang, "Nonline-of-sight error mitigation in mobile location," *IEEE Trans. Wireless. Commun.*, Vol. 4, 560–572, Mar. 2005.
4. Wang, K. and H. Ge, "Joint space-time channel parameter for DS-CDMA system in multipath Rayleigh fading channels," *IEE Electronics Letters*, Vol. 37, 458–460, Mar. 2001.
5. Alexiou, A. and M. Haardt, "Smart antenna technologies for future wireless systems: Trends and challenges," *IEEE Commun. Mag.*, Vol. 42, 90–97, Sep. 2004.
6. Tan, S. Y. and H. S. Tan, "Modelling and path loss measurements for microcellular communications in a corridor complex," *IEE Electronics Letters*, Vol. 30, 2008–2010, 1994.
7. Sun, Q., S. Y. Tan, and K. C. Teh, "Analytical formulae for path loss prediction in urban street grid microcellular environments," *IEEE Trans. Veh. Tech.*, Vol. 54, 1251–1258, July 2005.

Propagation of Ultra Wideband Signals in Automotive Environment

Ching-Ping Wang and Wen-Jiao Liao

National Taiwan University of Science and Technology, Taiwan

Abstract— This work investigates the propagation of the ultra wideband signal in automotive environment, especially the car underhood. Measurements as well as simulations using high frequency numerical electromagnetic tool are performed to examine the interaction between broadband signals and obstacles in the propagation environment. The measurement setup contains two broadband antennas. A large metal plane and a box with conducting surfaces are put in the environment in various configurations. Transmission magnitudes and phases are recorded with a network analyzer by sweeping from 2 to 6 GHz. By applying the Fourier transform, the frequency spectrum is converted into time domain to examine the propagation characteristics of the environment. With adjustments in the propagation environment setup, the location of primary scatters can be identified.

1. INTRODUCTION

UWB technology is a transmission scheme that spreads information over a large frequency band. It has many merits such as high data rate, low power consumption, low cost, high security, and anti-interference. For the past few years, its primary application target is aimed to high speed transmission in Wireless Personal Area Networks (WPAN), which facilitates the multimedia transmission need for video or audio contents. In practice, it can substitute connecting cables within 10 meters [1, 2].

This research is motivated by the need of wireless in-car communication. New electronic devices such as the tire pressure monitoring and collision avoidance radar were devised and implemented on cars. Due to the variation in sensor locations, the wiring issue may impose a critical problem for such an after market product [3]. Since the automotive environment is rich in scattering and reflection, quality of conventional narrow band communication can be sensitive to transmitter and receiver positions. Therefore, we are interested in evaluating the propagation characteristics of a broadband signal in a complicated environment to study its impact on impulse signals.

Behaviors of broadband signal propagation were constantly investigated. For example, the effect of laptop computer shielding on UWB signal propagation in an indoor environment has been studied in [2]. Measurements of UWB signal attenuation in car environment is conducted in [3]. Ultra wideband signal propagation in desktop environments is investigated in [4]. In this work, we conducted both simulations and measurements that emulating broadband signal propagation in an underhood environment. Instead of directly measuring an impulse signal, we utilized a network analyzer to sequentially record the S_{21} transmission in a broad frequency range form 2 to 6 GHz. The recorded frequency spectrum is then converted into the time domain via Fourier transform, which synthesizes an impulse signal of extremely short period to reveal the propagation environment characteristics. By adjusting the propagation setup, we can identify the locations of primary scatters and analyze the transmission behavior under different scattering environments.

2. BROADBAND ANTENNA DESIGN AND PERFORMANCE EVALUATION

In order to transmit broadband signals, antennas that comply with the UWB band were fabricated. A broadband monopole is adopted from [5] with a microstrip line feed. To validate its performance, the reflection coefficient spectrum and radiation patterns of the antenna are measured.

The prototype antenna shown in Figure 1(a) was fabricated with a 60×60 mm substrate. In order to tune the antenna into the desired operation band, copper pasters were applied around the radiator for tuning. Figure 1(b) shows the adjusted antenna. Figure 2 shows return loss curves of three monopoles in different sizes. The red curve yields a return loss below -7.14 dB from 2 to 6 GHz. The corresponding antenna, which is 4/3 times bigger than the original design, is then used as transmitting and receiving antenna as well.

The measured radiation patterns indicate that the maximum gain of co-polarization appears at $\phi = 0°$ and $150°$ on the E-plane. The polarization is vertical and the maximum gain is about 3.3 dB and is nearly omni-directional on the azimuth plane below 4.5 GHz. Note the realized gain values of cross-polarization on the Azimuth plane become lager from 4.5 to 6 GHz, which indicates

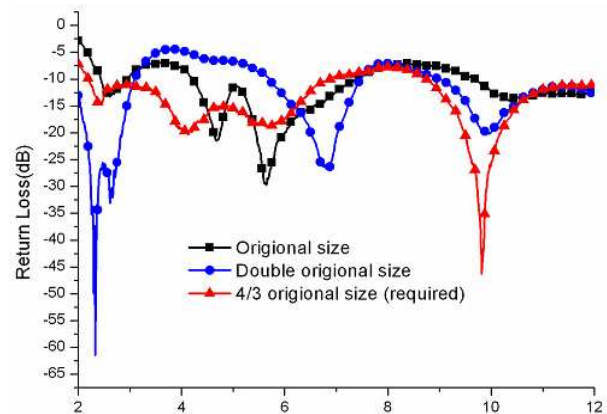(a) Original antenna   (b) Adjusted antenna

Figure 1: Fabricated broadband monopole antennas.

Figure 2: Return loss spectra of prototype antennas.

that the radiation has both vertical and horizontal components at the high band. Nevertheless, this antenna is vertically polarized in principle below 4.5 GHz.

3. EXPERIMENTAL MODAL SETUP AND SIMULATION METHOD

The propagation experiment was setup in an EMC chamber. Two aforementioned monopole antennas were used for signal transmission and reception. Measurements with various test setups were performed. The setup shown in Figure 3(a) is denoted as "No Ground". The one shown in Figure 3(b) is denoted as "Add Ground". The cart's length, width and height are 91 cm, 46 cm and 88 cm, respectively. Receiving as well as transmitting antenna were located above the cart's edge by 10 cm.

(a) No Ground   (b) Add Ground

Figure 3: Propagation measurement environment.

We also wrapped a paper carton with aluminum foil and used it as an obstacle between two antennas. The box was put in the middle of the ground plane. The dimensions of the box are $19 \, \text{cm} \times 26 \, \text{cm} \times 29 \, \text{cm}$ as shown in Figure 4. Because there are six posture conditions possible with the box, six cases were defined and measurements were performed according. The length, width and height of the box posture are labeled in Figure 4.

The experiment was performed in two steps. We first measured the amount of transmission with a network analyzer. A total of 801 points were measured from 2 to 6 GHz. We next performed Fourier transform to convert the frequency spectrum into the time domain to visualize the propagation delay introduced by the environment.

In addition to the propagation measurement described above, we also used a high frequency numerical tool to simulate the broadband propagation characteristics. The tool used here is Numerical Electromagnetic Code-Basic Scattering Code (NEC-BSC). This software adopts ray tracing techniques in conjunction with Uniform Geometrical Theory of Diffraction (UTD). It computes field contributions from various scattering mechanisms such as reflection and diffraction. The distribu-

tions of field intensity in far field as well as near field can be acquired.
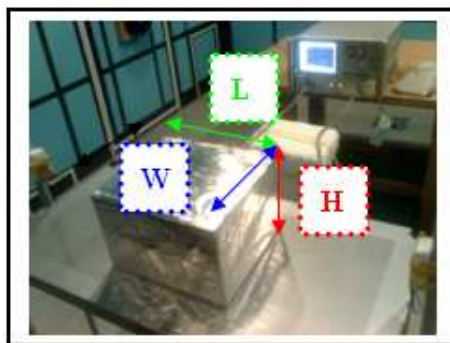

Figure 4: The metal box used in propagation measurement.

The simulation model was setup to emulate the measurement environment. The components incorporated include the transmitting antenna, receiving antenna, the ground plane and the metal box. In the simulation, a frequency sweep from 2 to 6 GHz was performed.

Note that the simulation model is different to the actual measurement environment in several ways. The transmitting antenna is a vertically polarized half wavelength dipole whose length is varied with the transmitting frequency. The receiving antenna is replaced with an observation point to record the electric fields. The cart in the measurement is replaced by a 100 cm × 50 cm metal plate in the simulation. The ceiling and walls of the EMC are omitted in the simulation.

Similar to the experiment process, the simulated results were also transformed into time domain to examine the interaction between the waves and obstacles in the environment.

4. OBSERVATIONS OF EXPERIMENT RESULTS

Figure 5 shows the measured transmission spectra of two setups. For both cases, the amount of transmission decreases as the frequency goes up. Though both curves exhibit volatile variations, one can tell that the curve of the "Add Ground" case is about 3 dB higher than the one of the "No Ground" case. This indicates that the cart is an effective reflecting surface.

From Figure 6, it can be found that the amount of transmission of all cases containing the metal box is much lower than the "No Box" case, which is referred as the "Add Ground" case in Figure 5. The difference is above 10 dB below 4.5 GHz. The reduction rate of the "No Box" case becomes larger above 4.5 GHz. This is because the antenna polarization is not strictly vertical above 4.5 GHz and therefore suffers gain reduction. The levels of all cases fall as the frequency goes up. The apparent gap between the "No Box" and cases with the metal box in place indicates that the box, which blocks the line-of-sight propagation, contributes substantial attenuation. Similar
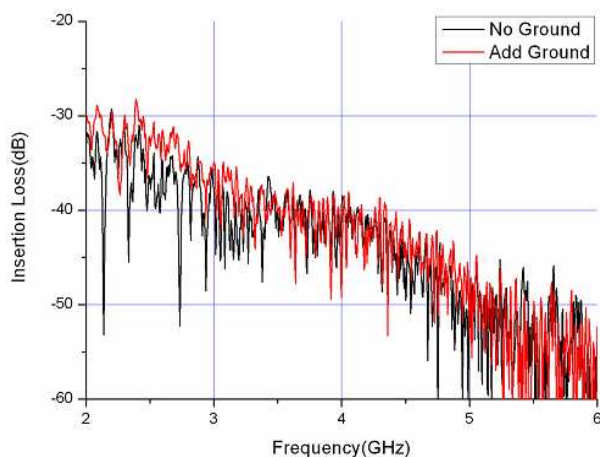

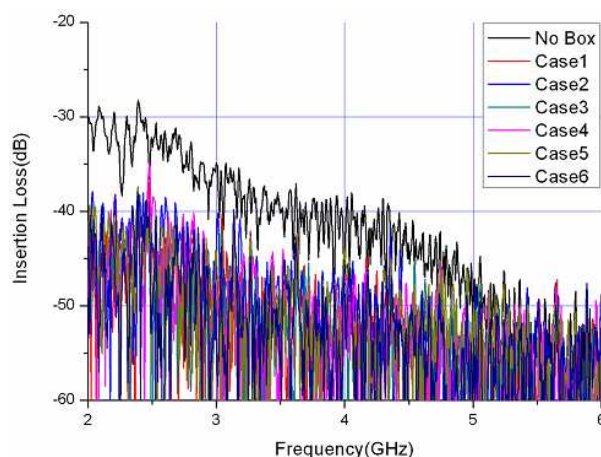Figure 5: Transmission spectra with and without the ground plane.

Figure 6: Transmission spectra with and without the metal box.

attenuation phenomenon of metal obstacles is found in [3].

We next converted the measured results into the time domain. Since we recorded transmission results of a relatively broad bandwidth, a fine time domain resolution can be achieved according to the Fourier transform relationship. Furthermore, because the velocity of electromagnetic wave is a constant in a homogenous medium, the signal's time domain variation is the same as its spatial domain variation.

The relationships between frequency bandwidth and time sampling interval as well as the one between frequency sampling interval and time duration is governed by the Fourier transform. According to the bandwidth and frequency resolution of the experiment, the associated spatial resolution is 7.5 cm and the total length is 6000 cm. According to the transform results, most of the energy comes in at 105 cm for the "Add Ground" case. Its magnitude is twice as strong as the one of "No Ground" case. In comparison, the transmitted energy for cases with the metal obstacle is much weaker than the ones without blockage. Details of the propagation features for cases with the metal box are plotted in Figure 7. According to the boundary condition of conducting surfaces, the electric field is normal to the metal plate. Because the transmitting antenna is vertically polarization in general, the waves are more likely to crawl over the top surface. Moreover, the surface wave is attenuated in proportion to the travelling distance. This phenomenon can be explained using plots in Figure 7. Because the distance (L) of the top surface in "Case 2" is 3 cm shorter than the distance in "Case 1", the attenuation is weaker and the peak is higher in "Case 2". The height (H) is higher in "Case 3", and the transmitted signal must impinge the top edge with a larger incident angle. Therefore, the amount of incidence is smaller and the peak is weaker than the one in "Case 2". As to "Case 4", although the height is 26 cm, the peak is still high since the length is 19 cm only. In general, the locations of the peaks do not change much. The first peak positions for "Case 1" to "Case 4" are 105 cm, 112.5 cm, 120 cm and 112.5 cm.

(a) H, W, L=19, 26, 29 cm

(b) H, W, L= H= 19, 29, 26 cm

(c) H, W, L=26, 19, 29 cm

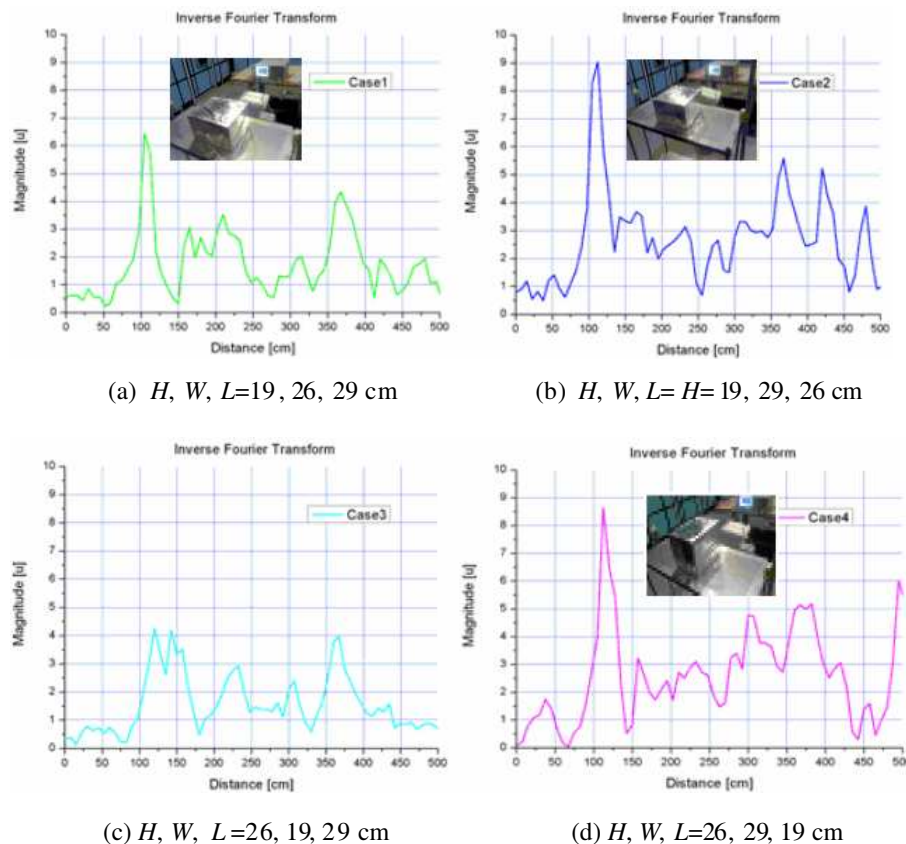(d) H, W, L=26, 29, 19 cm

Figure 7: Comparisons of spatial domain distributions with metal box.

The NECBSC simulation results are shown in Figure 8(a). The most obvious difference to the measured results is that the transmission spectra are very stable. This is because the simulation model, which contains no walls, is lack of multiple reflection terms. The "No Ground" case, which

contains antennas only, yields a flat line. In other cases, the transmission amplitudes, similar to the measurement, suffer more attenuation as the frequency goes up.

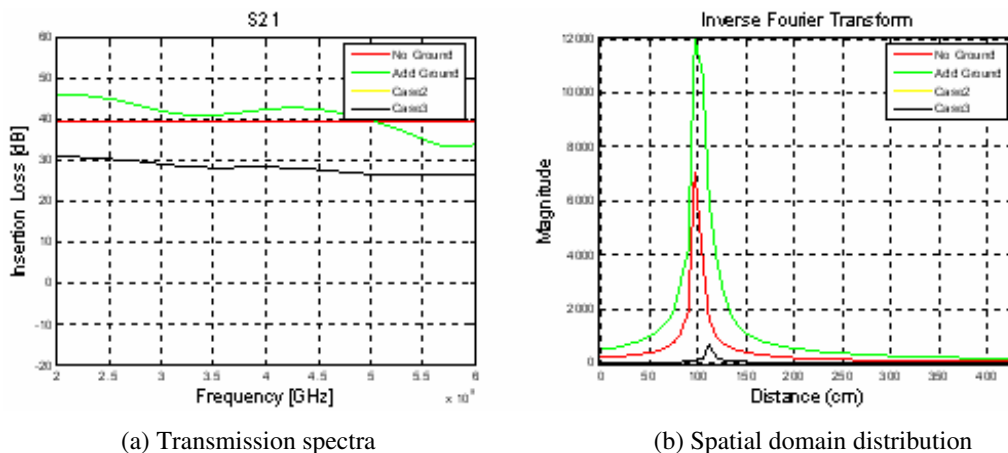(a) Transmission spectra   (b) Spatial domain distribution

Figure 8: Simulated frequency and time domain transmission results.

The spatial domain results are also derived by applying the Fourier transform. The peak values found in "No Ground", "Add Ground", and all cases with metal box are in proportion to the measurement results. The peak values of "Add Ground" and "No Ground" both occur at 97.5 cm. The peak value for the case with metal box appears at 112.5 cm.

5. EXECUTIVE SUMMARY

In this work, we have explored the UWB propagation characteristics in the presence of obstacles, which is common in automotive environment. The impulse signal responses at various setups were synthesized via Fourier transform on frequency domain results. Both measurement and simulation approaches were attempted. To facilitate the measurements, we first fabricated broadband monopole antennas. The return loss spectrum and radiation patterns of the prototype antenna show that its performance can comply with broadband signal transmission needs. In the measurement, the 2 to 6 GHz bandwidth provides a 7.5 cm spatial resolution to reveal the transmission path of UWB signal and the blockage effect of the metal obstacle. As to the simulation results, the changes in the transmission spectra behave similar to the measured ones. However, because multiple reflections due to room walls were not considered and the antenna used in the simulation has no frequency dependence, the volatile spectra shown in the measurements were not seen in the corresponding simulation results. A more detailed simulation model is needed for future exploration on propagation of UWB signal in the presence of obstacles.

REFERENCES

1. Chen, Z. N., "UWB antennas: Design and application," *6th International Conference on Information, Communications & Signal Processing*, 1–5, 2007.
2. Siripornnarachseema, P., S. Promwong, P. Sithiyopaskul, R. Muanghlua, and N. Sukutamtanti, "The effect of laptop computer shadowing on UWB propagation in an indoor environment," *IEEE International Symposium on Communications and Information Technology, ISCIT 2005*, Vol. 2, 975–978, 2005.
3. Berens, F., H. Dunger, S. Czarnecki, T. Bock, R. Reuter, S. Zeisberg, J. Weber, and J. F. Guasch, "UWB car attenuation measurements," *Mobile and Wireless Communications Summit*, 1–5, 2007.
4. Suzuki, Y. and T. Kobayashi, "Ultra wideband signal propagation in desktop environments," *IEEE Conference on Ultra Wideband Systems and Technologies*, 493–497, 2003.
5. Ma, T.-G. and S.-J. Wu, "Ultrawideband band-notched folded strip monopole antenna," *IEEE Transactions on Antennas and Propagation*, Vol. 55, No. 9, 2473–2479, 2007.

A Simplified Statistical Modeling of Radioclimatological Parameters for LOS Links in South Africa

P. K. Odedina and T. J. Afullo

University of KwaZulu-Natal Durban, P. O. Box 4000, South Africa

Abstract— Different techniques has been proposed by various authors in different regions of the world to model the solution to radio wave propagation problems, but most of these techniques have not considered modeling the primary parameters on which other radioclimatology factors are based. This paper has therefore focused on working from the basics, by considering the initial parameters on which radioclimatological factors were based. These initial parameters namely; temperature, pressure and humidity were modeled with vertical height increase using one year radio propagation data and six years radio propagation data in Durban, KwaZulu Natal province of South Africa. A simplified statistical technique is employed in the analysis of this data.

1. INTRODUCTION

The propagation of electromagnetic waves around the earth is influenced by the properties of the earth and the atmosphere [1–5]. The earth is an inhomogeneous body whose electromagnetic properties vary considerably as we go from one point to another, its properties varying with temperature, pressure and humidity.

According to ITU-R Recommendation P.530 [3], the propagation loss on a terrestrial line-of-sight path relative to free space loss is the sum of different contributions which includes: attenuation due to atmospheric gases; diffraction fading due to obstruction or partial obstruction of the path; fading due to multipath and attenuation due to precipitation. Each of these contributions has its own characteristic as a function of frequency, path length and geographical location [6].

Some of the basic parameters in the field of radioclimatology includes but not limited to the weather parameters such as; temperature, pressure, humidity and water vapour pressure. Also included are the variables that account for the earth curvature such as effective earth radius factor (k-factor), surface roughness, multipath effect and geoclimatic factor. Parameters such as k-factor have been adequately studied, modeled and applied to line of sight link in Southern of Africa [7–9]. Temperature, pressure, humidity, water vapour, saturated vapour pressure and refractivity as it affects signal degradation and how they are related were discussed in [1, 10]. This paper discussed a simplified statistical modeling of the these basic parameters mentioned above and the steps involved were explained in the next section.

2. MODELING OF RADIOCLIMATOLOGICAL PARAMETERS

One of the main challenges of radio propagation studies is the acquisition of data to verify different proposed models. The data used in this paper therefore are two parts, both of which are radiosonde measurements spanning different periods in Durban, KwaZulu-Natal Province of South Africa. The first part of the data is a one year radiosonde data measured over a height range of 0–500 m a.s.l., in 2004. From these set of data, parameters within the first 0–100 m height a.s.l were extracted. The second set of data is a radiosonde measurement for six years from (1990–1995) over a height range of 0–26000 m a.s.l. These sets of data come very handy in the statistical analysis for the model formulation.

The one year data was initially used to start the model formulation process, but more data is needed in the 0–100 m height level since this is where signal loss is experienced most in terrestrial line of sight link design [1]. To fill in this data gap, the spatial interpolation techniques discussed in [11] was implemented for each of the twelve months in 2004 for 0–500 m a.s.l. Having done this, the graph of each of the parameters were plotted against height for each of the twelve months in order to obtain a model equation.

The process explained above gives an initial model equation for each of temperature, pressure and humidity for the twelve calendar months. The second process embarked on is to modify and improve on these initial model equations. To do this, a simplified statistical approach was used. The Initial set of modeled equations was used to predict the missing parameters for the height range 0–26000 m a.s.l. Then the predicted parameters were matched against the measured parameters in order to observe how well the initially modeled equations predict the data points that exist from

measurement. It was observed from this exercise that the initially modeled equations were able to predict parameters fairly well with little disparity for the lower height ranges 0–200 m a.s.l., but data points in the higher height ranges specifically above 1000 m a.s.l. were not properly estimated by the model. At this point it was concluded that another method will be required to predict the higher height range parameters.

Finally, the data set within the height range 0–200 m a.s.l. for the six years were sorted and collated for both the measured and predicted parameters. Then the deviation between the measured and the predicted values for these data and the average of these deviations called mean temperature deviation (MTD), mean pressure deviation (MPD) and, mean humidity deviation (MHD) were determined. Next, the standard deviation of these parameters namely; standard deviation of temperature deviation (SDTD), standard deviation of pressure deviation (SDPD) and standard deviation of humidity deviation (SDHD) were also computed. These standard deviations then become the compensation factors which were used to modify the initial set of modeled equations for the twelve months.

The mean deviations give an indication whether the initially modeled equation over estimated or under estimated the true measured values. If the mean deviation for a particular parameter in a certain month is negative for instance, it means that modeled equation overestimated the true value for that month by a value given as the standard deviation for that month. Hence this standard deviation value is subtracted from the constant value in the initial equation to get the final modeled equation for that month and vice versa. The results obtained from this activity were presented in Table 1 in Section 3 for only three months in this paper. The equations used to arrive at the above solutions are stated below:

$$\text{MTD} = \frac{\sum (T_m - T_p)}{N} \tag{1}$$

$$\text{MPD} = \frac{\sum (P_m - P_p)}{N} \tag{2}$$

$$\text{MHD} = \frac{\sum (H_m - H_p)}{N} \tag{3}$$

where T_m = Measured temperature value (°C), T_p = Predicted temperature value (°C), P_m = Measured pressure value (hpa), P_p = Predicted pressure value (hpa), H_m = Measured humidity value (%), H_p = Predicted humidity value (%), N = Total Number of variables used in the analysis. Also for the compensation factor calculation, the following formulations are applicable:

$$\text{SDTD} = \sqrt{\frac{\sum (T_D - \text{MTD})^2}{(N-1)}} \tag{4}$$

$$\text{SDPD} = \sqrt{\frac{\sum (P_D - \text{MPD})^2}{(N-1)}} \tag{5}$$

Table 1: Compensation factor for temperature, pressure and humidity for sample months (1985–1990).

Month	Temperature (DegC)				Pressure (hpa)				Humidity (%)			
Jan	MTD	-0.37	Initial Model	T = -0.019h + 26.08	MPD	-5.88	Initial Model	P= -0.113h + 1019	MHD	11.50	Initial Model	H = -0.019h + 66.17
	SDTD	2.68	Final Model	T = -0.019h + 23.4	SDPD	3.95	Final Model	P= -0.113h + 1015.05	SDHD	10.99	Final Model	H = -0.019h + 77.16
Jun	MTD	-0.52	Initial Model	T = -0.040h + 22.69	MPD	-7.65	Initial Model	P= -0.120h + 1029	MHD	-4.11	Initial Model	H = -0.033h + 68.44
	SDTD	5.02	Final Model	T = -0.040h + 17.67	SDPD	6.50	Final Model	P= -0.120h + 1022.5	SDHD	17.58	Final Model	H = -0.033h + 50.86
Nov	MTD	-2.23	Initial Model	T = -0.029h + 25.99	MPD	6.33	Initial Model	P = -0.112h + 1010	MHD	-5.52	Initial Model	H = -0.020h + 81.76
	SDTD	2.99	Final Model	T = -0.029h + 23	SDPD	5.59	Final Model	P = -0.112h + 1015.59	SDHD	11.47	Final Model	H = -0.020h + 70.29

$$\mathrm{SDHD} = \sqrt{\frac{\sum (H_D - \mathrm{MHD})^2}{(N-1)}} \tag{6}$$

$$T_D = T_m - T_p, \quad P_D = P_m - P_p, \quad H_D = H_m - H_p, \tag{7}$$

where T_D = Temperature Deviation, P_D = Pressure Deviation and H_D = Humidity Deviation.

3. RESULTS

It should be noted from the sample results in Table 1. That H means humidity while h is the height variable. The three graphs in Figure 1 are some of the plots from thirty-six of such graphs from which our initially modeled equation in Table 1 was determined.
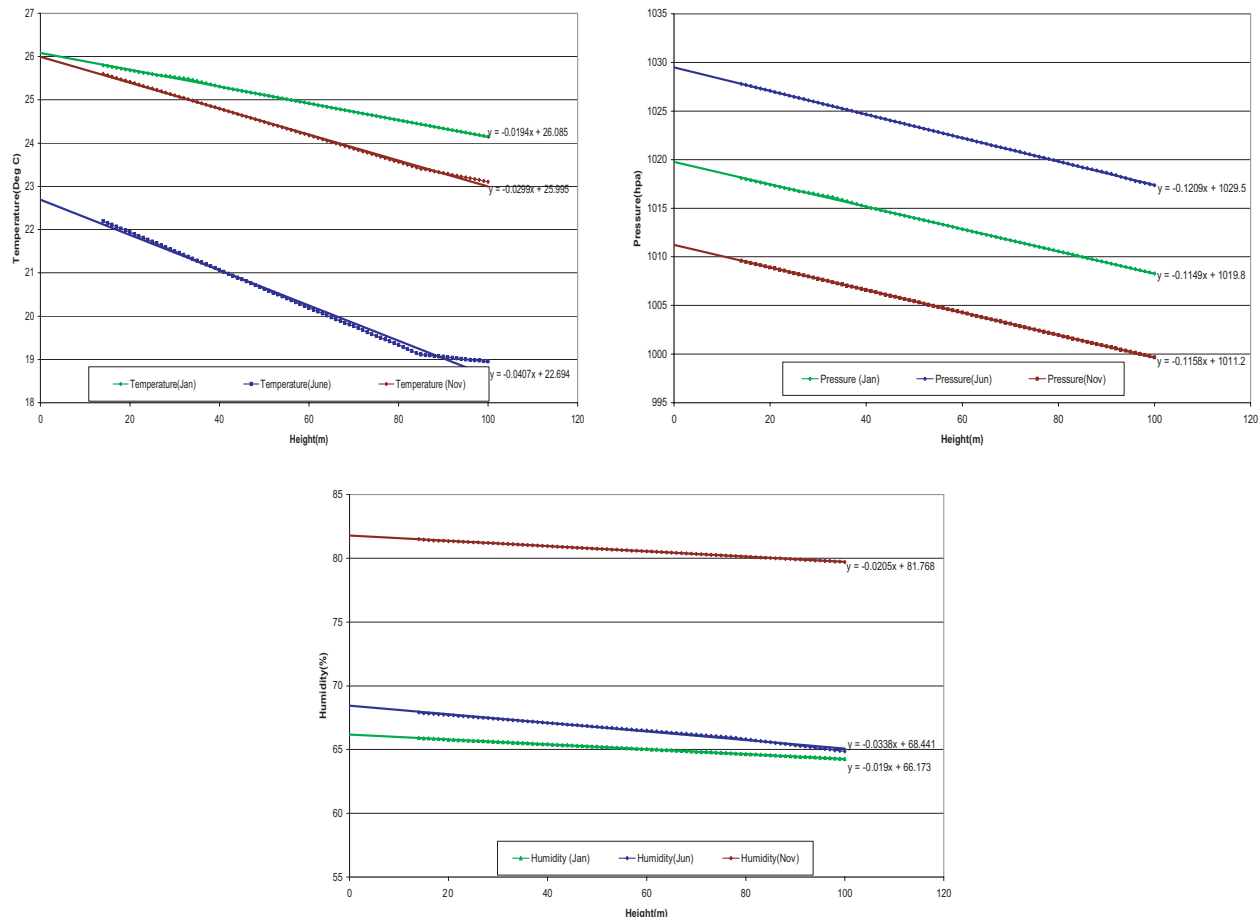
Figure 1: Plot of the different parameters (Temperature, Pressure, Heights) variation with height.

4. CONCLUSION

In this presentation, a simplified statistical approach had been combined with spatial interpolation technique to model three basic radioclimatological parameters for line of sight links in South Africa. The combination of these two techniques gives a reasonable prediction of temperature, pressure and humidity within the first 100 m a.s.l. The beauty of the approach is its simplicity; hence the approach can be implemented in any radioclimatological data analysis with less ambiguity. The results presented will be very used for radio system designer in South Africa. The model will be improved further with the availability of data in the other provinces of South Africa.

REFERENCES

1. ITU Radiocommunication Study Group III: "The radio refractive index: Its formula and refractivity data," ITU Radiocommunication Assembly, Document No. ITU-R P.453–7, July 11, 2000.

2. ITU Radiocommunication Assembly, "Propagation data and prediction methods required for the design of terrestrial line of sight systems," ITU-R, 530–10, 1982–1986–1990–1992–1994–1995–1997–1999–2000–2001.

3. ITU Radiocommunication Study Group III: "Effects of tropospheric refraction on radiowave propagation," ITU Radiocommunications Assembly, Document No. ITU-R, 834–4, 1992-1994-1997-1999-2003.

4. Olsen, R. L. and T. Tjelta, "Worldwide techniques for predicting the multipath fading distribution on terrestrial LOS links: Background and results of tests," *IEEE Transactions on Antenna and Propagation*, Vol. 47, No. 1, 157–170, January 1999.

5. Olsen, R. L., "Radioclimatological modeling of propagation effects in clear-air and precipitation conditions: recent advances and future directions," *Proceedings of the Third Regional Workshop on Radio Communications in Africa Radio Africa '99*, Gaborone Botswana, October 1999.

6. Afullo, T. J. and P. K. Odedina, "On the K-factor distribution and diffraction fading for Southern Africa," *South African Institute of Electrical Engineers Research Journal*, Vol. 97, No. 2, 172–181, June 2006.

7. Palmer, A. J. and D. C. Baker, "Predicting the long-term average of the effective earth radius factor for south africa using ground based observatons," *South African Institutte of Electrical Engineers Research Journal*, Vol. 97, No. 2, 182–185, June 2006.

8. Palmer, A. J. and D. C. Baker, "Predicting the monthly average cummulative distribution of the effecticve earth radius factor for south africa," *South African Institute of Electrical Engineers Research Journal*, Vol. 97, No. 2, 186–190, June 2006.

9. Odedina, P. K. and T. J. Afullo, "Effective earth radius factor (k-factor) determination and its application in Southern Africa," *Proceedings of the Second IASTED International Conference on Antennas, Radar and Wave Propagation*, 222–227, Banff, Alberta Canada, July 1–21, 2005.

10. Dabideen, A. S., M. Gopichund, and T. J. Afullo, "Study of clear-air ducting and fading in KwaZulu-Natal," *The Transactions of the South African Institute of Electrical Engineers Journal*, ISSN No.: 0038-2221, Vol. 96, No. 2, 121–132, March 2005.

11. Odedina, P. K. and T. J. Afullo, "Use of spatial interpolation technique for determination of geoclimatic factor and fade depth calculation in Southern Africa," *Proceeding of IEEE AFRICON 2007 Conference*, ISBN: 0–7803–8606–X. IEEE Catalog Number: 04CH37590C, Namibia, September 26–28, 2007.

Regular Polyhedron Antenna Array Design and Simulation for MIMO Systems

L. Wang and H. G. Wang

The Electromagnetics Academy at Zhejiang University, Zhejiang University
Hangzhou, China

Abstract— Some MIMO applications require antennas to be closely spaced, and it will cause the mutual coupling among antennas and high spatial correlation which will harmful for enhancing the system capacity. In order to solve this problem, many methods have been developed. The antenna array design is a feasible way to enhance the MIMO system performance. We developed the regular polyhedron antenna arrays. Regular polyhedron has excellent spatial symmetry, and it is vital in 3D environment. This kind of regular polyhedron antenna arrays exhibit lower mutual coupling effect and spatial correlation then the antenna array fixed on a plane. In numerical simulation, the efficient integral method MGLFIM is employed to fast calculate the input admittance and radiation pattern. Each regular polyhedron antenna array demonstrates excellent efficiency in enhancing the MIMO system capacity. To obtain the same system capacity, the distance between two monopoles can be smaller using regular polyhedron than using planar antenna array. With the same volume, the regular polyhedron antenna array can exhibit better performance.

1. INTRODUCTION

Recently, a lot of research works have shown that the capacity can be increased in multiples using MIMO systems. However, in some application of MIMO communication, the antennas should be closely spaced. Mutual coupling and correlation between antennas which are harmful to enhance diversity gain and capacity of MIMO system will be very high. To obtain high capacity of MIMO system is to design the channel with low correlation. Therefore, the design of antenna array plays a crucial role in enhancing antenna diversity and system capacity of MIMO system. Many methods have been developed to solve this problem e.g., MIMO cube. In this paper, we develop new kinds of antenna array, regular polyhedron antenna array and an antenna array constructed like the construction of carbon-60, which demonstrate high system capacity then planar antenna array. There are five kinds of regular polyhedrons: tetrahedron, hexahedron, octahedron, dodecahedron and icosahedron. We fixed the monopole at the center of each surface. We give the construction of regular polyhedron antenna array, and simulation examples using efficient integral method MGLFIM, and give the capacity of regular polyhedron array in MIMO systems. Finally, we give some conclusion.

2. POLYHEDRON ANTENNA ARRAY DESIGN AND ANALYSIS

In MIMO system, the channel model for 3D indoor environments can be expressed as Fig. 1 shown. As our recent research, the equivalent circuit of a MIMO system. We can obtain the channel matrix
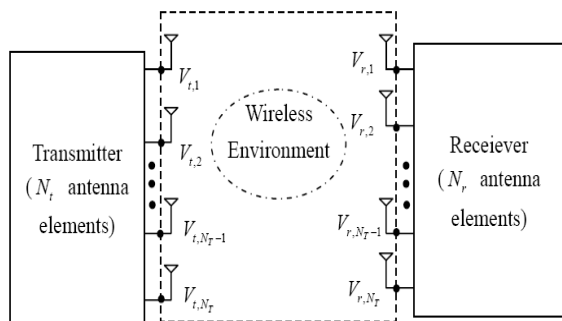

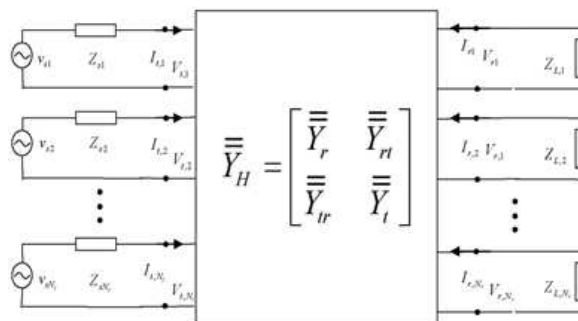Figure 1: The illustration of a generic MIMO system.

Figure 2: The equivalent circuit of a MIMO system.

as:

$$\bar{V}_r = -\left(\bar{\bar{Z}}_L^{-1} + \bar{Y}_r\right)^{-1} \cdot \bar{\bar{Y}}_{rt} \cdot \left(\bar{\bar{Z}}_s^{-1} + \bar{Y}_t\right)^{-1} \cdot \bar{\bar{Z}}_s^{-1} \cdot \bar{V}_s = \bar{\bar{H}} \cdot \bar{V}_s \tag{1}$$

$$\bar{\bar{H}} = -\left(\bar{\bar{Z}}_L^{-1} + \bar{Y}_r\right)^{-1} \cdot \bar{\bar{Y}}_{rt} \cdot \left(\bar{\bar{Z}}_s^{-1} + \bar{Y}_t\right)^{-1} \cdot \bar{\bar{Z}}_s^{-1} \tag{2}$$

The capacity of MIMO system can be expressed using "water-filling".

Therefore, it shown that transmit and receive admittance matrices plays a crucial role in MIMO system. In order to obtain high system capacity, we developed new kinds of antenna arrays as Fig. 3 shown. There five kinds of regular polyhedron, e.g., dodecahedron, hexahedron, octahedron, tetrahedron and icosahedron, so we can construct 4×4, 6×6, 8×8, 12×12, 20×20 MIMO systems. Regular polyhedron has excellent symmetry, which is vital in the indoor environment which has 3D uniform APS. In our simulation system, we use the same antenna array both at transmitter and receiver. Additionally, since the spatial construction of the antenna array, the correlation between monopoles should be lower than the planar antenna array has. Another kind of polyhedron antenna array we developed as Fig. 3(f) shown, is a 32 antenna array. It looks like the construction of carbon-60.

(a) tetrahedron   (b) hexahedron   (c) octahedron

(d) dodecahedron   (e) icosahedron   (f) carbon-60
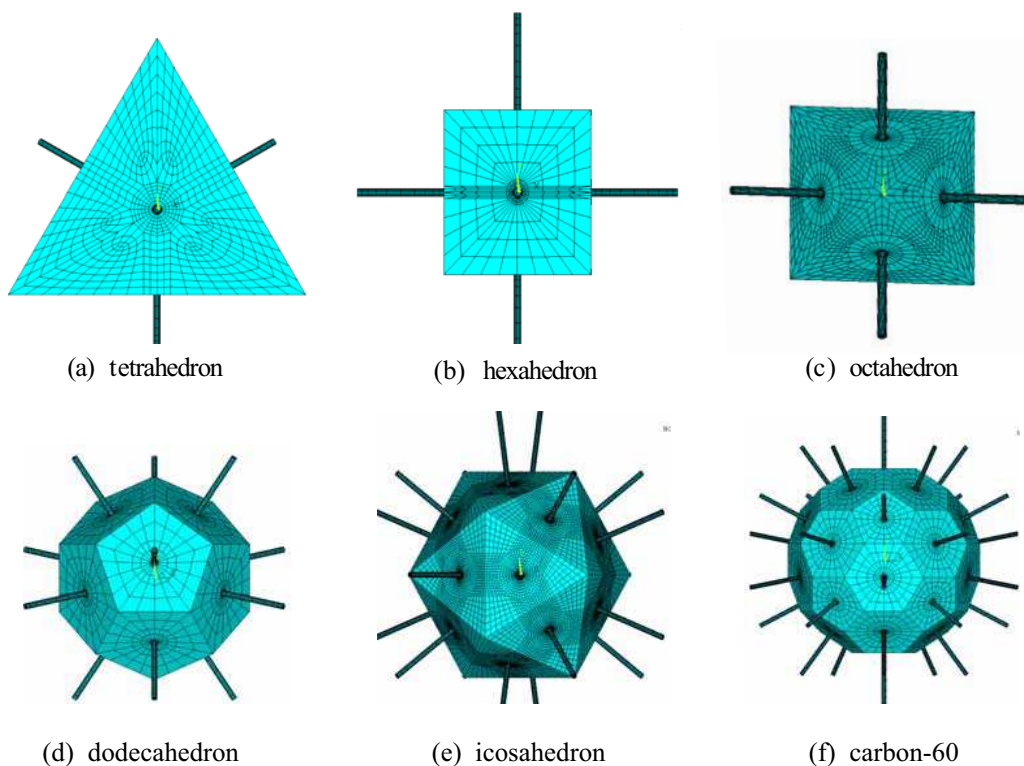
Figure 3: Regular polyhedron antenna array and carbon-60 antenna array.

3. NUMERICAL SIMULATION

In our numerical simulation, efficient integral method MGLFIM is employed to fast calculate the input admittance and radiation pattern. We set the work frequency of antenna array to be $1.95\,\text{GHz}$, the radius and height of each monopole are set to be $\lambda/100$ and $\lambda/4$. Fig. 4 shown the radiation gain patterns of five kinds of regular polyhedron antenna arrays, and the distance between nearest monopoles is set to be $\lambda/2$. Fig. 5 shows the mutual resistance and reactance of tetrahedron antenna array varying the antenna spacing from 0 to λ. Because of the tetrahedron's symmetry, the mutual impedances between any two monopoles are the same in Fig. 5, R1 and X1 are the self impedance's resistance and reactance, and R2 and X2 are mutual impedance's. Fig. 6(a) demonstrates the capacity of MIMO varying antenna spacing from 0 to λ. Fig. 6(b) shows the comparison of the system capacity between icosahedron antenna arrays and planar antenna array. The planar antenna array is set to be 4×5 monopoles antenna array, and the distance between each nearest monopoles is set to be $\lambda/2$.
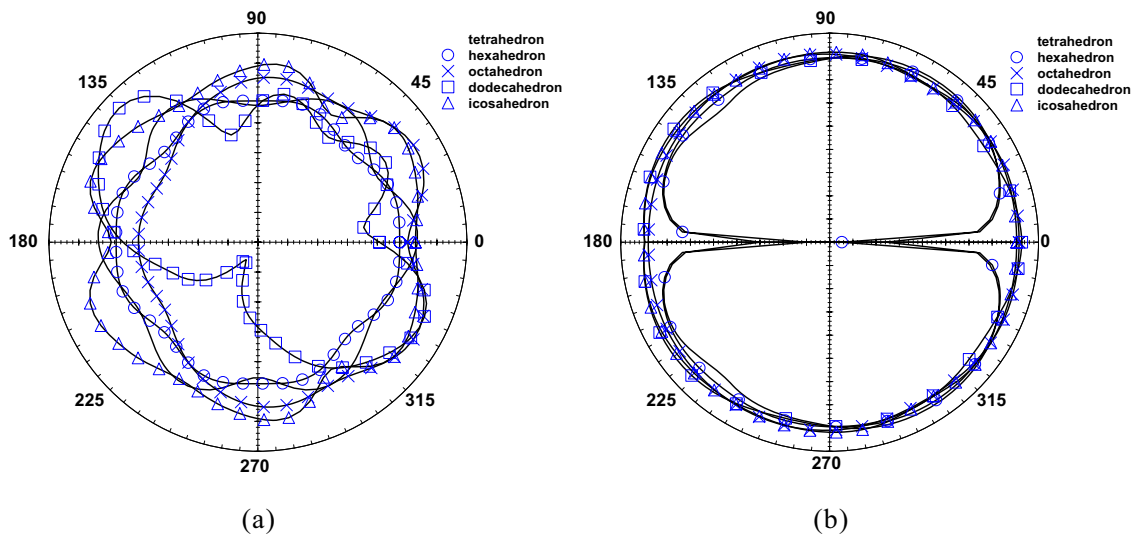
Figure 4: (a) H plane and (b) E plane pattern of the regular polyhedron array when the port of the monopole is excited by 1 volt voltage.
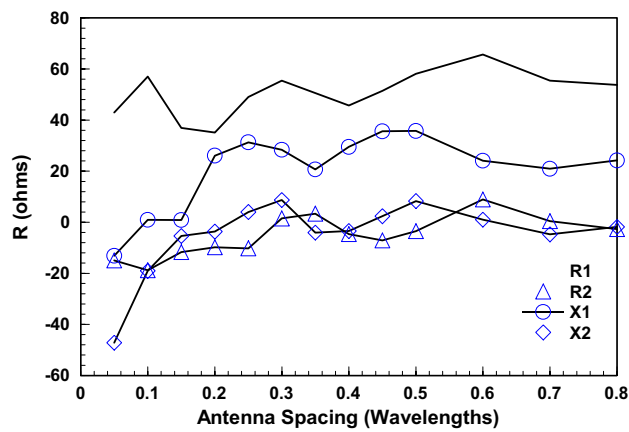

Figure 5: The mutual impedances of tetrahedron array (R) the mutual resistance and reactance (X) versus the spacing between two nearest monopoles.
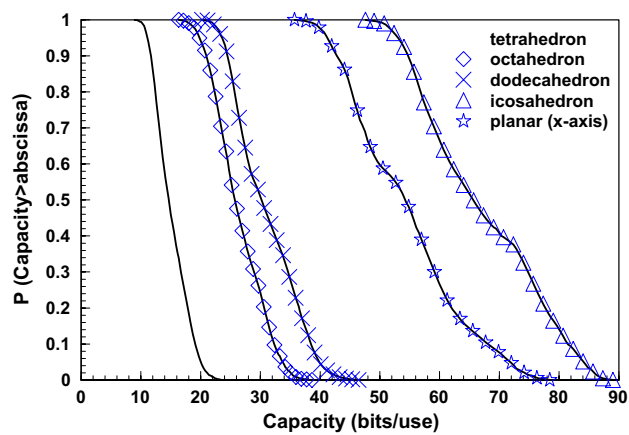

Figure 6: The capacity of five kinds of regular polyhedron antenna array (the antenna spacing between two nearest monopole is set to half wavelengths).

4. CONCLUSION

In this paper, we introduce a new kind of antenna array for MIMO system. Regular polyhedron antenna arrays can apparently enhance the performance of MIMO systems. In order to obtain higher system capacity, we can use the new kind of antenna array. Compare to planar antenna array, we can obtain higher capacity with less antennas and small physical volume of regular polyhedron antenna array. The fast MoM method MGLFIM we employed is efficient to calculate the input admittances and radiation patterns of this kind of very complex antenna.

ACKNOWLEDGMENT

REFERENCES

1. Khaleghi, A., "Diversity techniques with parallel dipole antennas: Radiation pattern analysis," *Progress In Electromagnetics Research*, PIER 64, 23–42, 2006.
2. Getu, B. N. and J. Bach Andersen, "The MIMO cube — A compact MIMO antenna," *IEEE Trans. on Wireless Comm.*, Vol. 4, No. 3, 1136–1141, May 2005.
3. Wang, H. G., "A novel numerical model for simulating three dimensional MIMO channels with complex antenna arrays," unpublished.
4. Wang, H. G. and C. H. Chan, "The implementation of multilevel Green's function interpolation method for full-wave electromagnetic problems," *IEEE Trans. Antennas Propag.*, Vol. 55, No. 5, May 2007.

A Source Localization Scheme Based on Unitary ESPRIT and the City Electronic Map

Hong Bing Song, **Hao Gang Wang**, **Li Wang**, and **Da Qing Liu**

The Electromagnetics Academy at Zhejiang University, Zhejiang University
Hangzhou 310027, China

Abstract— In this paper, a novel source localization scheme is proposed based on the unitary ESPRIT algorithm with back ray tracing technique and the city electronic maps. In this scheme, the unitary ESPRIT algorithm is employed to estimate the angles and delays of the arrival rays radiated from the source. Based on the obtained information, we devise a back ray tracing techniques according to the Geometrical Theory of Reflections and the 2D city electronic map. Error analyses have been made and show that the localization accuracy for a cell of about 80 m×40 m can reach 1 m when the SNR is greater than 15 dB.

1. INTRODUCTION

In recent years, interest in source localization services has grown rapidly. There are three typical methods for locating a source, viz., time-of-arrival (TOA), angle-of-arrival (AOA) approach [1] and hybrid technique [2]. The TOA and AOA need more than one base station and the estimation accuracy is low for non-line of sight (NLOS) source. In [2] a hybrid TOA/AOA positioning algorithm was proposed. Four different propagation environments, i.e., bad urban, urban, suburban, and rural, were conducted in [2]. However, it doesn't consider the real propagation path. In this paper a new hybrid localization method is proposed. Based on the city electronic map and high resolution signal parameter estimation algorithm-Unitary ESPRIT (UESPRIT) [3, 4] the performance of position system can be improved.

UESPRIT is a simple and highly efficient method for estimating signal parameters based on the translational invariance geometry of an array. In recent years, it has been extended to the solution of multidimensional cases, where it is used to estimate the direction of arrival (DoA) and time delay of arrival (TDoA). With the estimated DoA and TDoA, the arrival angles and the lengths of the paths of signals can be obtained. According to these angles, the Back Ray Tracing (BRT) emits signals from the receiver. Then the transmiting waves will move forward, hit the building, and be reflected. They will not stop until lengths of the path reach the estimated ones. The stop position is the source location. For the case the signals hit around the corner of the building, the estimated position may far away from the actual position, when the estimated DoA and TDoA deviate a little from the real DoA and TDoA. In order to minimize the estimated error, a selection scheme is employed in this paper. After eliminating the wrong estimated position, the average of the estimated position will be made, which is the ultimately estimated source location. The localization accuracy can reach 1 m when SNR is greater than 15 dB.

The organization of this paper is as follows. In the next section, the famous UESPRIT algorithm is reviewed. The BRT algorithm is presented in Section 3. Then, in Section 4, the error analyses are discussed. The numerical results are shown in Section 5.

2. A BRIEF DESCRIPTION OF UESPRIT

The receive antenna is a uniform linear array (ULA), as shown in Fig. 1. Assuming that the transmitter emits a sounding signal $e^{j\omega t}$ and the impinging wave fronts on the receiver are approximately planar. The received signal at each sensor of ULA reads:

$$y_{m,\,l}(t) = \sum_{i=1}^{N} u_i(t) e^{j2\pi(f_0 + m\Delta f)(t - \tau_i - \frac{dl\sin\theta_i}{c})} + n(t), \tag{1}$$

where m (from 0 to $M-1$), l (from 0 to $L-1$) denote indices of frequency points and antenna sensors of the ULA at Rx respectively, N denotes the number of propagation paths, f_0: center frequency, Δf: frequency gap, d: sensors interval, θ_i: the i-th path arrival angle, τ_i: the i-th path delay, $u_i(t)$: amplitude varying function of time, and $n(t)$: the noise. In the simulation $M \cdot \Delta f \cdot d \cdot L/c < 0.01$,
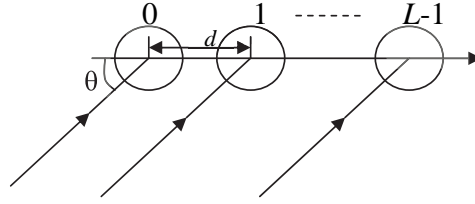
Figure 1: Uniform linear array (ULA).

so (1) can be simplified to [5]

$$y_{m,l}(t) = \sum_{i=1}^{N} u_i(t) e^{j2\pi f_0(t-\tau_i)} e^{-j2\pi m \Delta f \tau_i} e^{-j2\pi f_0 \frac{dl \sin \theta_i}{c}} + n(t). \tag{2}$$

(2) can be solved by 2-D UESPRIT. The smoothing technique will be used in the simulation, and the data will be reconstructed into M_s sMoverlapped snapshots (or sub arrays), each containing $M_{\text{sub}} = M - M_s + 1$ measurements. This set of $M_s L$ snapshots is used to estimate the parameters $\tilde{\tau}_N = [\tilde{\tau}_1 \dots \tilde{\tau}_N]^T$ and $\tilde{\theta}_N = [\tilde{\theta}_1 \dots \tilde{\theta}_N]^T$ with 2-D UESPRIT.

One of the most important things is to estimate the number of the paths. Here, a modified MDL [6, 7] is used.

3. BACK TRACING ALGORITHM FOR SOURCE LOCALIZATION

In Section 2 the parameters sets $\tilde{\tau}_N = [\tilde{\tau}_1 \dots \tilde{\tau}_N]^T$ and $\tilde{\theta}_N = [\tilde{\theta}_1 \dots \tilde{\theta}_N]^T$ have been derived by UESPRIT. With the estimated parameter DoA and TDoA, the arrival angles and the lengths of path which the signals have been traveled can be obtained. According to these angles, the Back Ray Tracing (BRT) emits signals from the receiver. Then the propagating waves will move forward, hit the building, and be reflected. They will not stop until the lengths of the path reach the estimated ones. The stop position is the source location.

4. ERROR ANALYSES OF THE ALGORITHM

According to the above discussion, the errors consist of two parts: the first one is caused by the finite accuracy of UESPRIT, e.g., the DoA and TDoA errors, while the second one is the terrain data error compared with the real scene.

If the signal paths hit around the corner of the building, as shown in Fig. 2, the first part of error will be extremely significant. In these situations (Fig. 2), the estimated position got from RT will be far away from the real source position. The terrain data error can also cause this problem. In order to eliminate this error, the estimated positions must be selected. Here we propose a selection scheme, using Eq. (3)

$$\begin{cases} \Delta = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} & i \neq j \quad i, j = 1 \dots N \\ \Delta < \Delta r \end{cases}, \tag{3}$$

where (x_i, y_i) and (x_j, y_j) are the estimated localizations, N denotes the number of path considered, and Δr denotes threshold used to determine whether the result is reliable. Consequently, the source localization is derived from the average of all the estimated positions that satisfy Eq. (3). Through out this paperr Δr is chosen as 2 m.
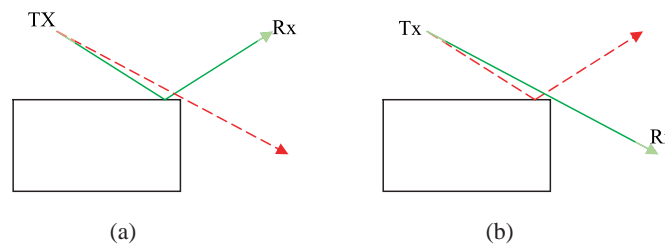

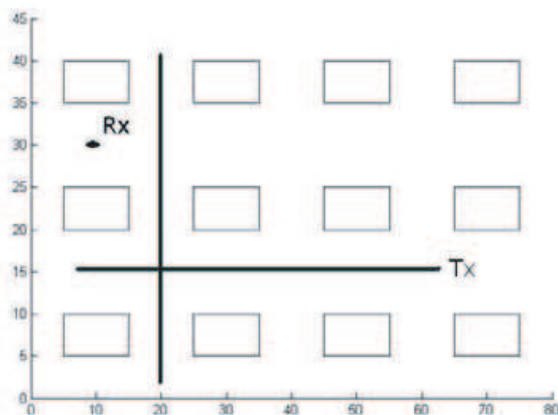Figure 2: Error model (solid line: real path, dash line: estimated path).

Figure 3: Simulation model.

5. NUMERICAL RESULTS

Two simulation examples are carried out in this paper. In the two examples the parameters are set to be $f_0 = 1\,\text{GHz}$, $\Delta f = 1 \times 10^5\,\text{Hz}$, $M = 30$, $L = 10$, and $d = 0.5\,\lambda$. The terrain data is shown in Fig. 3.

Simulation 1. The receiver was fixed in one place, and the transmitter moves along the solid line with a step of $0.5\,\text{m}$, as shown in Fig. 3. A total of 140 positions are sampled. In each transmit location the position estimation is carried out in all values of SNR, viz., SNR from 0 to 45 dB. The relation between the average estimation error and SNR was derived in Fig. 4. When SNR is greater than 15 dB, the average error $\bar{\varepsilon}$ is almost less than 1 m. the definition of $\bar{\varepsilon}$ is

$$\bar{\varepsilon} = \frac{\sum\limits_{i=1}^{n} \sqrt{(\hat{x}_i - x_i)^2 + (\hat{y}_i - y_i)^2}}{n},\tag{4}$$

where $(\hat{x}_i, \hat{y}_i)$ and (x_i, y_i) denote the estimated position and real position respectively. n denotes the number of the position.

Simulation 2. In this simulation the errors of the length and width of the building are assumed to be Gaussian distributed with zero-mean and variance $\sigma = 0.5\,\text{m}$. The algorithm is tested at three positions: $(70, 30)$, $(15, 15)$, and $(30, 2)$. The estimations are: $(69.996, 29.987)$, $(15.002, 15.017)$, and $(40.687, 2.004)$. In Fig. 5 the rectangle, circle, and cross denote the estimated positions with terrain error, estimated positions without terrain error, and actual location respectively.
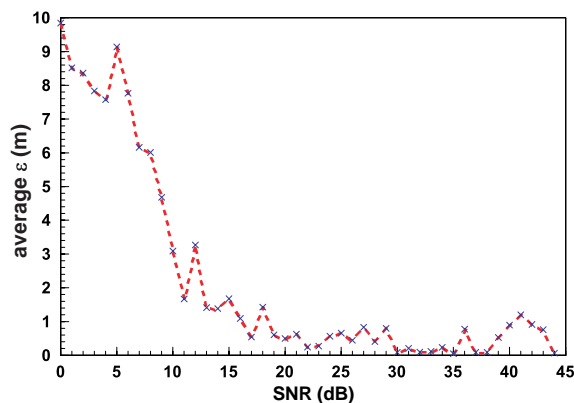
Figure 4: The estimated error versus SNR.
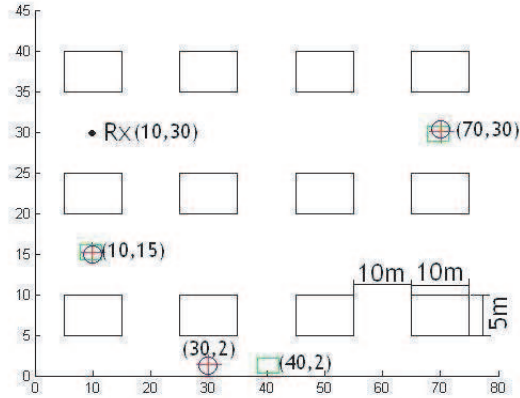
Figure 5: The position estimation under terrain errors.

6. CONCLUSION

In this paper, a source localization scheme is proposed based on the UESPRIT with BRT and the city electronic maps. In this scheme, the UESPRIT is employed to estimate the arrival angles and delays of the rays radiated from the source. Based on the obtained information, the source position can be obtained according to Ray Tracing program and the 2D city electronic map. The estimation error is analyzed, and two simulation examples have been carried out. According to the numerical results, the localization accuracy can reach 1 m when SNR is greater than 15 dB.

ACKNOWLEDGMENT

REFERENCES

1. Caffery, J. J., *Wireless Location in CDMA Cellular Radio Systems*, KAP, 1999.
2. Deng, P. and P. Z. Fan, "An AOA assisted TOA position system," *ICCT2000 Proceedings [C].*, 1501–1504, Beijing, August 2000.
3. Haardt, M. and J. A. Nossek, "Unitary ESPRIT: How to obtain increased estimation accuracy with a reduced computational burden," *IEEE Trans. Signal Processing*, Vol. 43, 1232–1242, May 1995.
4. Haardt, M. and J. A. Nossek, "Simultaneous schur decomposition of several nonsymmetric matrices to achieve automatic pairing in multidimensional harmonic retrieval problems," *IEEE Trans. Signal Processing*, Vol. 44, 161–169, January 1998.
5. Zwick, T., D. Hampicke, A. Richter, G. Sommerkorn, R. Thoma, and W. Wisebeck, "A novel antenna concept for double-directional channel measurements," *IEEE Trans. on Vehicular Technology*, Vol. 53, 527–537, March 2004.
6. Wax, M. and T. Kailath, "Detection of signal by information theoretic criteria," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, Vol. 33, 387–392, April 1985.
7. Xu, G., R. H. Roy, and T. Kailath, "Detection of number of sources via exploitation of centro-symmetry property," *IEEE Trans. Signal*, Vol. 42, 102–112, January 1994.

Performance Analysis of OFDM Communication System over Correlated Nakagami-m Fading Channel

Vivek K. Dwivedi[1], **Pradeep Kumar**[2], **and G. Singh**[2]

[1]Department of Electronics and Communication Engineering
Jaypee Institute of Information Technology University, Noida 201307, India
[2]Department of Electronics and Communication Engineering
Jaypee University of Information Technology, Solan 173215, India

Abstract— In this paper, we have analyzed the performance of OFDM communication system over correlated Nakagami-m fading channel using maximal ratio combining (MRC) diversity at the receiver. We have presented new closed form formulas for exact symbol error rate (SER) analysis of the MQAM OFDM communication systems over correlated Nakagami-m fading with arbitrary fading index m. Using a well known moment generating function (MGF) based analysis, we express the average SER in terms of higher transcendental function such as Appell hypergeometric function.

1. INTRODUCTION

The demand of wireless communication is growing exponentially and next generation of wireless broadband multimedia communication systems will integrate various function and application in same system which supports large data rates with sufficient robustness to radio channel impairments, requires careful choosing of modulation technique. The suitable choice is orthogonal frequency division multiplexing (OFDM) which is special case of multi-carrier communication system, where single data stream is transmitted over number of lower sub-carrier [1, 2]. The basic principle of OFDM communication systems is to split the high data stream into number of lower rate data streams which are transmitted simultaneously over a number of sub-carriers. High spectral efficiency and multipath immunity are two major features of OFDM systems. OFDM has been shown to be an effective technique to combat frequency selective multipath fading. Recently, the Nakagami-m fading channel model has received considerable attention due to its great versatility, in the sense that it has greater flexibility and accuracy in matching various experimental data more general than Rayleigh, log-normal or Rician distributions. Seaglione [3] claimed that the magnitude of channel frequency response can also be modeled as Nakagami-m random variable with same fading parameters m since the channel frequency response at each frequency bin in a linear combination of the channel multipath whose magnitude are Nakagami-m random variables. Nakagami fading channel model [4–8] has received considerable attention due to its great flexibility and accuracy. To analyze the performance, it is necessary to know the statistics of the channel in the frequency domain.

By using diversity techniques the performance of OFDM systems can be improved dramatically. Thus the OFDM systems with diversity have very significant role in wireless communication [7]. So it is necessary to investigate the performance over different fading channels. To analyze the performance of OFDM systems on Nakagami-m fading channels, it is necessary to know the statistics of the channels in the frequency domain. Previous work presented in [3, 5, 6] has assumed that the frequency domain channel response samples are also Nakagami-m distributed with the same fading parameters as the time domain channel which is not valid as shown in [7]. Kang et al. [7] shows that the magnitudes of frequency responses can be well approximated by Nakagami-m random variables with new parameters but it consider only two diversity path at the receiver.

In this paper, we have presented a novel method for symbol error rate (SER) for correlated Nakagami fading channel using maximal-ratio combining (MRC) diversity and we are considering $(D \geq 2)$ diversity path at the receiver and for $D = 2$ moment generating function (MGF) of presented method is similar with MGF in [7].

2. SYSTEM MODEL

Let $S_i(k)$ be the kth OFDM data block to be transmitted with N sub-carriers. These data are used to modulate N orthogonal sub carriers. After that the inverse discrete Fourier transform (IDFT)

is used to modulate the input signal. After modulation signal can be represented as [1]

$$x_i(n) = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} S_i(k)e^{j2\pi k/N} \qquad n = 0, 1, 2 \ldots N - 1$$

Cyclic prefix is inserted after IDFT modulation, which is removed before demodulation at the OFDM receiver. The resultant signal is up converted to RF before transmission and at the receiver end down converted. The received signal after removal of cyclic prefix can be demodulated using DFT. Output of DFT can be represented as

$$R_i(k) = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} r_i(n)e^{-j2\pi k/N} = H_i(k)S_i(k) + W_k \qquad k = 0, 1, \ldots N - 1$$

where W_k is additive complex Gaussian noise and N-point DFT of the channel impulse response at the nth frequency bin is defined as $H(k)$ in frequency domain channel impulse response.

$$H_i(k) = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} h_i(n)e^{-j2\pi k/N} = \frac{1}{\sqrt{N}} \left(e^H h_i(n)\right) \qquad (0 \le n \le N - 1)$$

In the above equation $h_i(n)$ is Nakagami-m distributed random variable.

$$e = (1, \exp(-j2\pi n/N), \ldots, \exp(-j2\pi n(N-1)/N))^T$$

where T is transpose of matrix.

3. BER ANLYSIS OF OFDM SYSTEM WITH MRC DIVERSITY

When receiving antennas are closely spaced then receiving signal are also correlated then SNR $\gamma_1, \gamma_2, \ldots, \gamma_M$ cannot be considered as independent random variable. Correlation coefficient between two receiving antenna is (assuming equal correlation between antennas ρ) [7]:

$$\rho = E\left(h_i(k)h_j(k)\right) \Big/ \sqrt{E\left(h_i(k)h_i^*(k)\right) E\left(h_j(k)h_j^*(k)\right)}.$$

Assuming that the average power signal as well as fading parameter in each channels of a maximum ratio combiner system are identical. The assumption of identical power is reasonable if the diversity channels are closely spaced and the gain of each channel is such that all noise power are equal [9]. The SNR at the output of MRC is given by [10].

$$\gamma_t = \frac{E_S}{\sigma^2} \sum_{i=1}^{D} |H_i|^2 = \frac{E_S}{\sigma^2} \sum_{i=1}^{D} (H_i)(H_i)^H = \sum_{i=1}^{D} \gamma_i$$

E_S is symbol energy and σ^2 is the variance of the zero mean complex mean Gaussian noise. Moment generating function (MGF) generating function over correlated Nakagami fading channel is given in [11].

$$\psi(s) = \left(1 - \frac{\bar{\gamma}_t(1 - \rho + D\rho)s}{m_f}\right)^{-m_f} \left(1 - \frac{\bar{\gamma}_t(1 - \rho)s}{m_f}\right)^{-m_f(D-1)} \tag{1}$$

where m_f is fading parameter as defined in [7].

For diversity path $D = 2$

$$\psi(s) = \left(1 - \frac{\bar{\gamma}_t(1 + \rho)s}{m_f}\right)^{-m_f} \left(1 - \frac{\bar{\gamma}_t(1 - \rho)s}{m_f}\right)^{-m_f}$$

Above equation is similar with Equation (9) in [7]. Average SER for coherent square MQAM signals is given by [8]:

$$\bar{P}_E = \frac{4q}{\pi} \int_0^{\pi/2} \psi\left(\underbrace{-\frac{g}{\sin^2\theta}}_{I_1}\right) d\theta - \frac{4q^2}{\pi} \int_0^{\pi/4} \psi\left(\underbrace{-\frac{g}{\sin^2\theta}}_{I_2}\right) d\theta \tag{2}$$

where $q = 1 - 1/\sqrt{M}$ and $g = 3/(2(M-1))$. Integral I_1 and I_2 are solved separately, by putting $s = -g/\sin^2\theta$ in Equation (1). Integrals I_1 and I_2 can be written as:

$$I_1 = \frac{1}{\pi}\int_0^{\frac{\pi}{2}} \psi\left(-\frac{g}{\sin^2\theta}\right) d\theta \quad \text{and} \quad I_2 = \frac{1}{\pi}\int_0^{\frac{\pi}{4}} \psi\left(-\frac{g}{\sin^2\theta}\right) d\theta$$

After some mathmatical manupulation and from [12], above integral can be expressed as:

$$I_1 = \frac{\psi(-g)}{2\pi}\frac{\Gamma(1/2)\Gamma(m_f D + 1/2)}{\Gamma(m_f D + 1)} F_1\left(\frac{1}{2}, m_f, m_f(D-1), m_f D + 1, \frac{1}{1+A}, \frac{1}{1+B}\right)$$

$$I_2 = \frac{1}{\pi}(A)^{-m_f}(B)^{-m_f(D-1)}2^{-(m_f D + 1/2)} F_1\left(m_f D + 1/2, m_f, m_f(D-1), m_f D + 1, -\frac{1}{2A}, -\frac{1}{2B}\right)$$

where $A = \frac{\bar{\gamma}_t(1-\rho+D\rho)}{m_f}$, $B = \frac{\bar{\gamma}_t(1-\rho)g}{m_f}$, $\bar{\gamma}_t$ = Signal to noise ratio, $F_1()$ Appell hyper geometric function given in [13], D = Number diversity path of MRC at receiver. By putting I_1 and I_2 in Equation (2) resultant average SER becomes:

$$\bar{P}_E = 2q\frac{\psi(g)}{\pi}\frac{\Gamma(1/2)\Gamma(m_f D + 1/2)}{\Gamma(m_f D + 1)} F_1\left(\frac{1}{2}, m_f, m_f(D-1), m_f D + 1, \frac{1}{1+A}, \frac{1}{1+B}\right)$$
$$- \frac{4q^2}{\pi}(A)^{-m_f}(B)^{-m_f(D-1)}2^{-(m_f D + 1/2)} F_1\left(m_f D + 1/2, m_f, m_f(D-1), m_f D + 1, -\frac{1}{2A}, -\frac{1}{2B}\right)$$

Total average SER of of multiple received antenna OFDM system can be expressed as:

$$P_{\text{etotal}} = 1 - \left(1 - \bar{P}_E\right)^N$$

4. RESULT

The Nakagami-m distribution has been employed as an important model for characterizing the amplitude of fading channels. Figure 1 shows the SER probability of MRC over correlated Nakagami fading channel with 4-QAM modulation for different correlation parameter. At $\rho = 1$ SER approaches 10^{-1} for 20 dB SNR and $\rho = 0$ SER approaches 10^{-3} for 20 dB SNR. As correlation increases SER performance of system is decreased. Figure 2 shows SER probability of MRC over correlated Nakagami fading channel with 4-QAM modulation for different number combined diversity path (D), Nakagami fading parameter for $m = 4$ and correlation coefficient $\rho = 0.5$ for analysis. SER performance is improved as no diversity path is increased.
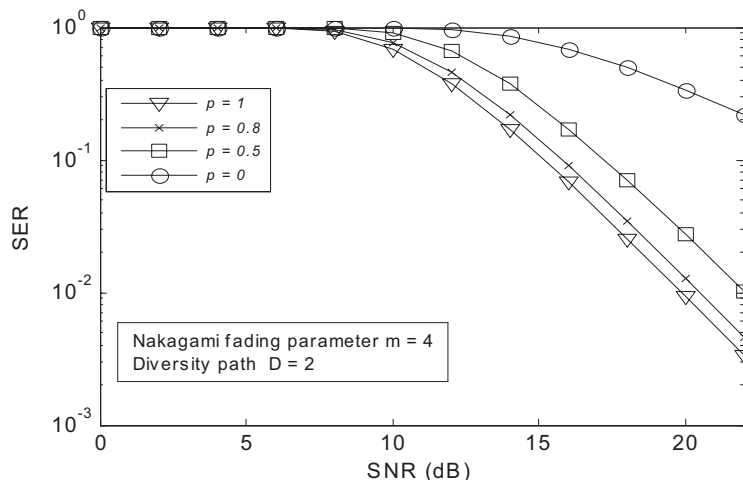

Figure 1: SER of MRC over correlated Nakagami-m fading with QAM OFDM system and different correlation coefficients.
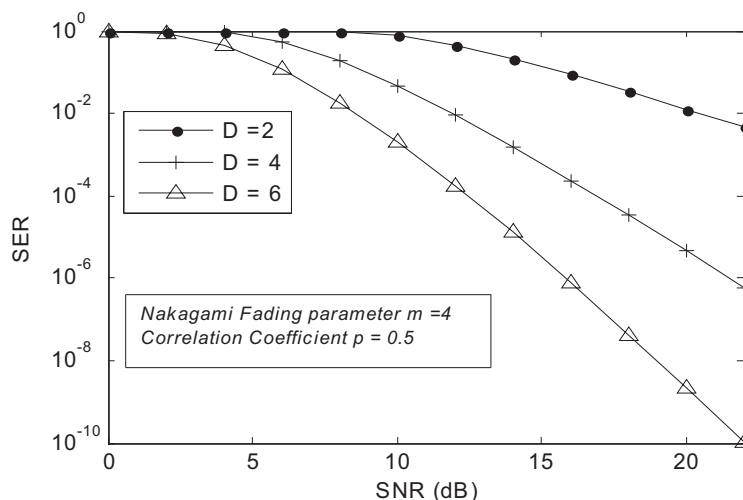
Figure 2: SER of MRC over correlated Nakagami-m fading with QAM OFDM system and different numbers of combined paths.

5. CONCLUSION

In this paper, we have investigated a novel method for SER performance analysis of OFDM system over correlated Nakagami-m fading channel using MRC diversity at the receiver. In proposed method we are considering ($D \geq 2$) no antenna at receiver so BER performance of OFDM system is improved greatly.

REFERENCES

1. Moose, P. H., "A technique for orthogonal frequency division multiplexing frequency offset correction," *IEEE Trans. Comm.*, Vol. 42, 2908–2914, Oct. 1994.
2. Dwivedi, V. K. and G. Singh, "An efficient BER analysis of OFDM systems with ICI conjugate cancellation method," *PIERS Proceedings*, 166–171, Cambridge, USA, Jul. 2–6, 2008.
3. Scaglione, A., S. Barbarossa, and G. B. Giannakis, "Optimal adaptive precoding for frequency-selective Nakagami-m fading channels," *Proc. 52nd IEEE Vehicular Technology Conf.*, Vol. 3, 1291–1295, 2000.
4. Nakagami, M., "Them-distribution — A general formula of intensity distribution of rapid fading," *Statistical Methods in Radio Wave Propagation*, W. G. Hoffman, Ed., Pergamon, Oxford, UK, 1960.
5. Gong, Y. and K. B. Letaief, "Performance of space-time Trellis coding over Nakagami fading channels," *Proc. 51st IEEE Vehicular Technology Conf.*, Vol. 2, 1405–1409, 2001.
6. Aalo, V. A. and J. Zhang, "Average error probability of optimum combining with a co-channel interferer in Nakagami fading," *Proc. Wireless Communications and Networking Conf.*, Vol. 1, 376–381, 2000.
7. Kang, Z., K. Yao, and F. Lorenzelli, "Nakagami-m fading modeling in the frequency domain for OFDM system analysis," *IEEE Comm. Lett.*, Vol. 7, 484–486, Oct. 2003.
8. Annamalai, A. and C. Tellambura, "Error rates for Nakagami-m fading multichannel reception of binary and M-ary signals," *IEEE Trans. on Comm.*, Vol. 49, 58–68, Jan. 2001.
9. Pierce, J. N. and S. Stein, "Multiple diversity with non independent fading," *Proc. IRE*, Vol. 48, 89–104, Jan. 1960.
10. Brennan, D. G., "Linear diversity combining techniques," *Proc. IRE*, Vol. 47, 1075–1102, Jun. 1959.
11. Dwivedi, V. K. and G. Singh, "Improved BER analysis of OFDM communication system on correlated Nakagami fading channel," *Proc. Int. Conf. on Recent Advances in Microwave Theory and Applications (Microwave-2008)*, 536–539, Jaipur, India, Nov. 21–24, 2008.
12. Gradshteyn, I. S. and I. M. Ryzhik, *Table of Integrals, Series, and Products*, 6th Ed., Academic Press, San Diego, 2000.
13. Erdelyi, A., *Higher Transcendental Function*, McGraw Hill, New York, 1953.

Technical Equipment for Research of EM Field and Biological Systems Interactions

Jan Vrba[1], **Luca Vannucci**[2], **Paolo Togni**[1], **and Lukas Visek**[1]

[1]Dept. of EM Field, Czech Technical University in Prague, Prague, Czech Republic
[2]Institute of Microbiology, Czech Academy of Sciences, Prague, Czech Republic

Abstract— Research of interactions between EM Field and biological systems is of growing interests elsewhere. Also here in Czech Republic there are several groups working in this field, often in international co-operations. We will describe here mainly basic technical equipment developed for 5 different research projects in the discussed area of interactions of EM field and biological systems.

1. INTRODUCTION

In present time four research institutions here in the Czech Republic run research projects focused on studies of interactions between EM field and biological systems. These institutions are technically supported by Department of EM Field of the Czech Technical University in Prague. In this contribution we would like to give more details about that projects and obtained technical results (i.e., description of developed exposition systems).

Three of discussed projects (1 in Germany and 2 here in Czech Republic) are basic research for simulation of the microwave hyperthermia treatment. Other two projects (both in Czech Republic are focused on simulation of the case of exposition by mobile phone.

In the modern view, cancer is intended as a complex illness, involving the cells that undergo to transformation, their environment, and the general responses at biochemical and biological levels induced in the host. Consequently, the anti-cancer treatment protocols need to be multi-modal to reach curative effects. Especially after the technical improvements achieved in the last 15 years by bio-medical engineering, microscopy devices, and molecular biology methods, the combinations of therapeutic procedures are growing in interest in basic and clinical research.

The combination of applied biological research together to the physical sciences can offer important perspectives in anticancer therapy (e.g., different methodologies and technical devices for application of energies to pathological tissues).

The modern bioengineering knowledge applied to traditional tools, as the microscopy, has largely renewed and expanded the fields of their applications (e.g., *in vivo* imaging), pushing the interest for direct morpho-functional investigations of the biomedical problems.

2. WAVEGUIDE APPLICATOR

Very good results of EM field expositions in biological experiments can be obtained by simple but efficient waveguide applicators, see example in Fig. 1.

Waveguide offer a very big advantage — in approximately of fifty percents of its aperture the irradiated electromagnetic field is very near to a plane wave, which is basic assumption for good homogeneity of the heating and optimal treatment penetration.

Here described system is being used (shared) for research projects by two two institutions (Institute of Radiation Oncology in Prague and Institute of Microbiology of the Czech Academy of Sciences).

Aperture of this waveguide is 4.8×2.4 cm and it is excited at frequency 2.45 GHz. Effective heating is in the middle of the real aperture — its size is approximately 2.4×2.4 cm. Waveguide is filled by teflon to reduce its cut-off frequency. Power from generator is possible to control from 10 to 180 W, in these experiments we work between 10 and 20 W mostly.

3. EVALUATION OF WAVEGUIDE APPLICATOR

To evaluate this applicator from technical point of view we made a series of experiments, see e.g., Fig. 2, where you can see example of measurement of temperature distribution by IR camera.

Here you can see temperature distribution obtained on surface of a model of mouse made from agar — with a simulated tumour on mouse back. Experiment has been done by heating phantom during 2 minutes delivering a power of 10 W. Maximum of temperature increase has been found

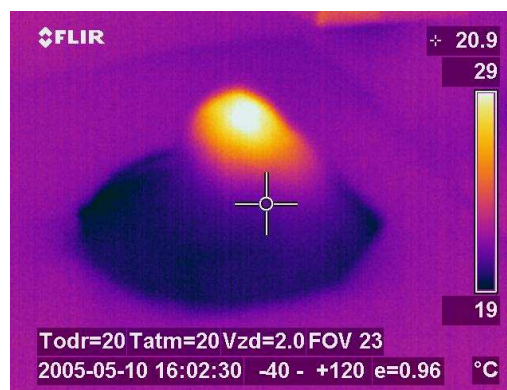Figure 1: Waveguide applicator for biological experiments.

Figure 2: Temperature distribution obtained on surface of a model of mouse.

approximately 10°C. Similar results with different increase in temperature we have got also in other technical experiments on phantom or live mouse when power or heating time was changed.

Next Fig. 3 gives example of temperature increase on the surface of the phantom and 1 cm in that phantom. Temperature was measured by our 4-channel thermometer. In this case with two thermoprobes.

Heating here is scheduled to 9 times repeated 30 s of heating and 30 s pause. Difference in temperature on the surface and under it is on the level of 1°C. That means very good homogeneity of temperature distribution in the treated area during planned biological experiments.
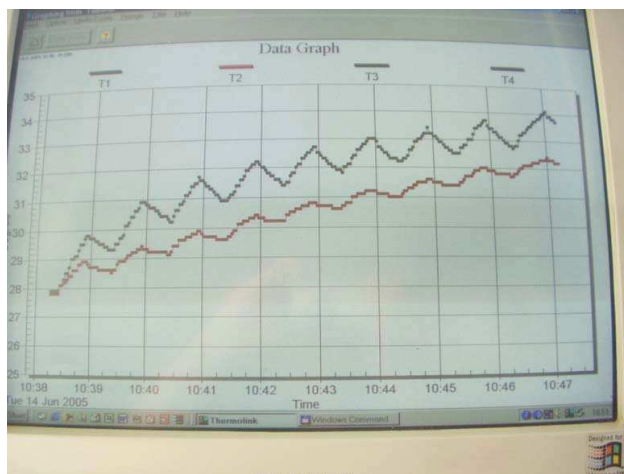
Figure 3: Temperatures during experiments.

4. ARRAY APPLICATOR

The main goal of the planned biological experiment is a hyperthermia treatment of the experimentally induced pedicle tumours of the rat to verify the feasibility of ultrasound diagnostics and magnetic resonance imaging respectively to map the temperature distribution in the target area of the treatment. That means to heat effective volume of approximately cylindrical shape (diameter approx. 2 cm, height approx. 3 cm). Temperature to be reached is 41°C or more (i.e., temperature increase of at least 4°C from starting point 37°C), time period of heating is 45 minutes.

Considering the necessary effective heating depth for the planned experiments, we have found 915 MHz to be suitable frequency. As an excellent compatibility of the applicator with non-invasive temperature measurement system (ultrasound or NMR) is a fundamental condition for our project, we should have to use non-magnetic metallic sheets of minimised dimensions to create the conductive elements of the applicator. Therefore the applicator itself (see Fig. 4) is created by two inductive loops tuned to resonance by capacitive elements [4, 5]. Dimensions of these resonant loops were designed by our software, developed for this purpose. Coupling between coaxial feeder and

resonant loops (not shown in Fig. 4) as well as a mutual coupling between resonating loops could be adjusted to optimum by microwave network analyser.
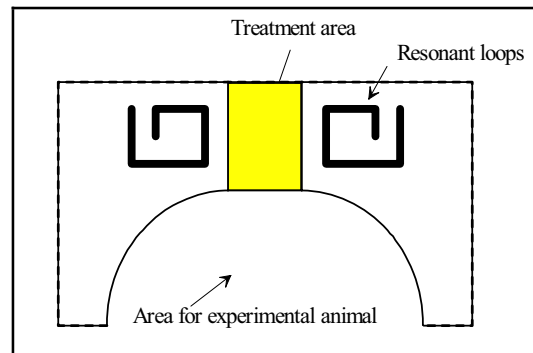

Figure 4: Arrangement of discussed microwave hyperthermia applicator.

The position of the loops is fixed by perspex holder. There is a special cylindrical space for experimental animal in lower part of this perspex holder. As the heated tissue has a high dielectric losses, both loops are very well separated and so no significant resonance in heated area can occur. From this follows, that either the position of the loops with respect to heated area or the distance between the loops is not very critical.

First measurements to evaluate the basic properties of the discussed applicator were done on agar phantom of muscle tissue:

- evaluation of basic microwave properties (transfer of EM energy to the tissue, reflections),
- evaluation of compatibility with US and NMR,
- calculation and measurement of SAR and temperature distribution and its homogeneity.

Exact tuning of the resonant loops to frequency 915 MHz has been easy and we could optimise the coupling between the coaxial feeder and resonant loops as well, reflection coefficient less than 0.1. We have tested the power to be delivered to the applicator to obtain sufficient temperature increase (approximately 4°C in less than 5 minutes is required). With power 10 W delivered to each loop for period of 2 minutes we succeeded to obtain the temperature increase of approximately 7°C. To keep the increased temperature for a long time, 2 W in each loops were sufficient. Similar values were obtained during first experiments on rats also. Even with higher level of delivered microwave power we did not observe the change of resonant frequency (caused by increased temperature of the loops).

This applicator has been developed for German Cancer Research Institute in Heidelberg. And it is being used there for a series of animal experiments to study effect of hyperthermia on tumours and possibility to combine hyperthermia with chemotherapy etc.

Compatibility of this applicator with a Magnetic resonance unit (MR) has been studied and it has been demonstrated.
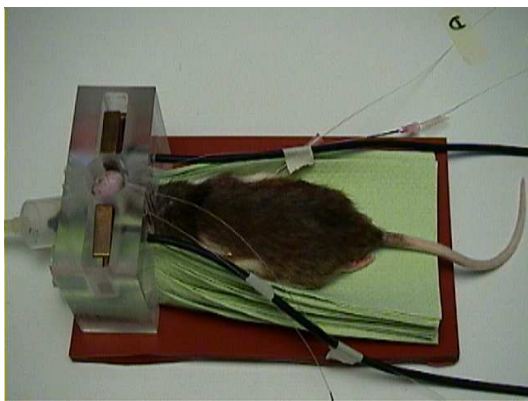

Figure 5: Photograph of the discussed applicator.

Figure 6: MR images of the discussed applicator.

We have tested the influence of the applicators on US diagnostics and NMR imaging and the result of this evaluation shows very good compatibility. Only a negligible deterioration of the US images have been observed when the incident power was kept under 100 W.

Details about influence of microwave power on MR imaging are given in Fig. 6. We can see here a sequence of images of the discussed applicator made by MR unit for four different cases. First case (upper left) is image for the case without power excitation of the applicator. Second case (left down) a power of 10 W has been delivered to each loop. We can see quite clear configuration of the applicator set-up. Third case (upper right) gives situation when 20 W has been delivered to each loop. Slight noise but still quite a clear configuration of the applicator set-up can be observed. Fourth case (right down) gives situation when 40 W has been delivered to each loop. In this case noise disturbed the possibility to observe the configuration of the applicator.
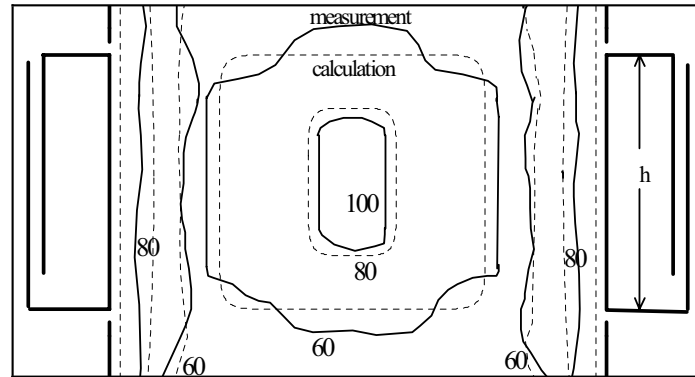
Figure 7: Normalised SAR distribution (calculated and measured) in the heated agar phantom.
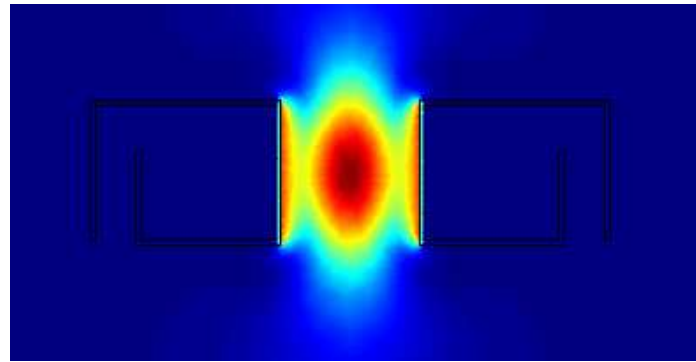
Figure 8: Numerical SAR analysis.
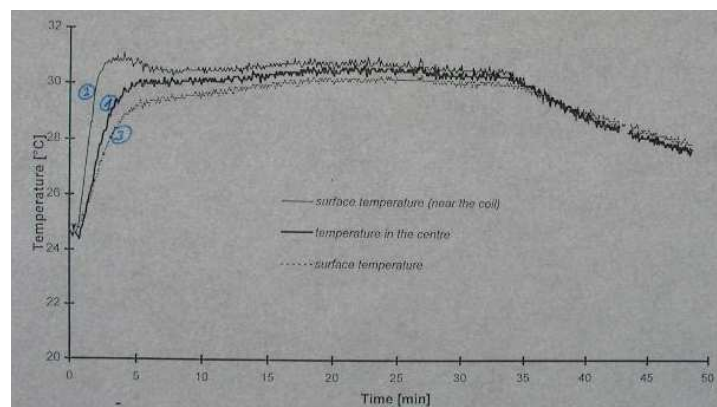
Figure 9: Temperature measurements.

In theoretical and experimental evaluation, the grade of homogeneity of the temperature distribution in the target area has been tested, see the Fig. 7. Our mathematical approach is based on idea of waveguide TM_{01} mode excited in the agar phantom under the given conditions (see the dashed lines). Measurement of SAR (full lines) has been done on agar phantom of the muscle tissue. Very good agreement have been obtained when verifying these results numerically (Fig. 8).

In Fig. 9, we can see temperature vs. time measurement in the case of agar phantom inserted in the studied applicator.

Next figure shows experimental setup of applicator and simple exposure chambers installed at Medical Faculty of Charles University in Pilsen.

Figure 10: Exposure system for research of EM field and biological system interactions.

5. CONCLUSIONS

In this paper we have described basic microwave technical equipment for biological experiments, i.e., waveguide applicator and so called array applicator, including its basic evaluations.

As a novel results of our work we could mention that the new types of microwave applicator for basis biological research of cancer treatment and EM field interactions with biological systems have been developed and evaluated. Evaluation procedures have shown, that this applicator is a very effective heating structure and a compatibility with US and MR has been approved as well.

Having approved these applicators in animal experiments, we are now working on development of their clinical versions to be used in clinical praxis soon.

ACKNOWLEDGMENT

REFERENCES

1. Vrba, J., *Medical Applications of Microwaves (in Czech)*, CTU Press, Prague, 2003.
2. Hand, J. and J. R. James (editors), *Physical Techniques in Clinical Hyperthermia*, Wiley, New York, 1986.
3. Franconi, C., J. Vrba, et al., "27 MHz hybrid evanescent-mode applicators," *Int. Journal of Hyperthermia*, Vol. 9, No. 5, 655–674, 1993.
4. Vrba, J., C. Franconi, et al., "Evanescent-mode applicators for hyperthermia," *IEEE Trans. on Biomedical Engineering*, Vol. 40, No. 5, 397–407, May 1993.

Microwave Medical Imaging and Diagnostics

Jan Vrba[1], Ladislav Oppl[1], Jan Vrba, Jr.[2], and David Vrba[1]

[1]Czech Technical University in Prague, Dept. of EM Field, Prague, Czech Republic
[2]RWTH Aachen University, Chair of Electromagnetic Theory
Kopernikusstraβe, Aachen 1652074, Germany

Abstract— Future trends in medical applications of microwave technique and technology can be seen in development of new diagnostic and imaging methods based on high frequency EM field. A significant importance for the future can be identified for the following methods: Microwave tomography, Microwave radiometry, Measurement of complex permittivity, Imaging in the Terahertz waves band and Microwave diagnostic radars.

1. INTRODUCTION

Interactions of EM field with biological systems are utilised in the area of therapy (oncology, physiotherapy, urology atp.) from late seventieth of last century. Wideutilization of microwave thermotherapy can be observed in the countries of EU, USA and Japan. Our activities in microwave thermotherapy in former Czechoslovakia started in the year 1981. Since 1990, we are member of ESHO (European Society for Hyprthermia Oncology), which co-operates with NAHS (North American Hyperthermia Society) and ASHO (Asian Society of Hyperthermia Oncology).

Recent trends in microwave medical applications are to study the possibilities to develop new diagnostics based on EM field resp. on microwave technique. A significant importance for the future can be identified for the next methods:

- Magnetic resonance,
- Microwave tomography,
- Microwave radiometry,
- Measurement of complex permittivity,
- Imaging with terahertz waves,
- Microwave diagnostic radar.

We will not talk here about magnetic resonance, as it is just well known and broadly used application of EM field in medical diagnostics. We will focus here on other above mentioned methods (excluding microwave diagnostics radars).

2. MICROWAVE APPLICATORS FOR MEDICAL IMAGING AND DIGNOSTICS

Since 1981, we develop waveguide applicators working in frequency band from 27 MHz up to 2450 MHz. These applicators were used for the treatment of more then 1000 patients with superficial or sub-cutaneous tumors (up to the depth cca 4–6 cm). Now, following new trends in this field, we continue our research in the important directions of deep local and regional applicators. We have found, that quite similar applicators are optimal to be used for medical imaging and diagnostics.

Examples of mentioned applicators are given in the Fig. 1. Both applicators have aperture 18×12 cm and are operating at 434 MHz. Waveguide applicators in Fig. 1 (on the left side) are filled by dielectric material in order to decrease their cut off frequency Evanescent mode applicator (Fig. 1 — on the right side) is excited under its cut-off frequency.

On the next figure, there is a sketch of applicator working at 70 MHz and its SAR distribution — including influence of water bolus between applicator and phantom of biological tissue.

Interesting possibility how to achieve relatively sharp beam from external applicator is to use a focusing principle. The aperture of a standard rectangular waveguide applicator is divided into 3 or 5 sectors with shifted excitation (i.e., different amplitude and phase). The basic schematics of this lens type applicator is shown in the following figure.

Figure 1: Dielectric filled waveguide applicator and Evanescent mode applicator.
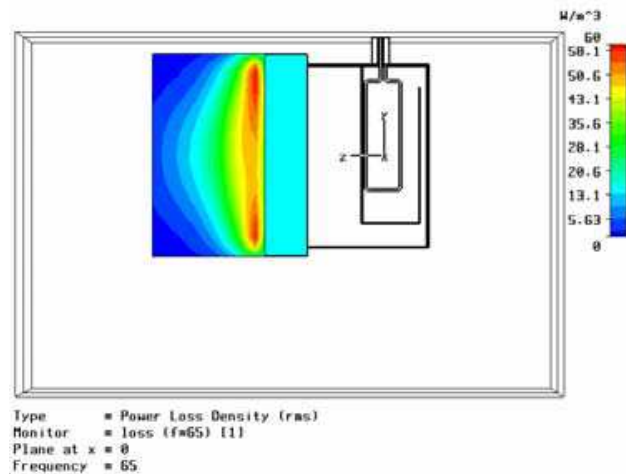
Figure 2: Evanescent mode applicator and its calculated SAR pattern.

3. MICROWAVE TOMOGRAPHY

Microwave tomography [1] is in general application of basic CT principals but by utilization of microwave band. Scattering of EM waves in non-homogeneous human body is however much more complicated than simple attenuation of ionising radiation. Therefore development of microwave tomography is conditioned by new theoretical approach, optimization of evaluation algorythms and more efficient computer technique.

An experimental setup is schematically shown on next figure. Studied object will be placed in water phantom. It will be irradiated by transmitting antenna while scattered EM field will be monitored by receiving antenna and evaluated by a network analyser. Receiving antenna will be scanning around studied object and/or it will be possible to move/rotate the studied object. Measured data will be then reconstructed on the basis of Fourier transform.
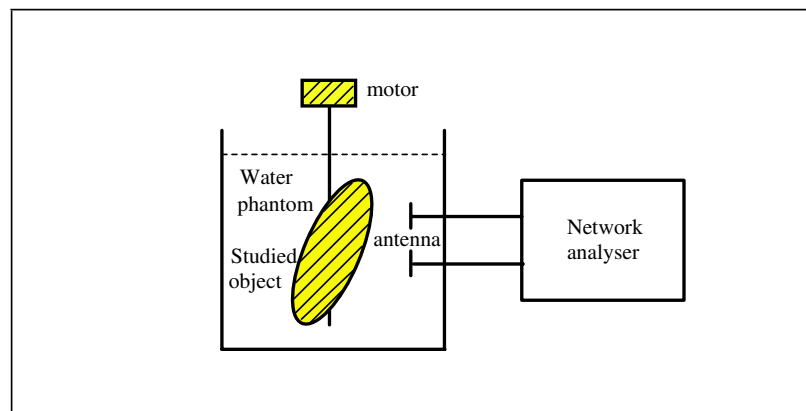
Figure 3: Schematics of experimental setup of microwave tomograph for biomedical imaging.

Microwave tomography represents applications of CT known principals to microwave frequency band, where in general situation is more complicated because of much more complicated propagation of EM waves. Therefore mathematical model should be created and optimized evaluation algorithms are needed and last but not least-powerful computers are necessary for the calculations.

4. MICROWAVE RADIOMETRY

Microwave radiometry is based on measurement of a very weak EM signal, which radiate any object (e.g., people), whose temperature is superior to absolute zero [1]. It is based on utilization of so-called Planck radiation law. Interest in microwave radiometry is given by possibility of its utilization at diagnostics of cancer and also of inflammatory disorder (e.g., appendicitis, arthritis, atp.) because tumors and inflammatory processes causes temperature rise. Microwave radiometer

as a tool for biomedical imaging applications has the possibility to "monitore" a thermal noise produced by objects with the temperature over absolute zero. Next figure gives a basic idea about experimental setup. Advantage of microwave radiometer is ability to "see" the temperature increase under the surface of human body. Therefore wee need to scan studied area of the tissue with a sensor and to evaluate the results of temperature measurements.
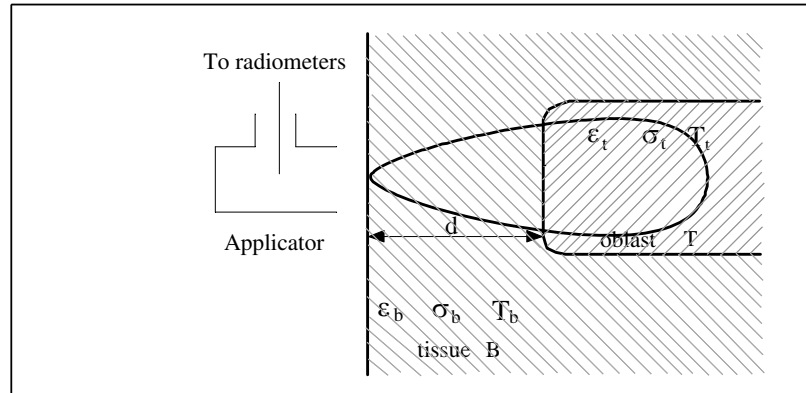

Figure 4: Principals of imaging by a microwave radiometer.

Let us suppose, that applicator (antenna of the radiometer) is situated on the layer B of the monitored biological tissue. Thickness of this layer is d and it temperature T_b. Its dielectric parameters will be ε_b and σ_b.

Further we have in this figure region of biological tissue T (in the depth d) where temperature is $T_t = T(z)$ and dielectric parameters are ε_t and σ_t. Then we can apply following equation to determine the temperature measured by discussed microwave radiometer.

$$T_r = 2?\beta(z) \cdot T(z) \cdot e^{-2?(z) \cdot dz} dz + T_d \cdot e^{-2?(z) \cdot dz}$$

where $\beta(z)$ is attenuation in the studied area. Often we can simplify this case in following way

$$\beta(z) = \delta.$$

Then initial equation can be rewritten as follows

$$T_r = 2?T(z) \cdot e^{-2?} dz + T_d \cdot e^{-2?}$$

If also temperature $T(z)$ will be constant and equal to T_b, then we can use modified expression
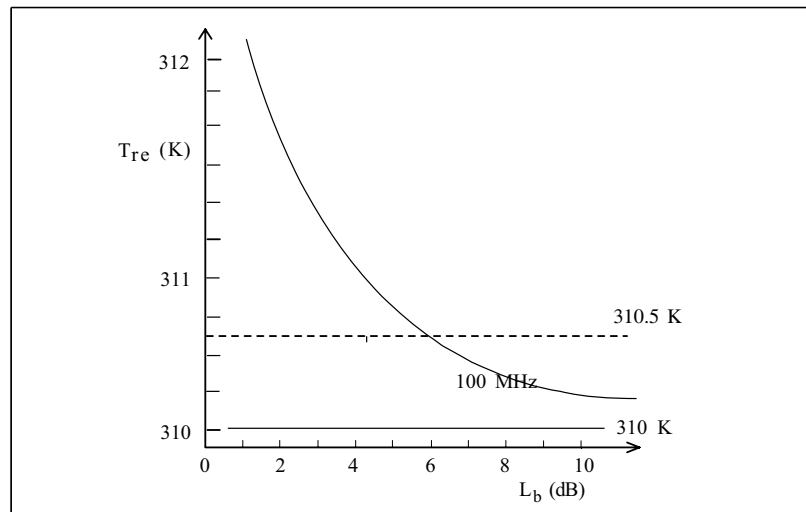

Figure 5: Temperature measured by radiometer with respect to tissue attenuation L_b.

to determine the temperature measured by discussed microwave radiometer

$$T_r = \left(1 - \frac{1}{L_b}\right) T_b + \frac{1}{L_b} T_t$$

where L_b is attenuation of the layer B for electromagnetic wave at given frequency.

In the case of low level of L_b, then radiometer will see the temperature of region T, i.e., T_t. If L_b will increase, then reading of radiometer will approach to value of T_b.

In the next figure, we can see relation between discussed quantities for the case, when $T_b = 310\,\text{K}$ and $T_t = 312\,\text{K}$ and sensitivity of a radiometer is $0.5\,\text{K}$ so we can detect increase in temperature for levels of $310.5\,\text{K}$. Temperature measured by radiometer is in this figure a function of L_b.

According to this figure attenuation in region B can be only $6\,\text{dB}$ at maximum, otherwise radiometer "will not see" increased temperature in region T.

At following figure, there are the graphs of the temperature which would be measured by a radiometer at different frequencies.
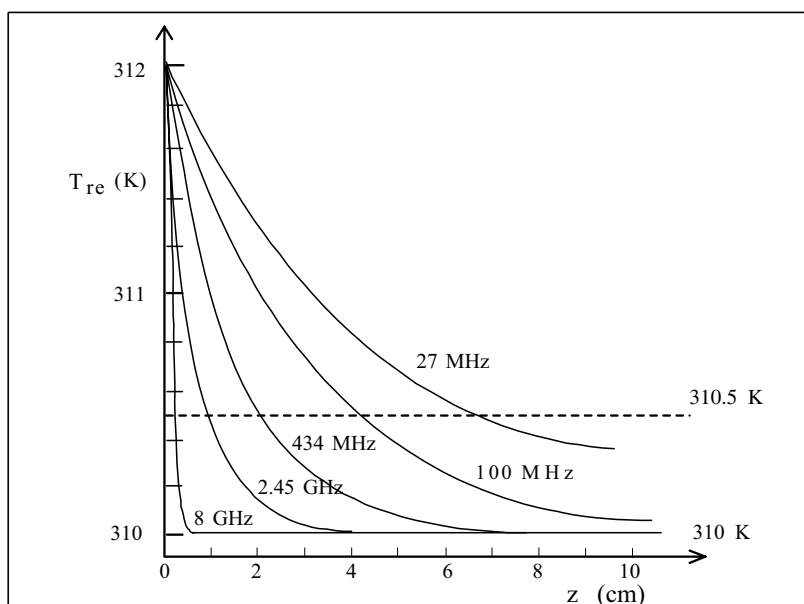
Figure 6: Temperature measured by radiometer with respect to tissue attenuation at different frequencies.

5. MEASUREMENTS OF COMPLEX PERMITTIVITY

Measurements of complex permittivity "in vivo" could be a suitable for biomedical imaging applications [1]. Usually an open end of coaxial line is used as a very suitable sensor for this measurement. Scanning the studied object by a such probe can bring us a map of the permittivity — we can then evaluate symmetry resp. unsymmetry of the measurement results and from this information we can make hypothesis about possible medical problems.
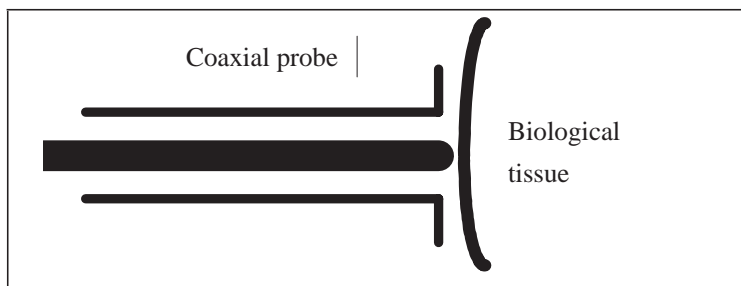

Figure 7: Coaxial probe for measurement of.

Characteristic impedance is Z_o. C_o and C_f are fringing capacities in the equivalent circuits, which can be determined by aid of calibration methods. Reflection coeficient can be determined as

$$" = |"|e^{j?}$$

Then for complex permitivity (its real and imaginery part) can be derived following equation

$$? = \frac{-2|"|\sin?}{?C_oZ_o(1 + 2|"|\cos + |"|^2)} - \frac{C_f}{C_o}$$

$$? = \frac{1 - |"|^2}{?C_oZ_o(1 + 2|"|\cos + |"|^2)}$$

6. IMAGING WITH TERAHERTZ WAVES

Frequency band of so called Terahertz waves (0.1–10 THz) is being studied during last years [1]. We can expect here a lot of new discoveries in material science and in biomedicine as well, especially for imaging purposes.

Theoretical models and a feasibility study can be based on similar principles as Microwave Radiometer and Infrared Camera Imaging.

Terahertz waves are "situated inbetween" microwave frequency band and infrared band. Therefore we can expect following properties:

1) Terahertz waves compared to microwaves will ofer better space resolution, but somewhat worse capability to recover image from deep positions.

2) On the contrary, comaprison of terahertz waves to infrared will result in better capabilty to recover image from deeper position but its space resolution will be worse.

7. CONCLUSIONS

As novel results of our work we could mention that waveguide applicators and prospective methods for microwave medical diagnostics have been described and discussed.

ACKNOWLEDGMENT

REFERENCES

1. Vrba, J., *Medical Applications of Microwaves*, (In Czech), 1st ed., 168, Issued by CTU, Prague, 2003.
2. Semenov, S. Y., et al., "Three-dimensional microwave tomography, initial experimental imaging of animals," *IEEE Transactions on BME*, Vol. 49, No. 1, 55–63, Jan. 2002.
3. Gabriel, S., R. W. Lau, and C. Gabriel, "The dielectric properties of biological tissue — II. Measurements in the frequency range 10 Hz to 20 GHz," *Phys. Med. Biol.*, Vol. 41, 2251–2269, 1996.
4. Fear, E. C. and M. A. Stuchly, "Microwave detection of breast cancer," *IEEE Trans. Microwave Theory Tech.*, Vol. 48, 1854–1863, 2000.
5. Vrba, J., et al., "Technical aspects of microwave thermotherapy," *RF Interaction with Humans: Mechanisms, Exposures and Medical Applications. IPEM Meeting*, Institute of Physics, London, Feb. 2003.
6. Vrba, J., C. Franconi, F. Montecchia, and I. Vanucci, "Evanescent mode applicators for subcutaneous hyperthermia," *IEEE Trans. on Biomedical Engineering*, Vol. 40, No. 5, 397–407, May 1993.
7. Franconi, C., J. Vrba, and F. Montecchia, "27 MHz hybrid evanescent-mode applicators with flexible heating field for deep hyperthermia," *International Journal of Hyperthermia*, Vol. 9, No. 5, 655–673, 1993.

Reversible Electroporation on a Microchip

H. S. Kim[1], **H. B. Kim**[2], **and J. H. Yi**[1]

[1]Department of Biomedical Engineering, Konkuk University, Republic of Korea
[2]Solco Biomedical Institute, Republic of Korea

Abstract— The efficiency of electroporation for the porcine aortic endothelial valve cells is studied using pulsed dc electric fields and a new microchip applicator. The results indicate that maximum electroporation efficiency of over 80% can be achieved by choosing a proper amplitude and number of pulses. The result shows that increased field strength provides higher electroporation efficiency and maintains high cell viability within a limited value of applied electric charge. Our study also shows that the electroporation is reversible when the electric field is within $1\,\text{kV/cm}$.

1. INTRODUCTION

Electroporation technique uses application of extracellular electric field and involves rapid structural rearrangement and formation of pores on the lipid bilayer of the cells. The technique is becoming an important technique for inserting biologically active molecules such as drug and gene into the cells [1–5]. The technique uses a non-thermal process and relatively low external electric field pulses of short duration thereby maintaining the integrity of the outer cell membrane and preventing further irreversible damage.

2. THEORY

Transmembrane potential can be enhanced in a cell membrane by applying external electric short pulses. The resulting transmembrane potential across the cell membrane induces localized structural rearrangements within the membrane, resulting in the formation of pores which perforate the membrane, and an attendant dramatic increase of ionic and molecular transport through the cell membrane. The threshold field strength required for the electroporation across a cell membrane of $5\,\text{nm}$ thickness is in the range of $0.2\,\text{V} \sim 1\,\text{V}$. For a suspended spherical cell of radius r_c, the transmembrane potential (V_m) is given by

$$V_m = \frac{1.5 E r_c \cos\theta}{\sqrt{1 + (2\pi f \tau)^2}}, \tag{1}$$

where E is strength of the homogeneous electric field over the space, f is pulse frequency, θ is a angle between the cell orientation and the electric field, and τ is the relaxation time constant of the cell membrane. The τ has value of micro second order. For dc fields and low frequency ac field, therefore, this equation can be reduced to:

$$V_m = 1.5 E r_c \cos\theta. \tag{2}$$

Equation (2) is useful when the resistance of the membrane is very high and the most cell membranes have high-enough resistance. Even if the electric field strength is the most important one, result of the electroporation can be influenced by the other physical parameters which are not included in the Eq. (1) such as pulse shape, number of pulses, pulse interval, thermal condition, and the other environmental factors. Since there is various parameters to affect the result of electroporation, parameter setting for an experiment is determined empirically and range of optimal or useful values for some parameters are published. It is difficult, however, to set the formulas for optimal parameter values. Finding a set of optimal parameter values through out a set of experimental procedures for a cell type is still worth to do.

3. METHODS

3.1. Cells

Aortic endothelial valve cells of swine were cultured in Dulbecco's Modified Eagle's Medium (DMEM) (Category D6046, Sigma), 2 mM L-glutamine (Category G3126, Sigma), $50\,\mu\text{g}\cdot\text{m}^{-1}\cdot\text{L}^{-1}$ uridine and 10% foetal calf serum (Sigma). To analyze an electroporation efficacy the cells were primarily cultured in a microchip applicator (Fig. 2) which has six identical electroporation sites on a Pyrex substrate and each site has an electrode pair of 1 mm inter-electrode distance. Incubation condition was 5% CO_2 at 37°C.

3.2. Dyes

Calcein-AM (Sigma Cat. No. C0875) 2 µm was mixed with the cell culture media DMEM before treatment. Since dead cells cannot be stained by the calcein-AM, the dye was used to confirm live cells from dead cells as a post-electroporation marker for live or dead. And propidium iodide (PI) was also mixed with the DMEM in a rate of 1 to 20 (v/v) before electric field application. The PI cannot cross an intact cell membrane without the pores on the membrane. Thus uptake of the PI can be a good index of electroporation efficiency. To measure cell viability after electroporation, we used a MTT assay in difference with the above to be sure massively the cells.

3.3. Electroporation Apparatus

We used a set of home-made electroporation devices named KS-500. The set is composed of 2 units, i.e., high-voltage pulse generator (Fig. 1) and a microchip applicator (Fig. 2). The pulse generator has 3 circuit blocks, i.e., a digital controller, a high-voltage booster, and a high-voltage switch. The digital controller generates single pulse or pulse train to trigger the high-voltage switch according to an application protocol of an electroporation experiment. The high-voltage booster boosts 12 volt dc to a high-voltage dc (50–500 V) and the switch generates a single high-voltage pulse or a pulse train by switching the boosted dc. Three thumb-wheel digital switch blocks are used in the generator to select pulse width (100 µs ∼ 1 ms), number of pulses (1–9 pulses), and pulse interval (100 µs ∼ 1 sec) and a 10-turn potentiometer is adopted to set the level of the high-voltage dc thereby defining magnitude of the high-voltage pulses.

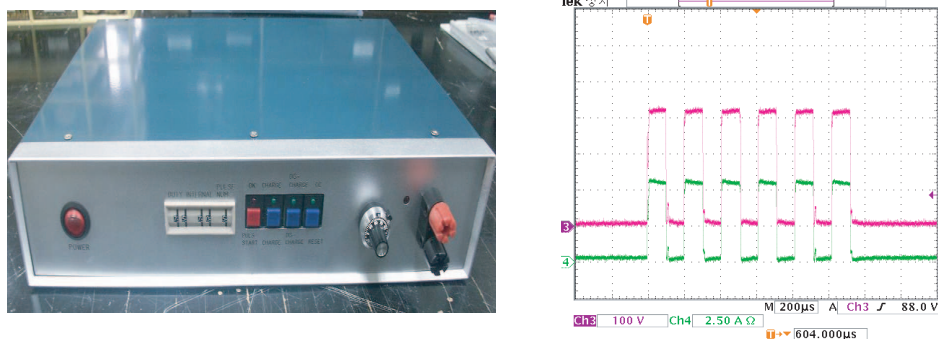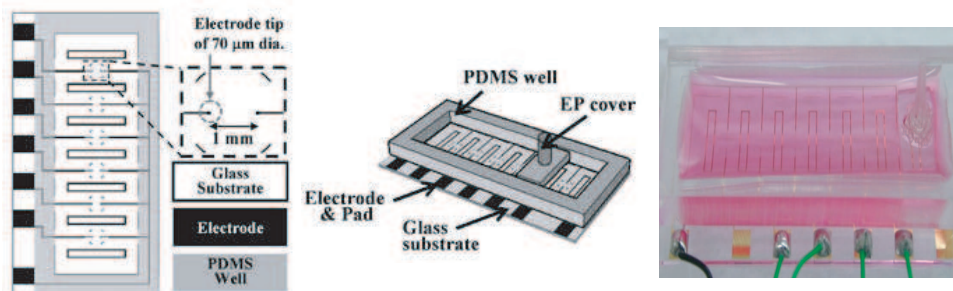Figure 1: A pulse generator and waveforms of the voltage and the current for 60 Ω load.

Figure 2: A microchip electroporation applicator.

The high-voltage booster runs when a user turns on the charge switch and finishes when the voltage reaches to the set-voltage with turning on a HV-OK LED (light emitting diode). The boosting operation takes less than 7 seconds. A user can trigger the designed high-voltage pulses using pulse start switch. A large capacitor (4,000 µF) in the high-voltage booster supports output current until a pulse train ends. There is an over current protection circuit in the high-voltage switch to prevent excessive current flow between the electrodes of the applicator. When the system is turned off, the capacitor in the booster circuit is automatically discharged through a resistor to prevented dangerous electric hazards. The capacitor can also be discharged by user using a discharge switch on the control panel of the system.

The microchip applicator is constructed on a 1 mm thick slide glass using soft lithography and it has six identical electroporation sites on a chip. The electrodes are fabricated using 0.2 μm gold layer. Each electrode has circular end of 70 μm diameter and inter-electrode distance is 1 mm. Shapes and dimensions are shown in Fig. 2. Volume of the cells and media is determined by the cover and the cover has a cavity for the cells and media and dimension of the cavity is 3.0 mm (depth) × 100 μm (width) × 10 μm (height). Six identical experiments with identical dimensions and electric field strength and distribution can be conducted using the new applicator.

3.4. Electroporation Protocol

To measure the electroporation efficacy, experiments using sequences of mono-phasic pulses were performed; pulse width = 100 μs, pulse interval = 100 μs, number of pulses = 2 and 5, and pulse amplitudes = 375, 750, 1 k, and 1.3 kV/cm, respectively. The pulses of different values were applied across the cells between two electrodes on the microchip site so that we can investigate a relationship between electroporation efficacy and the pulse parameters. Results were evaluated using a fluorescence microscope by calculating areas of total cells and electroporated cells. While un-pored cells can be measured by green only, the pored cells can be measured by both green and red. The EP (electroporation) efficiency is defined by

$$\text{EP efficiency} = \frac{\text{red fluorescence area of electroporated cells}}{\text{green flulorescence area total cells}} \times 100[\%] \qquad (3)$$

where green fluorescence area of total cells and red fluorescence area of electroporated cells are measured within a view-area of $8.7 * 8.7\,\mu\text{m}^2$. The areas were evaluated using Image J [6].

4. RESULTS

Different numbers of pulses and electric field strengths were used to investigate the relationship between the cell electroporation and pulse parameters. We used a fixed pulse frequency for simplicity of the experiment and we couldn't find any changes in the results when we changed the frequency by double or half. The quantitative assess were carried out as described below.

To investigate effect of electric field strength on the electroporation, we applied different electric field strengths (375 ∼ 1.3 kV/cm) to the cells using developed apparatus with fixed pulse width and interval, i.e., 100 μs, and 2 different number of pulses, i.e., 2 and 5. The results show that the EP efficiency has proportionality to electric field strength and supplied electric charge (5 pulse experiment shows better result at the same field) and there is a threshold of saturation. When the field strength was 1 kV/cm with 5 pulses, the efficiency had 88.9%. The two-pulse and five-pulse experiments showed big difference of the efficiency, i.e., 15.9% and 88.9%, respectively. All the experiment showed cell viability better than 91%. Results of decreased cell viability for increased number of pulses were presented from the previous studies. However, we couldn't find evidence of reduces cell viability by increasing number of applied pulses until a saturation value of applied charge. The applied electric charge is proportional to the number of pulses when the pulse width and amplitude are fixed. Our experiment shows that 5 pulses of 100 μs width at 1 kV/cm is a near-saturation condition. The high value of cell viability means that the electroporation we had conducted is reversible process and the electroporated cell membrane can be resealed within 3 min.
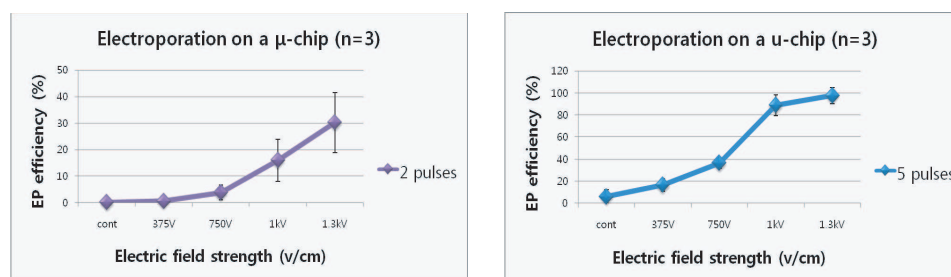

Figure 3: Results of electroporation EP efficiency EP efficiency.

5. CONCLUSIONS

In vitro electroporation experiments were performed using a new microchip applicator and efficiency of the electroporation was evaluated. Our study shows that the electroporation efficiency has proportionality to applied electric charge until it reaches its saturation limit and there is no strong relationship between cell viability and applied charge of field strength. The new microchip applicator can be a useful device for electroporation study and to assess electroporation-induced drug delivery for cancer treatment in vitro because of its consistency of the

REFERENCES

1. Canatella, P. J., J. F. Karr, J. A. Petros, and M. R. Prausnitz, "Quantitative study of electroporaiton-mediated molecular uptake and cell viability," *Biophys. J.*, Vol. 80, 755–64, 2001.
2. Jordan, D. W., et al., "Effect of pulsed, high-power radiofrequency radiation on electroporation of mammalian cells," *IEEE Trans. Plasma Sci.*, Vol. 32, 1573–8, 2004.
3. Chang, D. C., "An efficient method of microinjecting attached cells using electroporation," *Mol. Biol. Cell*, Vol. 3, A90, 1992.
4. Kirson, E. D., et al., "Disruption of cancer cell replication by alternating electric fields," *Cancer Res.*, Vol. 64, 3288–95, 2004.
5. Kotnik, T., et al., "Cell membrane electropermeabilization by symmetrical bipolar rectangular pulses: Increased efficiency of permeabilization," *Bioelectrochemistry*, Vol. 54, 83–90, 2001. http://rsb.info.nih.gov/ij/

Wideband Differential Phase Shifters Using Waveguides Filled by Inhomogeneous Dielectrics

Mohammad Khalaj-Amirhosseini

Iran University of Science and Technology, Iran

Abstract— A new structure is proposed for wideband differential phase shifter. This structure utilizes a waveguide filled by a longitudinally inhomogeneous dielectric along with a simple uniform waveguide. To optimally design the structure, the electric permittivity function of the inhomogeneous waveguide is expanded in a truncated Fourier series, firstly. Then, the optimum values of the coefficients of the series are obtained through an optimization approach to have low phase shift error and low reflection coefficient in a desired frequency bandwidth. The usefulness of the proposed structure is studied using some examples in the X-Band.

1. INTRODUCTION

Differential Phase Shifters (DPSs) have many applications in microwave circuits such as phase discriminators, beam forming networks, frequency translators, power dividers and phase array antennas. The main disadvantage of the conventional waveguide DPSs, which utilize dielectric cards or slabs in the waveguides [1, 2], is their narrow band property. In this paper, we propose utilizing Longitudinally Inhomogeneous Waveguides (LIWs) as a wideband DPS in a desired frequency range. LIWs are waveguides whose inside is totally filled by longitudinally inhomogeneous dielectric or several thin dielectric layers [1–10]. To optimally design LIWs, the electric permittivity function of them is expanded in a truncated Fourier series, first. Then, the optimum values of the coefficients of the series are obtained through an optimization approach. The usefulness of the proposed structure is verified using some examples.

2. THE PROPOSED DPS

In this section the new structure for DPSs is proposed. The proposed structure, as shown in Fig. 1, consists of two transmission lines, one hollow waveguide as a phase reference and one Longitudinally Inhomogeneous Waveguides (LIW). The waveguides are of dimensions a and b and the length of the hollow waveguide and LIW are d_0 and d, respectively. Also, the inside of LIW has been filled by a longitudinally inhomogeneous dielectric $\varepsilon_r(z)$, which can be constructed by several thin dielectric layers. The phase differences between two matched ends of the hollow waveguide and LIW at frequency f are as follows, respectively

$$\varphi_0(f) = \text{Angle}\left(\exp(-j2\pi f\sqrt{1-(f_c/f)^2}d_0/c)\right) \tag{1}$$

$$\varphi(f) = \text{Angle}\left(S_{21}(f)\right) \tag{2}$$

where c is the velocity of the light and $f_c = c/(2a)$ is the cutoff frequency of the hollow waveguide. We would like that the difference between these phase differences ($\Delta\varphi = \varphi - \varphi_0$), calling it the phase shift, be equal to the desired value $\Delta\varphi_d$ at all frequencies in the desired frequency bandwidth, i.e., from $f_{\min}$ to $f_{\max}$.
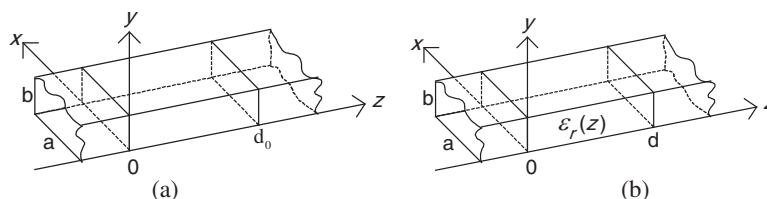

Figure 1: The proposed DPS. (a) A hollow waveguide as a phase reference. (b) A waveguide filled by a longitudinally inhomogeneous dielectric (LIW).

3. SYNTHESIS OF DPS

In this section a general method is proposed to design optimally the proposed DPS. Firstly, we consider the following truncated Fourier series expansion for the electric permittivity function.

$$\ln(\varepsilon_r(z) - 1) = \sum_{n=0}^{N} C_n \cos(n\pi z/d) \tag{3}$$

An optimum designed DPS has to have constant phase shift in a defined frequency range. The optimum values of the coefficients C_n and the length d could be obtained through minimizing the following error function related to K frequencies between $f_{\min}$ and $f_{\max}$.

$$\text{Error} = \sqrt{\frac{1}{K} \sum_{k=1}^{K} |\text{unwrap}(\varphi(f_k)) - \text{unwrap}(\varphi_0(f_k)) - \Delta\varphi_d|^2} \tag{4}$$

where unwrap (p) is a function that unwraps phases p by changing absolute jumps greater than π to their 2π complement. Moreover, defined error function should be restricted by some constraints such as low input reflection and easy fabrication, which could be written as follows

$$\max(|S_{11}(f)|) \le \rho_{\max} \tag{5}$$

$$\max(\varepsilon_r(z)) \le \varepsilon_{r,\max} \tag{6}$$

where $\rho_{\max}$ is the maximum allowable reflection coefficient and also $\varepsilon_{r,\max}$ is the maximum available dielectric constant. To obtain the phase difference $\varphi(f)$, we have to analyze the LIW. There are some methods to analyze the LIWs such as finite difference [4], Taylor's series expansion [5], Fourier series expansion [6], the equivalent sources method [7], the method of Moments [8] and approximate analytic solution [9]. Of course, the most straightforward method is subdividing LIWs into many uniform or linear electrically short sections [10] with length Δz, which

$$\Delta z \ll \lambda_{\min} = \frac{c}{f_{\max}\sqrt{\varepsilon_r}} \tag{7}$$

Then the $ABCD$ matrix of the LIW will be as the multiplication of the $ABCD$ matrices of all short sections. The parameters A, B, C and D could be used to find the S parameters of the LIW as follows

$$S_{11} = \frac{AZ_0 + B - CZ_0^2 - DZ_0}{AZ_0 + B + CZ_0^2 + DZ_0} \tag{8}$$
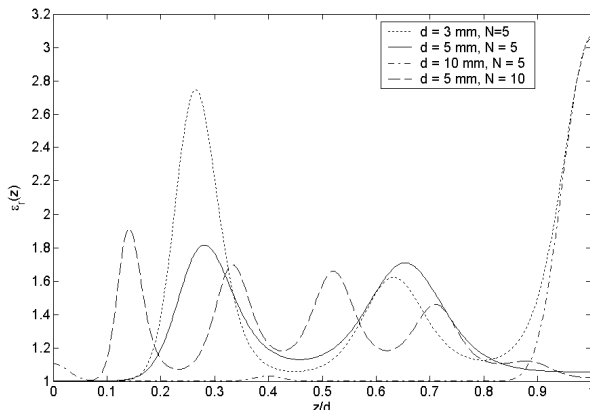
$$S_{21} = \frac{2Z_0}{AZ_0 + B + CZ_0^2 + DZ_0} \tag{9}$$

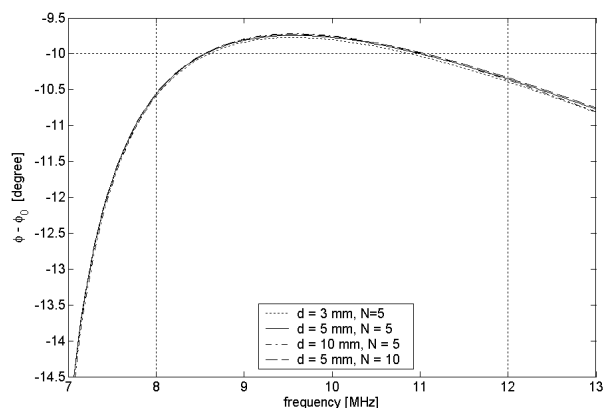Figure 2: The permittivity function of the designed DPS with $\Delta\varphi_d = -10°$.

Figure 3: The phase shift $\varphi - \varphi_0$ versus frequency for $\Delta\varphi_d = -10°$.

It is of worth to mention that it is possible to consider both transmission lines in Figs. 1(a) and 1(b) inhomogeneous, i.e., two LIWs, calling them Double DPS. In these structures, one has to use two individual truncated Fourier series expansion (3) for each LIW.

4. EXAMPLES AND RESULTS

We would like to design an LIW as a DPS with $\Delta\varphi_d = -10°$ in a frequency range from 8.0 to 12.0 GHz (X-Band), considering $a = 0.9$ inch, $b = 0.4$ inch (WR-90), $\varepsilon_{r,\max} = 10$, $\rho_{\max} = -14$ dB, $d_0 = 3$, 5 or 10 mm and $N = 5$ or 10. Figs. 2–4 illustrate the permittivity function $\varepsilon_r(z)$, $\varphi - \varphi_0$ and $|S_{11}|$, respectively. It is seen that the absolute of the phase error is less than 5.6% and also the absolute of the input reflection coefficient is less than -14 dB for all d_0s. It is evident that as the length of the structure is chosen larger the fabrication becomes easier. Now, we would like to design a DPS with $\Delta\varphi_d = -90°$ in X-Band considering $d_0 = 45$ mm and $\rho_{\max} = -13$ dB. This phase shifter could be designed by cascading of $M = 9$ similar sections with $d_0 = 5$ mm. Figs. 5–7 illustrate the permittivity function $\varepsilon_r(z)$, $\varphi - \varphi_0$ and $|S_{11}|$, respectively. It is seen that the absolute
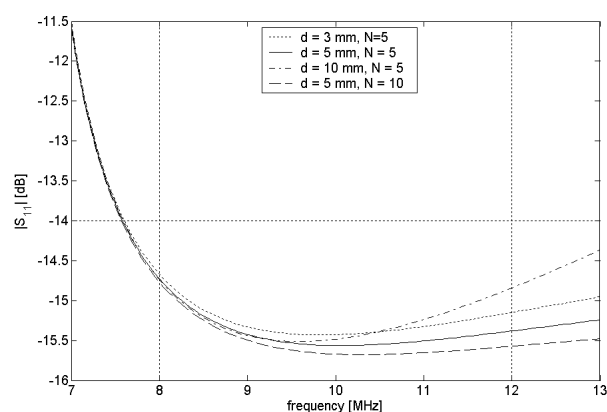

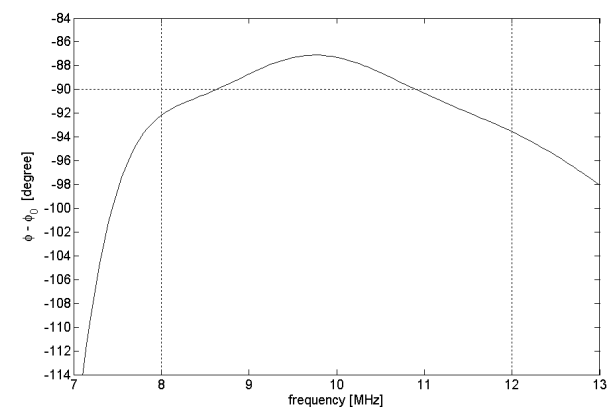Figure 4: The absolute of S_{11} versus frequency for $\Delta\varphi_d = -10°$.
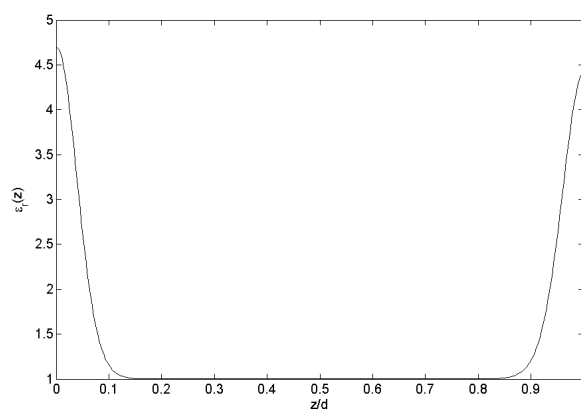

Figure 5: The permittivity function of one section of the designed DPS with $\Delta\varphi_d = -90°$.

Figure 6: The phase shift $\varphi - \varphi_0$ versus frequency for $\Delta\varphi_d = -90°$.

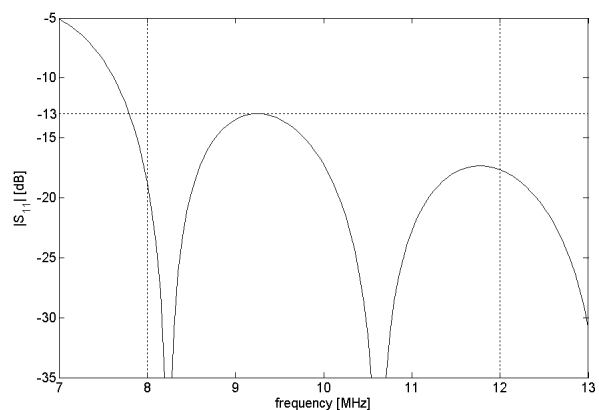Figure 7: The absolute of S_{11} versus frequency for $\Delta\varphi_d = -90°$.

Table 1: Optimum values of the coefficients C_n for $N = 5$.

	C_0	C_1	C_2	C_3	C_4	C_5
$d_0 = 3$ mm ($\Delta\varphi_d = -10°$)	-2.3189	-2.3746	-1.6393	-2.1184	-1.6773	-1.8705
$d_0 = 5$ mm ($\Delta\varphi_d = -10°$)	-2.3734	-1.3343	-2.1337	-0.9762	-1.5977	-0.9041
$d_0 = 10$ mm ($\Delta\varphi_d = -10°$)	-10.000	0.2919	5.0145	-6.0672	4.2148	4.2855
$d_0 = 9 \times 5$ mm $= 45$ mm ($\Delta\varphi_d = -90°$)	-9.9979	-0.9094	9.5853	1.1280	1.6807	-0.1802

of the phase error is less than 4.3% and also the absolute of the input reflection coefficient is less than -13 dB. Tables 1–2 show the optimum values of the parameters C_n and d, respectively, for all the designed DPSs. Using the above examples, one may satisfy about the capability of designing LIWs as DPSs.

Table 2: Optimum value of the length d.

	d (mm)
$d_0 = 3\,\mathrm{mm}\ (\Delta\varphi_d = -10°)$	2.96
$d_0 = 5\,\mathrm{mm}\ (\Delta\varphi_d = -10°)\ (N = 5,\ N = 10)$	4.96
$d_0 = 10\,\mathrm{mm}\ (\Delta\varphi_d = -10°)$	9.92
$d_0 = 45\,\mathrm{mm}\ (\Delta\varphi_d = -90°)$	$9 \times 4.74 = 42.66$

5. CONCLUSION

A new structure was proposed as differential phase shifter in a desired frequency range. This structure consists of two transmission lines, one hollow waveguide as a phase reference and one Longitudinally Inhomogeneous Waveguides (LIW). The LIW is a waveguide filled by a longitudinally inhomogeneous dielectric. To optimally design LIW, the electric permittivity function of it is expanded in a truncated Fourier series, first. Then, the optimum values of the coefficients of the series are obtained through an optimization approach. The usefulness of the proposed structure was verified using some examples in the X-Band. It is observed that the solutions yield a good constant differential phase shift and also low input reflection coefficient in the desired frequency band.

REFERENCES

1. Collin, R. E., *Field Theory of Guided Waves*, IEEE Press, New York, 1991.
2. Collin, R. E., *Foundations for Microwave Engineering*, McGraw-Hill, 1996.
3. Khalaj-Amirhosseini, M., "Microwave filters using waveguides filled by multi-layer dielectric," *Progress In Electromagnetics Research*, PIER 66, 105–110, 2006.
4. Khalaj-Amirhosseini, M., "Analysis of longitudinally inhomogeneous waveguides using finite difference method," *Int. Conf. on Inform. and Comm. Tech.: From Theory to Appl. (ICTTA 2006)*, 799–800, Damascus, Syria, Apr. 24–28, 2006.
5. Khalaj-Amirhosseini, M., "Analysis of longitudinally inhomogeneous waveguides using Taylor's series expansion," *Journal of Electromagnetic Waves and Applications*, Vol. 20, 1093–1100, 2006.
6. Khalaj-Amirhosseini, M., "Analysis of longitudinally inhomogeneous waveguides using the Fourier series expansion," *Journal of Electromagnetic Waves and Applications*, Vol. 20, 1299–1310, 2006.
7. Khalaj-Amirhosseini, M., "Analysis of longitudinally inhomogeneous waveguides using equivalent sources method," *Asia-Pacific Microwave Conf. (APMC 2007)*, Bangkok, Thailand, Dec. 12–14, 2007.
8. Khalaj-Amirhosseini, M., "Analysis of longitudinally inhomogeneous waveguides using the method of moments," *Progress In Electromagnetics Research*, PIER 74, 57–67, 2007.
9. Khalaj-Amirhosseini, M., "A closed form solution for longitudinally inhomogeneous waveguides," *Int. Symp. Antennas, Propagation and EM Theory (ISAPE 2008)*, Kunming, China, Nov. 2–5, 2008.
10. Khalaj-Amirhosseini, M., "Analysis of longitudinally inhomogeneous waveguides by cascading short linear sections," *Int. Symp. Antennas and Propagation (ISAP 2008)*, Taipei, Taiwan, Oct. 27–30, 2008.

The Design of Triple-Mode Low Noise Amplifier for SDR System

Yang Liu, **Sungju Choi**, **Sangho Lee**, **and Hyeongdong Kim**
Microwave Engineering Lab., Electrical & Computer Engineering Department
Hanyang University, Seoul, South Korea

Abstract— A fully integrated triple-mode CMOS low noise amplifier applicable to Software defined Radio (SDR) handset is presented. The target frequency bands include 1.84 GHz–1.87 GHz for CDMA 2000 1x, 2.11 GHz–2.17 GHz for WCDMA, and 2.3 GHz–2.4 GHz for WiBro. The selection of operating frequency band is implemented by four switched-capacitor circuits, where every two of them located at the input and output matching network, each switch is realized by NMOS transistor. The post layout simulation displays the input reflection coefficient and narrow band voltage gain of the LNA in three frequency bands of interest. The analysis of input matching, noise matching, and linearity of the LNA is discussed in this paper.

1. INTRODUCTION

With the wireless communications have been rapidly proceeding, the research of convergence technology which can implement various communication services on a single terminal is desired. SDR is the technology widely recognized as one piece of flexible radio hardware is reutilized for different applications and standards under software control. As the entire standards heavily depend on digital signal processing, making CMOS the first technology of choice. In order to realize multi-standard wireless communication according to the operation mode of SDR system, a new switchable fully integrated triple-mode CMOS LNA with excellent performance at each frequency band has been designed in this paper.

Based on simultaneous noise and input matching technology, the noise performance of single-band amplifier can be extremely enhanced. The triple-mode LNA in this work, however, it need to exhibit three resonance frequencies by controlled switches. Thus, it is difficult to optimize noise performance at each frequency band of interest. For this case, the methodology to design the LNA to achieve high performance of noise figure and input reflection coefficient, and how to lower the power consumption is introduced in Section 2. Section 3 describes circuit techniques to improve linearity. The post layout simulation results of the LNA are given in Section 4. Finally, conclusion is presented in Section 5.

2. TRIPLE-MODE LNA DESIGN

2.1. The Operating Theory of the Triple-mode LNA

We can indeed appreciate the excellent inductive source degeneration LNA topology which has been presented in [1], because the source inductor generates a real part to hold simultaneously input impedance and noise matching, and it does not introduce more noise than any other passive component while this topology can also enhance the linearity and ensure the stability of the amplifier. Figure 1 illustrates the input stage of this LNA topology.

A simple analysis of the input impedance shows that

$$Z_{in} = s(L_s + L_g) + \frac{1}{sC_{gs}} + \left(\frac{g_{m1}}{C_{gs}}\right)L_s \tag{1}$$

The resonance frequency can be easily defined as

$$\omega = \frac{1}{\sqrt{C_{gs}(L_s + L_g)}} \tag{2}$$

At resonance, the first two terms of Equation (1) vanish, and the left term $\frac{g_{m1}}{C_{gs}}L_s$ totally depending on the bias is to make $50\,\Omega$ for impedance matching. The key point for designing triple-mode LNA in Equation (2) is that L_s and L_g are fixed to capture the real part and cancel out the imaginary part of the input impedance respectively, and then, the C_{gs} can be adjusted to change the operating frequency.

Figure 2 shows the complete schematic of the designed LNA. As can be seen from the picture, the cascode amplifier topology is also employed in this work because it is well suitable for high
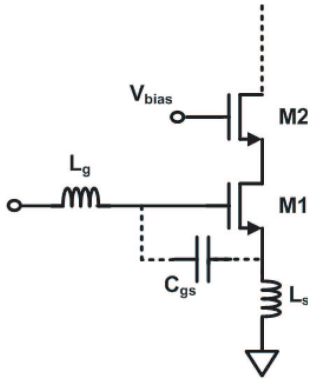
Figure 1: The input stage of source degeneration topology.
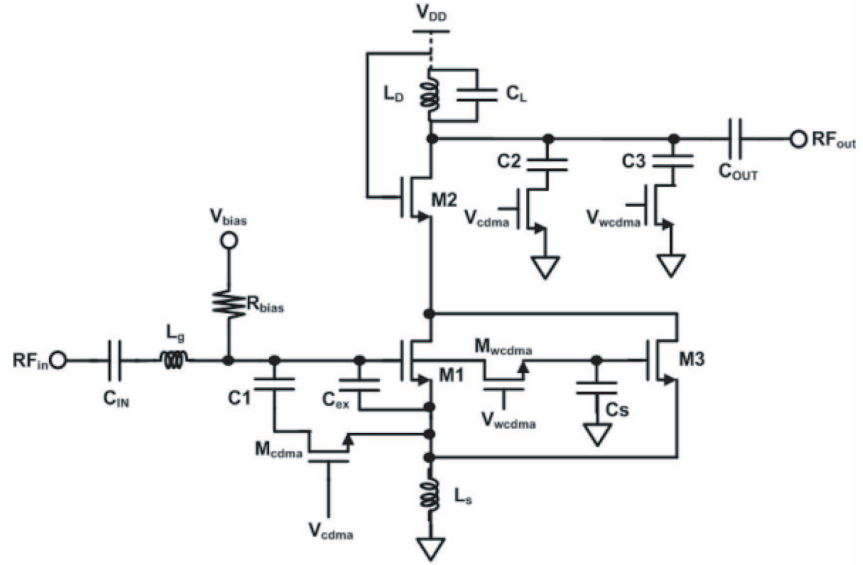

Figure 2: The complete schematic of the proposed LNA.

gain amplifier design and the Miller effect can be suppressed to help to reduce the effective input capacitance.

In the proposed LNA, the CDMA 2000 1x band is selected when the transistor M_{cdma} and M_{wcdma} is turned on at the same time, and it can be changed to WCDMA band by turning off M_{cdma} and turning on M_{wcdma} only, and when both the switches are turned off, the WiBro band can be selected.

2.2. Input Impedance Matching and Noise Matching

As mentioned above, the operating frequency of the LNA can be changed by adjusting the value of gate to source capacitance of input transistor M1. Therefore, the procedure of the design process would be started from the highest frequency band, WiBro, 2.3 GHz–2.4 GHz, and gradually moved to lower frequency band by increasing C_{gs}. Consider the first case when the LNA operates at WiBro band. Although both the switches are turned off, the parasitic elements of the switches M_{cdma} and M_{wcdma} need to be taken into account because the MOS switch couples the clock transitions through its gate-drain and gate-source overlap capacitance [2], and these parasitic elements can directly affect the noise figure. Another aspect should also be noted here, that is M3 is in series with M1. Therefore, the input impedance calculation from Equation (1) should be modified as

$$Z_{in} = s(L_s + L_g) + \left(\frac{1}{sC_t} + \frac{1}{sC_{p\_M3}} \middle\| sC_s \right) + g_{m1} \left(\frac{1}{C_t} + \frac{1}{C_{p\_M3}} \middle\| C_s \right) L_s \qquad (3)$$

where

$$C_t = C_{gs\_M1+C_{ex}+C_1} \qquad (4)$$

and C_{p_M3} in Equation (3) is the parasitic capacitance of M3. In Equation (4), it is not the total capacitance of C1. This expression is only in order to emphasize an existence of the error capacitive elements due to the overlap capacitance of MOS switch M_{cdma} need to be considered. Thus, based on the Equations (2) and (3), the input impedance can be matched by properly tuning the value of the L and C. Actually, as the M_{cdma} connected in series with C1 and M_{wcdma} associated in series between M1 and M3, the parasitic capacitive elements do not affect input matching so much, but these elements should be minimized for noise performance. Thus, it is better to choose a relatively smaller size of switches to reduce high frequency loss through the parasitic capacitor. Although the small device size will increase resistance R_{on}, the parasitic capacitive elements affect on noise far more than resistivity. The final M_{cdma} and M_{wcdma} size are chosen to 60 μm and 40 μm respectively, and simulation result shows a fact that the smaller size the switch transistors have, the better the noise performance is.

For the case when the switch M_{wcdma} is turned on, the transistor M3 converts to parallel with M1, and hence the width of the input transistor increase, lower the frequency to operate at

WCDMA band. It is easy to recognize that more DC current is introduced than the LNA operates at WiBro band because of the increased width of transistor. To solve this problem, shunt grounded capacitor, C_s, plays an important role here. Without this capacitor, M3 must be chosen to a large size enough to shift the frequency from 2.35 GHz to 2.14 GHz. This will eventually cause more power consumption when the switch M_{wcdma} turns on. In addition, as discussed in Equation (3), the parasitic capacitive elements of M3 connected in series with input transistor M1 when the switches are turned off. Thus, larger size of M3 can directly degrade noise figure while the LNA operating at WiBro band. However, with C_s, the width of M3 can be chosen to a relatively smaller size. Simulation results show the DC current consumption is only 10 μA difference between the LNA operating at WiBro band and WCDMA band, and the noise figure at 2.35 GHz is improved by 0.2 dB compared to that without this capacitor.

Because of the extra on-resistance of the switch, the input impedance should be rewritten as

$$Z_{in} = s(L_s + L_g) + \frac{1}{s\left(C_t + C_s + C_{p\_M3}\right)} + \frac{g_{m1}}{C_t + C_s + C_{p\_M3}} L_s + R_{on\_\text{wcdma}} \quad (5)$$

where the parasitic capacitance of M_{wcdma} is neglected and the error capacitive elements generated by M_{cdma} should also be taken into account.

Similarly, when the M_{cdma} is turned on, the operating frequency is shifted to CDMA 2000 1x band depending on the C1, and the input impedance can be expressed as

$$Z_{in} = s(L_s + L_g) + s\left(\frac{1}{C_t + C_s + C_{p\_M3}}\right) + \frac{g_{m1}}{C_t + C_s + C_{p\_M3}} L_s + R_{on\_\text{wcdma}} \| R_{on\_\text{cdma}} \quad (6)$$

From the Equations (3), (5), and (6), the imaginary part can be eliminated by choosing the suitable values of L_g, L_s, and the gate to source capacitance. Because the real part term in Equation (3) is much different from the other two equations, the best input impedance matching performance is designed for WiBro band in this work, and making the other two frequency bands around the 50 Ω matching point on the Smith Chart.

As can be visualized from Figure 3, point 1 which is at 2.35 GHz (WiBro) obtains a good input matching, and point 2 at 2.14 GHz (WCDMA) and point 3 at 1.85 GHz (CDMA 2000 1x) exhibit a similar result of input reflection coefficient corresponding to Equations (5) and (6).
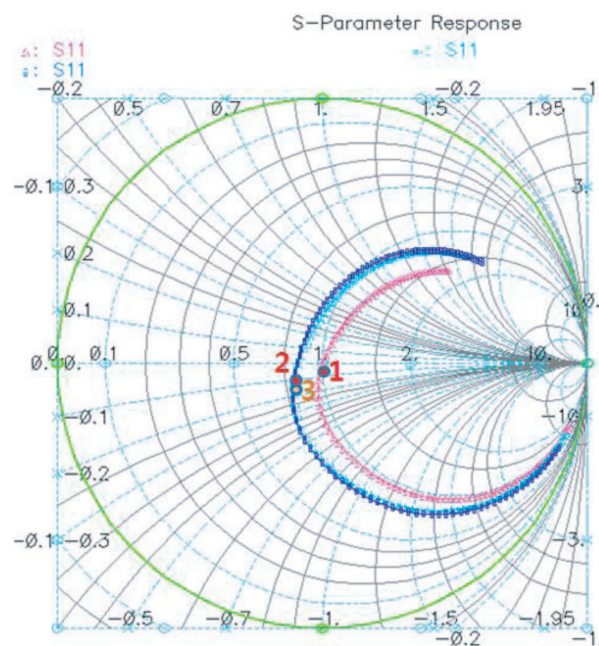

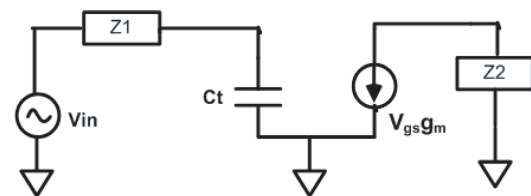Figure 3: Simulation result of S_{11} on the Smith Chart.

Figure 4: Circuit model for nonlinear analysis.

According to [1], the quality factor Q of the input matching network is very important to determine the noise performance. The CMOS RF passive devices, especially for spiral inductor,

however, cannot achieve high enough Q because of the substrate loss. Therefore, it is a big challenge to improve noise performance by using CMOS on-chip inductor for fully integrated LNA design. Based on the characteristic in terms of the four-noise parameters [4], shows that

$$F = F_{\min} + \frac{\lfloor (G_s - G_{opt})^2 + (B_s - B_{opt})^2 \rfloor R_n}{G_s} \tag{7}$$

where the R_n is the effective noise resistance, and G_s and B_s are real and imaginary parts of Z_{in}, respectively, and G_{opt} and B_{opt} are real and imaginary parts of noise impedance, respectively. This equation indicates that the circuit yields the best achievable noise performance when G_s adjusted to G_{opt}. If G_s differs from G_{opt}, its impact on F is amplified by R_n. However, in the triple-mode LNA design, it is difficult to make the input impedance matching and noise matching perfectly at one time. Thus, in this work, the main methodology for optimizing noise performance is to reduce R_n. Also, the four-noise parameters express the effective noise resistance as [4]

$$R_n \approx \frac{\gamma g_{do}}{g_m^2} \tag{8}$$

and

$$R_n \propto 1/W \tag{9}$$

where g_{d0} in Equation (8), is the bias dependence. These two equations imply that the larger device width offers the best chance of lowering noise figure. For a given drain current in this work, lower gate biasing and use a larger device width, the noise figure is eventually less sensitivity to a noise mismatching, and the topology of that set a transistor M3 parallel with M1, is also in order to pursue this idea.

2.3. Output Matching and Voltage Gain

In order to achieve three narrow band voltage gains, there is no any resistive component at the drain terminal. From Figure 1, the switches M_{cdma} and M_{wcdma} also exhibit at output matching network to alter the value of capacitance so as to achieve voltage gain and match the output impedance at different frequency band.

3. TRIPLE-MODE LNA LINEARITY ANALYSIS

In RF circuit design, the linearity is another important issue to consider. Nonlinearity of common source FET amplifier mostly comes from transconductance nonlinearity in the driving MOSFET transistor [2]. Employing Taylor series expansion, the drain current of a common source can be expressed as

$$i_{DS} = I_{dc} + g_m v_{gs} + \frac{g_m'}{2!} v_{gs}^2 + \frac{g_m''}{3!} v_{gs}^3 \tag{10}$$

It is well known that the coefficient of v_{gs}^3 in Equation (10) plays an important role in determining the IMD3 of an RF amplifier [5]. The coefficient g_m'' could be minimized by linearly superposing several CS FET transistors with proper bias and size in parallel [6]. However, this method to improve linearity need to employ larger DC gate bias to cancel the negative peak of the g_m'' causing more power consumption.

For linearity analysis purpose, the equivalent small signal circuit of LNA is depicted in Figure 4. The output impedance seen at the drain of M1, Z_2, is added in the model. In a typical common source FET amplifier, the magnitude of Z_1 is on the order of $1/(sC_{gs})$. Using the volterra series expression, the drain current is given as

$$i_d = A_1(s)v_{in} + A_2(s_1, s_2)v_{in}^2 + A_3(s_1, s_2, s_3)v_{in}^3 \tag{11}$$

Since the inter-modulation products are very close to the fundamental frequency, therefore, $s \approx s_1 \approx s_2$ can be assumed. Based on this assumption, the dependence of IM3 with inverse of the term [7]

$$\left[sC_t R_{in} + g_{m1} \left(1 + \frac{1}{Z_2} \right) \right]^3 \tag{12}$$

where Z_2 can be approximately recognized as $1/g_{m2}$. As can be seen in (12), the linearity can be improved by increasing input transconductance g_{m1}, total gate to source capacitance, and g_{m2}.

In the proposed LNA, capture a relatively low gate biasing and choose large device size for three operating frequency bands, not only the noise performance improved, but linearity enhanced as well. But, it should be note that excessively increase the total gate to source capacitance, both the noise and linearity will be degraded, and as mentioned earlier, if increase the size of transistor M3, the linearity is getting better, but the noise figure will become worse at all of the frequency bands of interest.

4. SIMULATION RESULT

The circuit layout of the triple-mode LNA is displayed in Figure 4. It is implemented by Cadence SpectreRF using TSMC RF 0.18 μm CMOS process.
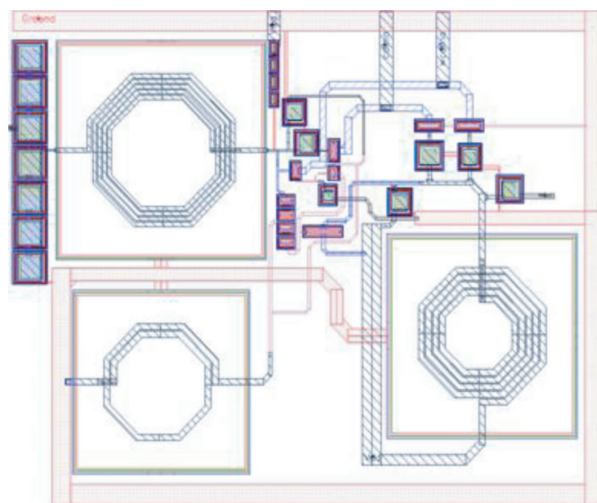
Figure 5: Layout of the proposed triple-mode LNA.

Figure 6 shows the simulation results of input reflection coefficient, S_{11}, and voltage gain, S_{21} operating at three frequency bands, and Figure 7 gives the noise figure of the proposed LNA.
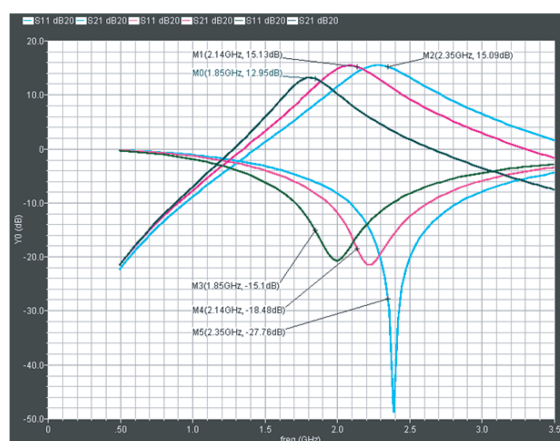

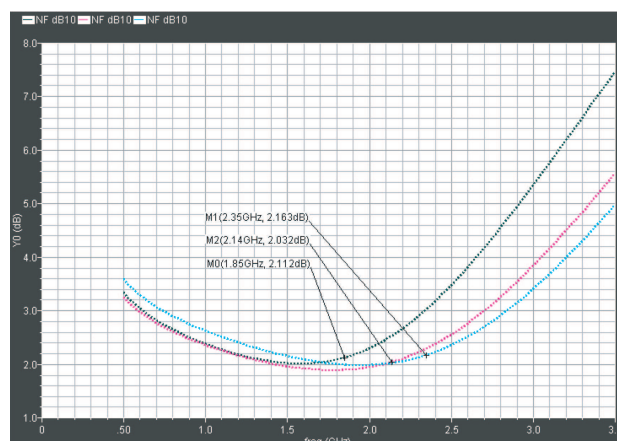Figure 6: Voltage gain and S_{11} in triple-mode CMOS LNA.

Figure 7: Noise figure in triple-mode CMOS LNA.

Simulation results shows 2.11 dB noise figure, −15.1 dB input reflection loss, and 12.9 dB voltage gain at 1.85 GHz for CDMA 2000 1x application; 2.03 dB noise figure, −18.4 dB input reflection loss, and 15.2 dB voltage gain at 2.14 GHz for WCDMA application; 2.16 dB noise figure, −27.6 dB input reflection loss, and 15.1 dB voltage gain at 2.35 GHz for WiBro application.

Figure 7 illustrates third-order inter-modulation, IIP3, at each operating frequency band. Two-tone simulation shows that the IIP3 are −2.7 dBm, −1.8 dBm, and −0.2 dBm at 1.85 GHz, 2.14 GHz, and 2.35 GHz respectively.

With a 1.8 V supply voltage, the circuit consumes 9.4 mW for the total process. The performances of the triple-mode low noise amplifier are summarized in Table 1.
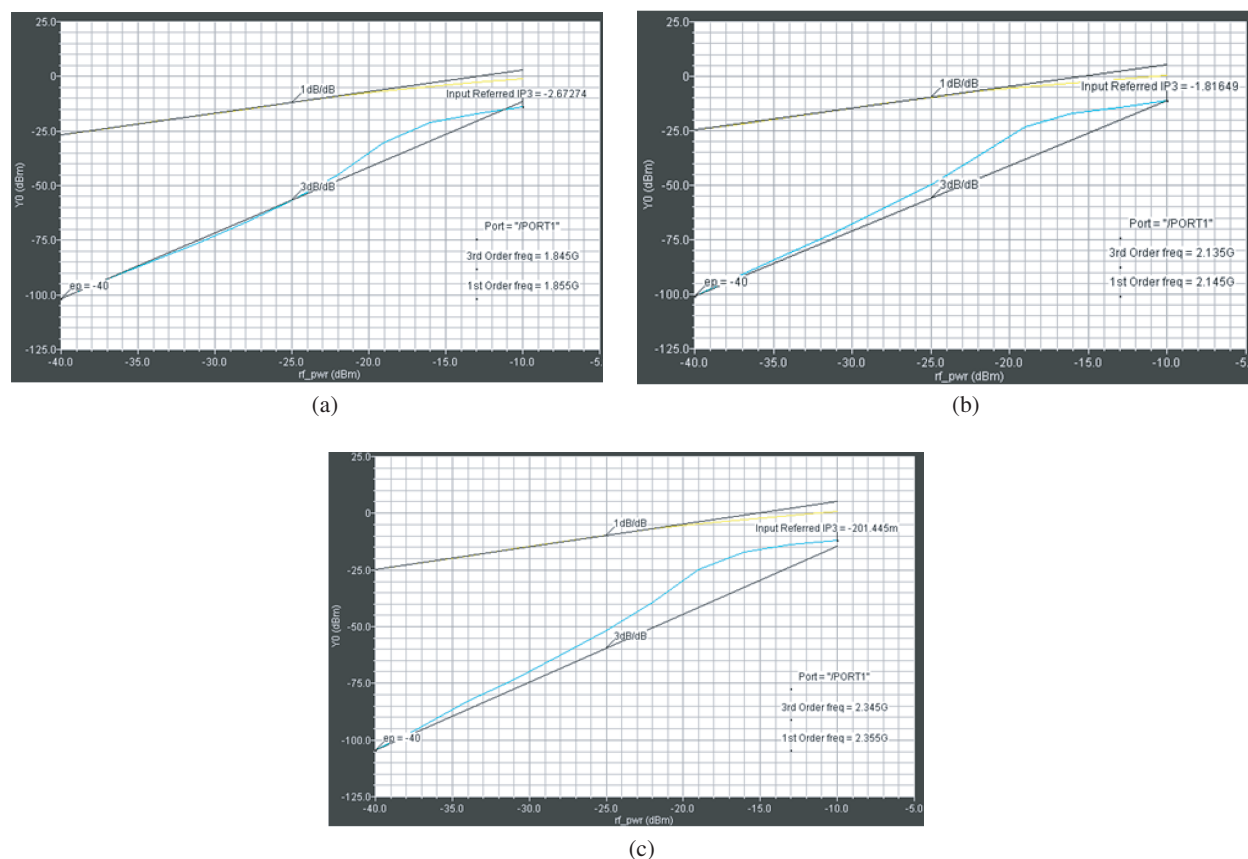
(a)

(b)

(c)

Figure 8: (a) IIP3 at 1.85 GHz; (b) IIP3 at 2.14 GHz; (c) IIP3 at 2.35 GHz.

Table 1.

	CDMA 2000 1x	WCDMA	WiBro
Frequency	1.85 GHz	2.14 GHz	2.35 GHz
Gain (dB)	12.9	15.2	15.1
S_{11} (dB)	−15.1	−18.4	−27.6
S_{22} (dB)	−8.7	−10.7	−12.2
P1dB (dBm)	−15	−16	−15
IIP3 (dBm)	−2.7	−1.8	−0.2
NF (dB)	2.11	2.03	2.16
Power consumption	9.4 mW		

5. CONCLUSION

The fully integrated triple-mode switchable CMOS low noise amplifier using TSMC RF 0.18 μm process is designed at CDMA 2000 1x, WCDMA, WiBro frequency bands for SDR system. With the same low power consumption for the total operating process, the LNA achieved high narrow band gain and input reflection coefficient at the three frequency bands, and also obtained high noise performance with similar noise figure at each frequency band of interest. Based on these advantages, the proposed LNA will be applied with excellent performance for multi-band multi-standard wireless communications.

REFERENCES

1. Shaeffer, D. K., et al., "A 1.5 V, 1.5 GHz CMOS low noise amplifier," *IEEE Journal of Solid-stage Circuits*, Vol. 32, 745–758, May 1997.
2. Razavi, B., *Design of Analog CMOS Integrated Circuits*, Hill Companies, Inc, 2001.
3. Goo, J.-S., "A noise optimization technique for integrated low-noise amplifiers," *IEEE Journal of Solid-state Circuits*, Vol. 37, No. 8, August 2002.
4. Kim, T. W., "Highly linear receiver front-end adopting MOSFET transconductance linearization by multiple gated transistors," *IEEE Journal of Solid-state Circuits*, Vol. 39, No. 1, January 2004.
5. Webster, D., J. Scott, and D. Haigh, "Control of circuit distortion by the derivative superposition method," *IEEE Microwave Guided Wave Lett.*, Vol. 6, 123–125, March 1996.
6. Nguyen, T. K., "CMOS low noise amplifier design optimization technique," *IEEE International Midwest Symposium on Circuits and Systems*, 2004.

Unconditional Stability Criteria for Microwave Networks

Eng Leong Tan[1], **Xiaofeng Sun**[1], **and Kian Sen Ang**[2]

[1]School of EEE, Nanyang Technological University, Singapore
[2]DSO National Laboratories, 20 Science Park Drive, Singapore

Abstract— This paper presents a review of unconditional stability criteria for analysis and design of microwave networks. A linear two-port network is said to be unconditionally stable (in even mode) if for all passive load and source terminations, the moduli of the input and output reflection coefficients are smaller than one. There are many stability criteria stating the necessary and sufficient conditions for such unconditional stability. One classical set of these criteria is based on the combination of Rollett condition ($K > 1$) together with one auxiliary condition (e.g., $|\triangle| < 1$, etc.), which constitutes a family of dual-parameter stability criteria. Another family of stability criteria involves merely a single-parameter condition, which is often more convenient to analyze and utilize. Besides the popular geometrically-derived criterion ($\mu > 1$ or $\mu' > 1$), many other single-parameter criteria can be deduced that feature various specific advantages. One such criterion that is found to be most advantageous is the quasi-invariant single-parameter criterion ($K_t > 1$). With its simplistic form (like K), it can be shown that $K_t > 1$ if and only if a two-port network is unconditionally stable. The parameter K_t can be derived from the classical matrix invariants and the familiar stability criteria, leading to the straightforward proof of its unconditional stability condition. Moreover, it is invariant under arbitrary lossless termination and reactive matching at either port when unconditional stability exists. Such invariant property is retained so long as $|\Delta| \leq 1$, which occurs more frequently than $|\Delta| > 1$ for most microwave devices. The parameter can also be aptly related to the maximum available gain of an unconditionally stable device. Application of the quasi-invariant single-parameter criterion is extended for simplified graphical analysis of linear three-port stability. In particular, only one plot of a single-parameter condition needs to be constructed in each termination reflection plane. Such plot is useful to aptly depict the terminations at each port that result in unconditional stability between the remaining two ports. The three-port stability plots are exemplified to illustrate the convenience using the single-parameter condition. These stability plots find direct application in the stabilization of potentially unstable microwave networks.

1. INTRODUCTION

A linear two-port network is said to be unconditionally stable if the input and output port reflection coefficients are smaller than one ($|\Gamma_{\text{in}}| < 1$ and $|\Gamma_{\text{out}}| < 1$) for all passive source and load impedances [1–3]. There exist necessary and sufficient criteria that can be used to determine such unconditional stability. One classical set of these criteria involves the Rollett's condition, which is [4, 5]

$$K = \frac{1 - |S_{11}|^2 - |S_{22}|^2 + |\Delta|^2}{2\,|S_{12}S_{21}|} > 1 \tag{1}$$

along with one of the following auxiliary conditions:

$$B_1 = 1 + |S_{11}|^2 - |S_{22}|^2 - |\Delta|^2 > 0 \tag{2}$$

$$B_2 = 1 + |S_{22}|^2 - |S_{11}|^2 - |\Delta|^2 > 0 \tag{3}$$

$$\beta_1 = 1 - |S_{22}|^2 - |S_{12}S_{21}| > 0 \tag{4}$$

$$\beta_2 = 1 - |S_{11}|^2 - |S_{12}S_{21}| > 0 \tag{5}$$

$$|\Delta| = |S_{11}S_{22} - S_{12}S_{21}| < 0 \tag{6}$$

Recently, alternative unconditional stability criteria have been proposed that involve only single parameters. The most popular one is the geometrically-derived criterion [6, 7]

$$\mu_1 = \frac{E_1}{|C_1| + |S_{12}S_{21}|} > 1 \tag{7}$$

$$\mu_2 = \frac{E_2}{|C_2| + |S_{12}S_{21}|} > 1 \tag{8}$$

where

$$C_1 = S_{11} - S_{22}^* \Delta \tag{9}$$
$$C_2 = S_{22} - S_{11}^* \Delta \tag{10}$$
$$E_1 = 1 - |S_{22}|^2 \tag{11}$$
$$E_2 = 1 - |S_{11}|^2. \tag{12}$$

Either μ_1 or μ_2 alone will be able to determine whether a two-port network is unconditionally stable or potentially unstable.

This paper reviews various alternative single parameters that are useful to determine the unconditional stability of linear two-ports. All the single parameters considered will be necessary and sufficient to determine if a two-port network is unconditionally stable. Moreover, we highlight an alternative criterion involving a quasi-invariant single parameter which exploits many advantages associated with the K factor. It should be noted that all the unconditional stability criteria concern only the stability of even modes (assumed henceforth). More rigorous statements of the criteria imposing Rollett proviso should be called upon to ascertain the absence of oscillations [5].

2. ROLLETT-BASED SINGLE-PARAMETER CRITERIA

In this section, we present the single-parameter criteria based on the dual Rollett and auxiliary conditions. Combining the K factor and its auxiliary condition given above in terms of B_1, B_2, β_1, β_2 or $|\Delta|$, several single parameters can be obtained as [8]

$$K_{B1} \equiv \frac{E_1}{|\delta_1| + |S_{12}S_{21}|} > 1, \quad K'_{B1} \equiv \frac{E_1 - |\delta_1|}{|S_{12}S_{21}|} > 1 \tag{13}$$

$$K_{B2} \equiv \frac{E_2}{|\delta_2| + |S_{12}S_{21}|} > 1, \quad K'_{B2} \equiv \frac{E_2 - |\delta_2|}{|S_{12}S_{21}|} > 1 \tag{14}$$

$$K_{\beta1} \equiv \frac{2E_1 - \delta_1}{|\delta_1| + 2|S_{12}S_{21}|} > 1, \quad K'_{\beta1} \equiv \frac{2E_1 - \delta_1 - |\delta_1|}{2|S_{12}S_{21}|} > 1 \tag{15}$$

$$K_{\beta2} \equiv \frac{2E_2 - \delta_2}{|\delta_2| + 2|S_{12}S_{21}|} > 1, \quad K'_{\beta2} \equiv \frac{2E_2 - \delta_2 - |\delta_2|}{2|S_{12}S_{21}|} > 1 \tag{16}$$

$$K_\Delta \equiv \frac{E_1 + E_2}{|\delta_1 + \delta_2| + 2|S_{12}S_{21}|} > 1, \quad K'_\Delta \equiv \frac{E_1 + E_2 - |\delta_1 + \delta_2|}{2|S_{12}S_{21}|} > 1 \tag{17}$$

where

$$\delta_1 = |S_{11}|^2 - |\Delta|^2 + |S_{12}S_{21}| \tag{18}$$
$$\delta_2 = |S_{22}|^2 - |\Delta|^2 + |S_{12}S_{21}|. \tag{19}$$

The (unprimed) parameters given in the left of (13)–(17) have been defined originally in [8]. They may be redefined to resemble the K factor more as given (primed) in the right of (13)–(17).

For more generalized auxiliary conditions of $|\Gamma_{\text{out}}| < 1$ for any $\Gamma_S = e^{j\phi_1}$ or $|\Gamma_{\text{in}}| < 1$ for any $\Gamma_L = e^{j\phi_2}$ (arbitrary real ϕ_1 and ϕ_2), the corresponding single parameters can be expressed as

$$K_{\Gamma1} \equiv \frac{E_1 - \text{Re}\left[C_1 e^{j\phi_1}\right]}{|\delta_1 - \text{Re}\left[C_1 e^{j\phi_1}\right]| + |S_{12}S_{21}|} > 1, \quad K'_{\Gamma1} \equiv \frac{E_1 - \text{Re}\left[C_1 e^{j\phi_1}\right] - |\delta_1 - \text{Re}\left[C_1 e^{j\phi_1}\right]|}{|S_{12}S_{21}|} > 1 \tag{20}$$

$$K_{\Gamma2} \equiv \frac{E_2 - \text{Re}\left[C_2 e^{j\phi_2}\right]}{|\delta_2 - \text{Re}\left[C_2 e^{j\phi_2}\right]| + |S_{12}S_{21}|} > 1, \quad K'_{\Gamma2} \equiv \frac{E_2 - \text{Re}\left[C_2 e^{j\phi_2}\right] - |\delta_2 - \text{Re}\left[C_2 e^{j\phi_2}\right]|}{|S_{12}S_{21}|} > 1. \tag{21}$$

These parameters can be reduced further for two special cases of (20) and (21), namely short- and open-circuit terminations. In particular, we have for short circuit ($\phi_1 = \pi$ and $\phi_2 = \pi$),

$$K_{Y1} \equiv \frac{E_1 + \text{Re}\left[C_1\right]}{|\delta_1 + \text{Re}\left[C_1\right]| + |S_{12}S_{21}|} > 1, \quad K'_{Y1} \equiv \frac{E_1 + \text{Re}\left[C_1\right] - |\delta_1 + \text{Re}\left[C_1\right]|}{|S_{12}S_{21}|} > 1 \tag{22}$$

$$K_{Y2} \equiv \frac{E_2 + \text{Re}\left[C_2\right]}{|\delta_2 + \text{Re}\left[C_2\right]| + |S_{12}S_{21}|} > 1, \quad K'_{Y2} \equiv \frac{E_2 + \text{Re}\left[C_2\right] - |\delta_2 + \text{Re}\left[C_2\right]|}{|S_{12}S_{21}|} > 1 \tag{23}$$

and for open circuit ($\phi_1 = 0$ and $\phi_2 = 0$),

$$K_{Z1} \equiv \frac{E_1 - \mathrm{Re}\,[C_1]}{|\delta_1 - \mathrm{Re}\,[C_1]| + |S_{12}S_{21}|} > 1, \quad K'_{Z1} \equiv \frac{E_1 - \mathrm{Re}\,[C_1] - |\delta_1 - \mathrm{Re}\,[C_1]|}{|S_{12}S_{21}|} > 1 \tag{24}$$

$$K_{Z2} \equiv \frac{E_2 - \mathrm{Re}\,[C_2]}{|\delta_2 - \mathrm{Re}\,[C_2]| + |S_{12}S_{21}|} > 1, \quad K'_{Z2} \equiv \frac{E_2 - \mathrm{Re}\,[C_2] - |\delta_2 - \mathrm{Re}\,[C_2]|}{|S_{12}S_{21}|} > 1 \tag{25}$$

All the subscripted K and K' parameters above constitute the family of Rollett-based single-parameter criteria for unconditional stability. This fairly large family provides us much room to identify and exploit the potential advantages of certain members in various aspects. For instance, the (unprimed) parameters in (22)–(25) have been demonstrated to be invariant under lossless shunt and series terminations respectively [8].

3. QUASI-INVARIANT SINGLE-PARAMETER CRITERION

Here we present the quasi-invariant single-parameter criterion that involves a symmetrical parameter defined analogous to (1) as [9]

$$K_t \equiv \frac{3 - 2\,|S_{11}|^2 - 2\,|S_{22}|^2 + |\Delta|^2 - \left|1 - |\Delta|^2\right|}{4\,|S_{12}S_{21}|}. \tag{26}$$

This criterion is necessary and sufficient to determine the unconditional stability of a linear two-port network, i.e., $K_t > 1$ if and only if a two-port network is unconditionally stable. Subsequently, the parameter is derived from the classical matrix invariants and the familiar stability criteria, leading to the straightforward proof of its unconditional stability condition [10]. In particular, the trace of certain dissipation matrix [10] (expressible in terms of β_1 and β_2) can be an auxiliary condition alternative to (4) or (5):

$$\beta_1 + \beta_2 > 0. \tag{27}$$

Combining this auxiliary condition (27) with the Rollett condition (1) then gives rise to the quasi-invariant single-parameter criterion involving (26) directly.

The quasi-invariant parameter is invariant under arbitrary lossless termination and reactive matching at either port when unconditional stability exists, or as long as $|\Delta| \leq 1$ which occurs more frequently than $|\Delta| > 1$ for most microwave devices. Furthermore, since K_t coincides with K when $K_t > 1$, many of the desirable properties inherent in K can be exploited. Indeed, we find that K_t is related to the maximum available gain (MAG) more aptly than K,

$$\mathrm{MAG} = \frac{|S_{21}|}{|S_{12}|\left(K_t + \sqrt{K_t^2 - 1}\right)} \quad \text{with} \quad K_t > 1. \tag{28}$$

This expression in terms of K_t preserves the convenience for analyzing trade-offs between gain and stability when $K_t > 1$.

Table 1: S parameter data sets, $|\Delta|$, K, K_t and MAG.

Set	S_{11}	S_{12}	S_{21}	S_{22}	$|\Delta|$	K	K_t	MAG
1	$0.2\angle 120°$	$0.05\angle 120°$	$3\angle 40°$	$0.5\angle -50°$	0.24909	2.573	2.573	12.13
2	$0.75\angle -60°$	$0.3\angle 70°$	$6\angle 90°$	$0.5\angle 60°$	2.1562	1.344	0.3299	-
3	$1.05\angle 20°$	$0.05\angle 120°$	$3\angle 40°$	$0.5\angle -50°$	0.67323	0.3358	0.3358	-
4	$0.1\angle 0°$	$0\angle 0°$	$0\angle 0°$	$0.3\angle 0°$	0.03	∞	∞	0
5	$1.2\angle 0°$	$0\angle 0°$	$0\angle 0°$	$0.3\angle 0°$	0.36	$-\infty$	$-\infty$	-
6	$0.1\angle 0°$	$0\angle 0°$	$0\angle 0°$	$1.3\angle 0°$	0.13	$-\infty$	$-\infty$	-
7	$0.5\angle 0°$	$0.025\angle 180°$	$2\angle 0°$	$0.1\angle 0°$	0.1	7.5	7.5	5.357
8	$0.95\angle -22°$	$0.04\angle 80°$	$3.5\angle 165°$	$0.61\angle -13°$	0.57205	0.188	0.188	-
9	$0.69\angle -123°$	$0.11\angle 48°$	$1.29\angle 78°$	$0.52\angle -77°$	0.25388	1.12	1.12	7.215

To compare various single parameters for unconditional stability, let us revisit the examples in [8, 9] whose S parameters are listed in Table 1 along with the values of $|\Delta|$, K, K_t and MAG. The gain values are included for reference by noting that K_t is related simply to MAG when $K_t > 1$. Table 2 lists the values of new subscripted-K' parameters, which signify unconditional stability if and only if each parameter value is greater than 1. While one can show that K_t is invariant under arbitrary lossless termination so long as $|\Delta| \leq 1$ [9], it is not certain yet about the invariant properties of these subscripted-K' parameters.

Table 2: Comparison of various single parameters for unconditional stability.

Set	K'_{B1}	K'_{B2}	$K'_{\beta 1}$	$K'_{\beta 2}$	K'_{Δ}	K'_{Y1}	K'_{Y2}	K'_{Z1}	K'_{Z2}
1	4.147	4.147	4.147	4.147	4.147	4.147	4.147	4.147	4.147
2	-0.8537	-1.201	0.4167	0.2431	-1.027	-0.7149	-2.228	-0.9926	-0.1734
3	-0.3285	-1.038	-0.3285	-0.6833	-0.3285	-0.3285	-3.086	-0.3285	-0.3285
4	∞	∞	∞	∞	∞	∞	∞	∞	∞
5	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$
6	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$
7	14	14	14	14	14	14	14	6	14
8	-0.624	-0.624	-0.624	-0.624	-0.624	-0.624	-0.624	-0.624	-0.624
9	1.241	1.241	1.241	1.241	1.241	1.241	1.241	1.241	1.241

Apart from being convenient for linear two port stability investigation, the single-parameter criterion can also be useful for linear three-port stability analysis [10]. Such analysis considers the stability between two ports with a fixed termination at the third port. Assume that Γ_3 is the corresponding reflection coefficient when a three-port is terminated at port 3. Due to the close relation between power gain and stability as dictated by (28), it is more advantageous to adopt the symmetrical quasi-invariant parameter $K_{t3}(\Gamma_3)$ rather than the asymmetrical geometrical stability parameter $\mu_3(\Gamma_3)$ [11]. The three-port stability condition reads

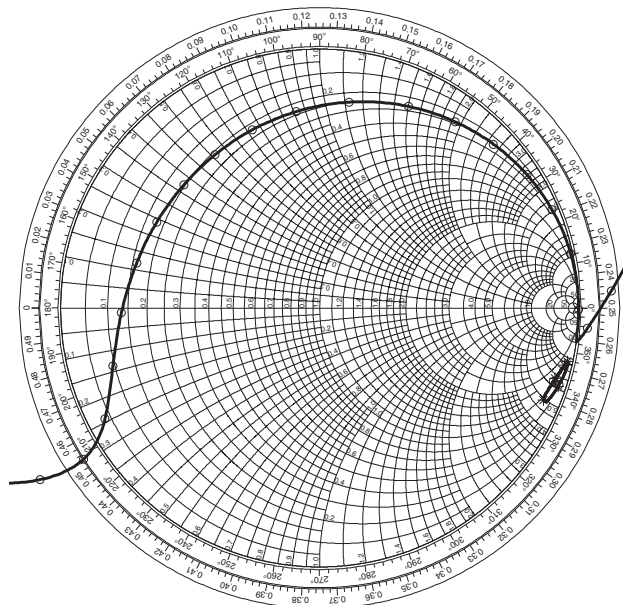$$K_{t3}(\Gamma_3) \equiv \frac{N(\Gamma_3)}{D(\Gamma_3)} > 1 \tag{29}$$

Figure 1: Stability plots viewed in the Γ_3 plane. The interiors of the plots marked with 'o' and '∗' represent the regions for $K_3(\Gamma_3) > 1$ and $K_{t3}(\Gamma_3) > 1$ respectively.

where

$$N(\Gamma_3) = 3\left|1 - S_{33}\Gamma_3\right|^2 - 2\left|S_{11} - \Delta_{22}\Gamma_3\right|^2 - 2\left|S_{22} - \Delta_{11}\Gamma_3\right|^2$$
$$+ \left|\Delta_{33} - \Delta_3\Gamma_3\right|^2 - \left|\left|1 - S_{33}\Gamma_3\right|^2 - \left|\Delta_{33} - \Delta_3\Gamma_3\right|^2\right|, \tag{30}$$

$$D(\Gamma_3) = 4\left|(S_{12} + \Delta_{21}\Gamma_3)(S_{21} + \Delta_{12}\Gamma_3)\right|, \tag{31}$$

Δ_{ij} is the cofactor of (i, j)-th element and Δ_3 is the determinant of three-port S-parameter matrix. To illustrate the three-port stability plot, consider a three port network with S-parameters as

$$\begin{bmatrix} 1.012\angle-31.79° & 0.385\angle65.21° & 0.185\angle96.79° \\ 1.720\angle139.77° & 0.703\angle-36.28° & 1.879\angle-21.70° \\ 1.563\angle-21.69° & 0.280\angle13.81° & 0.886\angle144.81° \end{bmatrix}. \tag{32}$$

Figure 1 shows the plots of $K_3(\Gamma_3) > 1$ and $K_{t3}(\Gamma_3) > 1$ in the Γ_3 plane. Observe that the large (upper) region of $K_3(\Gamma_3) > 1$ does not meet the unconditional stability condition. Using the single-parameter condition $K_{t3}(\Gamma_3) > 1$ enables one to single out the proper region directly. Such plot provides a quick and apt graphical representation of the Γ_3 terminations that result in unconditional stability between ports 1 and 2.

4. CONCLUSION

This paper has presented a review of single-parameter criteria that are necessary and sufficient to determine the unconditional stability of linear two-port networks. Many of these criteria can be derived based on the dual Rollett and auxiliary conditions. Moreover, an alternative criterion involving a quasi-invariant single parameter K_t has been highlighted which exploits many advantages associated with the Rollett K factor. The parameter is invariant under arbitrary lossless termination and reactive matching at either port when unconditional stability exists, or as long as $|\Delta| \leq 1$. It can also be aptly related to the maximum available gain of an unconditionally stable device. Application of the quasi-invariant single-parameter criterion has been extended for simplified graphical analysis of linear three-port stability. The three-port stability plots find direct application in the stabilization of potentially unstable microwave networks.

REFERENCES

1. Ha, T. T., *Solid State Microwave Amplifier Design*, Wiley, New York, 1981.
2. Vendelin, G. D., A. M. Pavio, and U. L. Rohde, *Microwave Circuit Design using Linear and Nonlinear Techniques*, Wiley, New York, 1990.
3. Gonzalez, G., *Microwave Transistor Amplifiers: Analysis and Design*, 2nd. ed., Upper Saddle River, Prentice Hall, NJ, 1997.
4. Rollett, J. M., "Stability and power gain invariants of linear two-ports," *IRE Trans. Circuit Theory*, Vol. 9, No. 3, 29–32, Mar. 1962.
5. Woods, D., "Reappraisal of the unconditional stability criteria for active 2-port networks in terms of S parameters," *IEEE Trans. Circuits Syst.*, Vol. 23, No. 2, 73–81, Feb. 1976.
6. Edwards, M. L. and J. H. Sinsky, "A new criterion for linear 2-port stability using a single geometrically derived parameter," *IEEE Trans. Microwave Theory Tech.*, Vol. 40, No. 12, 2303–2311, Dec. 1992.
7. Tan, E. L., "Simple derivation and proof of geometrical stability criteria for linear two-ports," *Microwave Opt. Technol. Lett.*, Vol. 40, No. 1, 81–83, Jan. 2004.
8. Tan, E. L., "Rollett-based single-parameter criteria for unconditional stability of linear two-ports," *IEE Proc. — Microw. Antennas Propag.*, Vol. 151, No. 4, 299–302, Aug. 2004.
9. Tan, E. L., "A quasi-invariant single-parameter criterion for linear two-port unconditional stability," *IEEE Microw. Wireless Comp. Lett.*, Vol. 14, No. 10, 487–489, Oct. 2004.
10. Tan, E. L., "Quasi-invariant single-parameter criterion for unconditional stability: Review and application," *2006 Asia-Pacific Microwave Conference*, 429–432, Dec. 2006.
11. Tan, E. L., "Simplified graphical analysis of linear three-port stability," *IEE Proc. — Microw. Antennas Propag.*, Vol. 152, No. 4, 209–213, Aug. 2005.

A Low Cost and High Performance Design of X-band Short Range Doppler Radar Transceiver

Jingzhou Luo and Lixin Ran

Information Science and Engineering College, Zhejiang University, China

Abstract— This paper presents a systematic integration and circuits design scheme of X-band Doppler radar for short-range use purpose. With the help of powerful microwave design tool ADS, the authors designed a complete CW Doppler radar transceiver and made a test for verification. Firstly, we establish a system model by conventional radar equation. Secondly, design the schematics of main modules including an oscillator, a mixer, and antennas. Finally, perform system integration using the designed circuits diagrams and parameters. The performance is found fairly satisfactory by test verification.

1. INTRODUCTION

Doppler radars play an important part in many industrial, scientific, and medical areas. Recently, more and more Doppler radar products are especially found in some commercial short-range use circumstances like check on work attendance system, family guard against theft and alarm system, and human body's vital sign detection system, etc. The features of these Doppler radar products are with high sensitivity of target detection; low power radiation to reduce harm for human body; low cost to facilitate wide applications; easy of circuits alignment to fit the batch production; and small circuit size, etc. Based on this consideration, the authors made a CAD for a continuous wave (CW) Doppler radar transceiver. The radar works at X band. It has an optimized system performance in virtue of ADS simulation techniques. Because of using microstrip circuit and microwave integrated circuit techniques, the transceiver PCB size can be less than $40 \times 40 \, \text{mm}^2$.

In the following, we firstly establish a system model by conventional the radar equation, and then design module schematics including an oscillator, a mixer, and TX/RX antennas. Finally, we perform system integration and make a simple Doppler experiment to verify the design.

2. RADAR TRANSCEIVER SYSTEM DESIGN

The diagram of designed radar transceiver is shown in the left part of Figure 1. It mainly consists of an oscillator, a mixer, and two antennas, also including other parts like a RF coupler and matchers, a base band filter and amplifier, etc. The right part is a radar target model. Similarly refer to Reference [1], we set up the parameters of target model including frequency sources and a phase-shifter relative to Doppler radar application except for the attenuator, as the latter represents the loss in free space and radar cross section (RCS).

Notice that the most important parameters of the transceiver are oscillator output power P_t, transmitting and receiving antenna gain G_t and G_r, and mixer conversion gain (or loss) G_m. Of cause this depends upon application purpose. Since our Doppler radar is limited in the case of short-range target detection, we use neither power amplifier nor receiving low noise amplifier. We start from the classic radar equation [2]:

$$P_r = \frac{P_t G_t G_r \lambda^2 S}{(4\pi)^3 R^4}. \tag{1}$$

where λ is the wavelength of the RF signal; S is the RCS of the target; R is the line of sight distance from radar to target. Considering the dielectric and polarization properties, the RCS should be revised to

$$S = r\xi\eta S_1. \tag{2}$$

where r, ξ and η represent target backward power scattering loss, power absorption loss, and wave polarization loss, respectively. S_1 is the RCS in ideality (target loss = 0), which depends on geometry of the targets. In the case of Doppler radar, S_1 includes the information of the target. Referring to (1) and (2), and the characters of the signal generated by a double sideband carrier suppression AM modulator, i.e., a direct multiplier of two signal, the total loss of the attenuator can be written as

$$L = \frac{\lambda^2 S}{(4\pi)^3 R^4}, \tag{3}$$

Thus, we need to determine the circuit specifications P_t, G_m, G_t and G_r, which can be done by ADS system simulation.

3. CIRCUITS DESIGN

3.1. Oscillator Design

Comparing to other microwave oscillators, Dielectric Resonator Oscillator (DRO) has many advantages such as high Q factor, small size, and low cost, etc. Although its phase noise performance is not very satisfactory, it can still meet the requirement of Doppler radar applications [3, 4]. In fact, through simulation described in Figure 1, we can show that even if the phase noise of oscillator reaches the upper value ($-10\,$dBc/Hz), the Doppler signal still can be demodulated satisfactorily. So choosing DRO as the design scheme is reasonable.
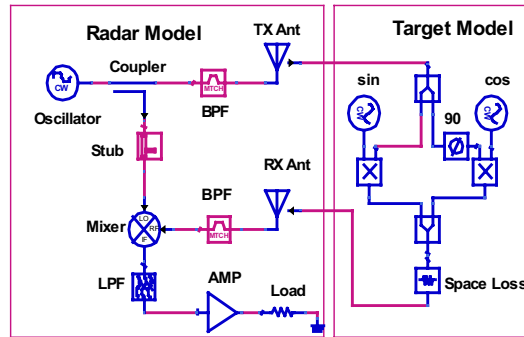
Figure 1: Radar transceiver and target model.

Figure 2(a) gives a schematic level of the simulation. We choose GaAs MESFET NE71084 as the main oscillation component. The circuit is of parallel feedback oscillation type. The feed back capacitance can be obtained by the coupling between gate and drain microstrip transmission line. Figure 2(b) gives Harmonic Balance simulation result. From the figure we can see the oscillator output power at 10.07 GHz is near 12 dBm.

(a) Circuit model   (b) Harmonic balance simulation result
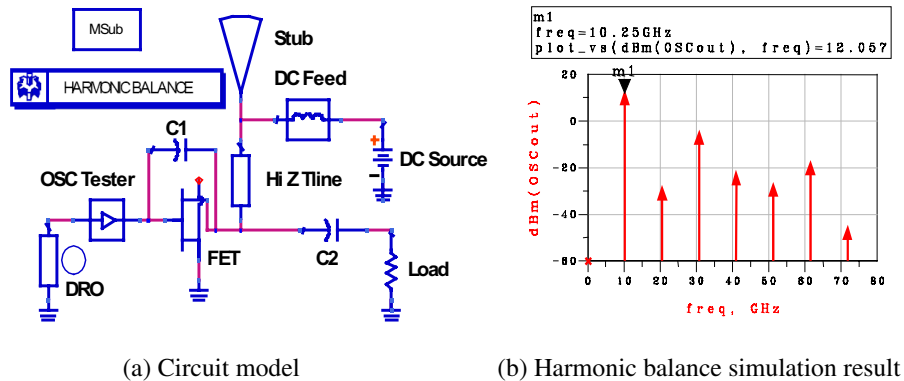
Figure 2: Oscillator simulations.

3.2. Mixer Design

As we know a double balanced diode mixer has 5~7 dB loss, and it also needs large physical size, because it uses a microstrip coupler. While FET mixer has less combine harmonics and higher output power saturation point. Figure 3(a) shows our mixer circuit design schematic. To be different from a normally FET mixer, the RF input signal is not added at the gate but at the drain. This design also poses the advantages of decreasing PCB size and improving mixer performance. Besides, we use a 5 k-ohm resistor as output load, since for most operational amplifier the input impedance must be high.

By measuring and optimizing the level difference between the baseband output and RF input we can get conversion gain G_m of 13.6 dB at 10.07 GHz. This is a very good result. To show this process the input RF and demodulated IF spectrum are shown in Figures 3(b) and 3(c). Notice

that in these two figures the frequency coordinate origin has deferent meaning: in 3(b) it is the oscillation frequency and in 3(c) it is the DC component.
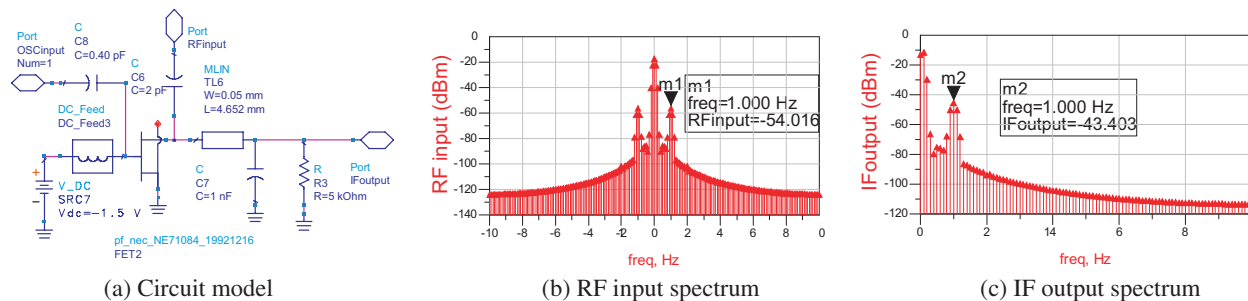

(a) Circuit model   (b) RF input spectrum   (c) IF output spectrum

Figure 3: Mixer circuit simulations.

3.3. Antenna Design

The key point for design of antennas is also based on improving antenna gain. As a compromise consideration between performance and circuit complexity, a two-element microstrip array is chosen. The procedure for this design is very typical: firstly design a single patch element and a matching circuit of transmission line, and then combine two elements into an array. The layout parameters are got by ordinary way, but unlike off center feeding, the feeding point is chosen at the middle of the patch wide side. This is for convenience of establishing phase centre. The simulation results show this way can also lead a good performance.

Figure 4 shows the antenna gain and input impedance characteristics. We can see the gain at working frequency $10.07\,\mathrm{GHz}$ is above $8\,\mathrm{dB}$, and S_{11} is less than $-32\,\mathrm{dB}$. This is a satisfactory result.

(a) Directivity and gain characteristics @ 10.07GHz   (b) Impedance characteristics
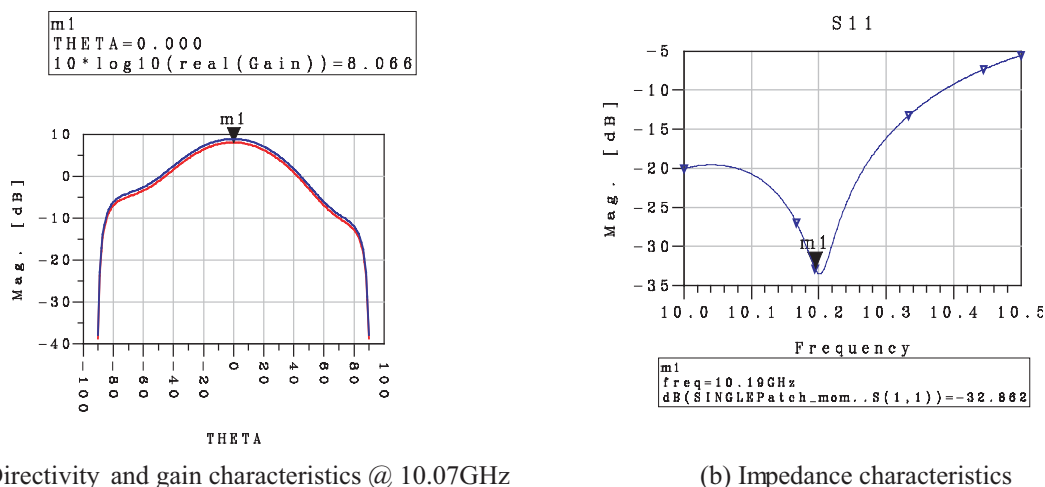
Figure 4: Antenna optimizations.

4. SYSTEM SIMULATIONS AND VERIFICATION

To check the design a system level simulation is done and compared to a test. The designed circuits of the oscillator, mixer, and antenna have been integrated into a FR-4 type PCB with size of $40 \times 40\,\mathrm{mm}^2$ shown in the right part of Figure 5. The parameters for simulation are set to following: $R = 1\,\mathrm{m}$; Doppler frequency shift $= 1\,\mathrm{Hz}$; $P_t = 10\,\mathrm{mW}$; $G_t = G_r = 8\,\mathrm{dB}$; $S = S_1$ (considering no target loss case); the moving target is a metal plate of $200 \times 200\,\mathrm{mm}^2$. We substitute these schematics and values into the system in Figure 1 and get the waveform with about $4\,\mathrm{mV}$ peak-to-peak values as shown in Figure 6(a). This signal amplitude is enough for signal to noise ratio (SNR) requirement of a low pass filter amplifier.

For verification purpose we also made an experiment by a test board composed of the transceiver daughter board and a self made digital signal processing (DSP) main board shown in Figure 5. The

latter is the main hardware to be used for low pass filtering, pre-amplifying (with a 20 dB LPF amplifier), data sampling, memorizing, system controlling and processing, on which a TMS320 UC5402 DSP chip is used as the central controller and processor. However, for our comparison purpose here we need only the sampled output baseband data signal. This measured signal result is shown on Figure 6(b). We can see the measured Doppler waveforms agree fairly with the simulation in fundamental frequency and amplitude. Notice that in the simulation no amplifier is accounted, so the measured signal should be 20 dB larger than simulated one. For length limit we omit the details of the test procedure here.
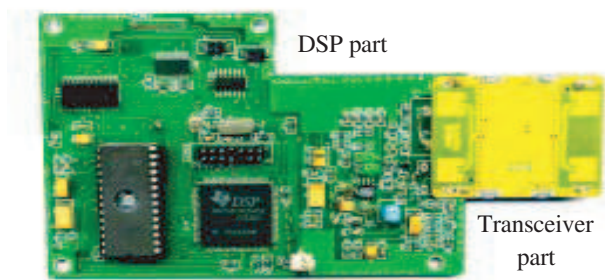
Figure 5: Test board for design verification.

(a) Simulation signal
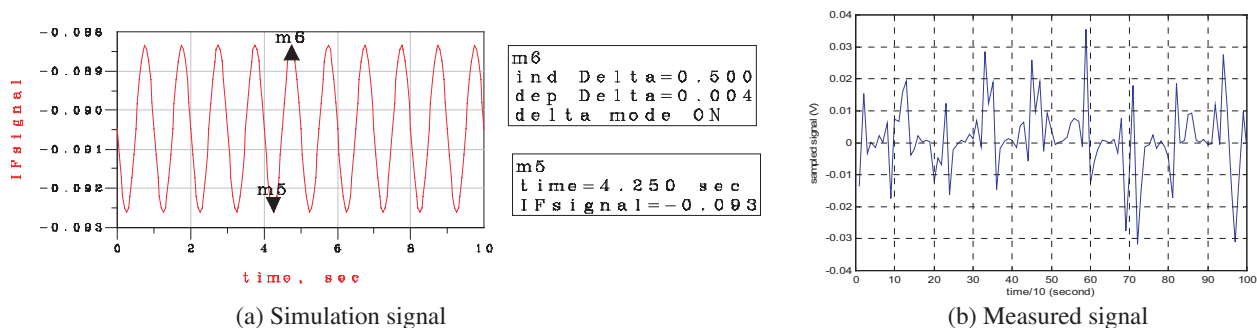
(b) Measured signal

Figure 6: Simulation and measurement results of baseband signal.

In fact, the transceiver is worked with high sensitivity of target detection. The maximum range of large object detection like a moving human body can reach up to 10 meters, provided that appropriate adjusting the PCB layout and optimizing the components of the circuit is carried out.

5. CONCLUSIONS

The authors have presented a system integration and circuit design scheme of X-band Doppler radar for short-range use purpose. The system is modeled by conventional Doppler radar theory and designed by ADS. Typical designed circuit specifications are achieved: the oscillator output is 12 dBm, the mixer conversion gain is 13 dB, and the antenna gain is 8 dB. The maximum range of large object detection can reach up to 10 meters. Simulations and experiments show this design scheme can meet our application requirement.

REFERENCES

1. ADS Documentation, Agilent Technologies, 2003A.

2. Skolnik, M. I., *Introduction to Radar Systems*, McGraw-Hill Book Company Inc., 1962.

3. Boric-Lubecke, O., et al., "Wireless IC Doppler radars for sensing of heart and respiration activity," *6th International Conference on Telecommunications in Modern Satellite, Cable and Broadcasting Service*, Vol. 1, 337–344, Oct. 2003.

4. Droitcour, A. D., O. B. Lubccke, V. M. Lubccke, and J. Lin, "0.25 μm CMOS and BiCMOS singlechip direct-conversion Doppler radars for remote sensing of vital signs," *ISSCC Digest*, Vol. 1, 348–349, 2002.

Photonic Crystal Based Subwavelength Imaging and Cloaking Optical Devices

O. Vanbesien[1], **N. Fabre**[1], **X. Mélique**[1],
L. Lalouat[2], **B. Cluzel**[2], **F. de Fornel**[2], **and D. Lippens**[1]

[1]Institut d'Electronique, de Microélectronique et de Nanotechnologie (IEMN — UMR CNRS 8520)
Université des Sciences et Technologies de Lille
Avenue Poincaré, BP 60069, 59655 Villeneuve d'Ascq Cedex, France
[2]Institut Carnot de Bourgogne (ICB — UMR CNRS 5209)
9, Avenue A. Savary, BP 47870, 21078 Dijon, France

Abstract— Potentialities of 2.5D dielectric photonic crystals for subwavelength imaging or invisibility cloaking operation at optical wavelengths are studied numerically and experimentally. First, a negative refraction based flat lens is designed, fabricated and characterized. Using scanning near-field optical microscopy, subwavelength imaging (0.8λ) for a deduced refractive index of -1 is obtained at $1.55\,\mu\text{m}$. Second, an original design aimed to operate as a cloaking device is proposed by combining photonic lattices operating under different regimes, exploiting both pass- and stop-band properties.

1. INTRODUCTION

Photonic crystals have been widely studied since two decades for their outstanding potentialities to control the motion of light [1]. Formerly exploited for their band gap properties associated with point or linear defects for wave guiding or multiplexing, they found a second breathe when metamaterials became of prime concern in the electromagnetic research community [2]. The concept of negative refraction, based in the lower part of the spectrum on permittivity and permeability engineering [3, 4], can also be evoked in the case of photonic crystal as pass-bands are considered. Indeed, a band slope reversal with respect to the wave vector, if isotropic, can be interpreted as a negative refractive index. Compared to metal-dielectric metamaterials, full dielectric crystals appear more mature for the optical wavelength range owing to standard technological processes of nanoelectronics for their fabrication [5–8]. Conversely, less versatility will be possible in terms of subwavelength patterning since dispersion properties in such crystals are fixed by the geometrical parameters, namely the periodicity and the filling factor.

To quantify the potentialities of 2.5D photonic crystals, we will investigate first the subwavelength imaging by a $n = -1$ flat lens [9], with special attention to the wave matching [10], and second we will propose an original design for a cloaking device by combining photonic lattices operating under pass- and stop-band conditions [11].

2. SUBWAVELENGTH IMAGING BY A $n = -1$ PHOTONIC CRYSTAL BASED FLAT LENS

To design a photonic crystal (PC) based $n = -1$ flat lens, we start from a two dimensional triangular air hole lattice patterned in a semiconductor matrix. The refractive index of this matrix corresponds to the first guided optical mode confined in a vertical InP/GaInAsP/InP heterostructure for a wavelength close to $1.55\,\mu\text{m}$. The PC geometrical parameters are fixed using an optimization procedure we develop base on the exploitation of the cavity modes defined by the finite lens surrounded with air. As we want to demonstrate the focusing properties of such a lens illuminated by a quasi-point source, this procedure allow us to enhance the lens transmission for various incidence angles. A typical band structure is given in Figure 1(a) which corresponds to a filling factor of 38%. The second TE band shows a negative slope and crosses the light line for $a/\lambda = 0.31$ (a is the period of the crystal and λ the wavelength in air). As shown in Figure 1(b), the corresponding iso-frequency curve is circular leading to a refractive index value of -1. Moving to a real prototype, the third dimension of space has to be considered. As illustrated by the schematic side view of Figure 2(c), the PC has a finite depth of about $2\,\mu\text{m}$ and a quasi point source is defined using a ridge waveguide ($0.6\,\mu\text{m}$ width) to illuminate the lens. The source/lens distance is half the lens thickness. A three-dimensional FDTD calculation at $\lambda = 1.55\,\mu\text{m}$ has been conducted and Figures 2(a) and 2(b) gives the E-field density for two specific planes corresponding to a top view at heterostructure level and

a side view at source location. These two plots show that the focusing of light is present even if a large reflection dominates at the first interface. Moreover, the spot location behind tends to show a shift of the refractive index value from −1 when the third dimension is included.
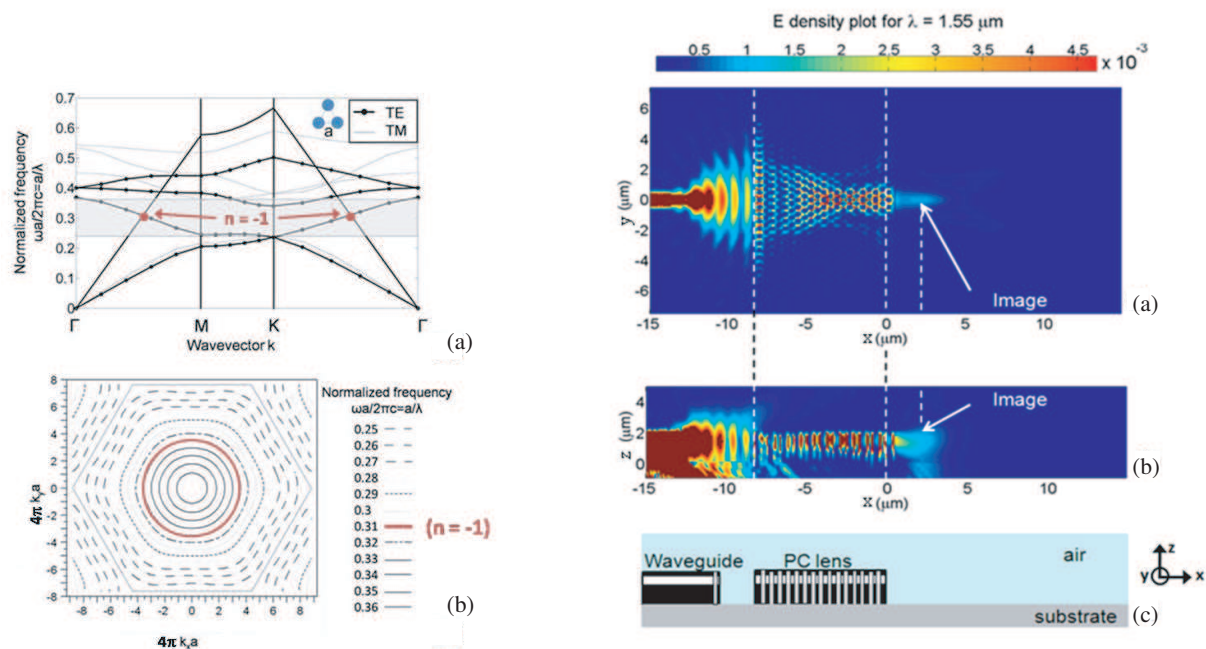

Figure 1: (a) Band structure of the triangular photonic crystal lattice described in text; (b) Iso-frequency surface for the second band.

Figure 2: 3D FDTD simulation of the photonic crystal based $n = -1$ flat lens operating at 1.55 μm; (a) top view, (b) side view, (c) schematic side view.

Figure 3 illustrates the experimental results obtained with such an approach. On the left side, a SEM top view of the prototype is given and on can observe the high quality obtained for the lens in terms of periodicity and regularity for the PC. The light is injected on the left and arrows indicate the theoretical positions of focus spots within and behind the lens for $n = -1$ using ray tracing issued from Snell-Descartes law generalized to negative values of n. On the right side, a SNOM picture of the lens for a wavelength of 1.525 μm is given [12]. This value is chosen to give the brightest spot behind the lens. This image is recorded in air at an altitude corresponding to the confining plane of the vertical heterostructure, approximately 1.7 μm above the InP substrate. First of all, one can note a very high intensity of the optical field at the first lens interface; this reveals a high level of reflection and thus a strong impedance mismatch between air and the lens that remains to be optimized. Second, one can observe interference fringes close to the second interface. This can have various origins related to out of plane dispersion losses or to interaction between the lens interfaces and the collecting probe. Despite all these limitations, a clear spot is visible behind the lens closed to the expected position. Quantitatively, the transmission level is around −30 dB. This spot remains between 1.53 μm and 1.54 μm and the optimal resolution obtained here, obtained after correction induced by the finite size of the probe is 0.8λ. If this value is subwavelength, it doesn't break the Rayleigh criterion (0.5λ) which was expected with our experimental set-up of collection of propagating optical waves in air. This result is nevertheless very close to the optimal value of 0.66λ calculated by theoreticians with similar configurations. It has to be mentioned that, despite our 2D design procedure, high transmission (brightest spot) and $n = -1$ (spot position) do not coincide. This can be due to small variations of the real geometrical parameters achieved around the nominal theoretical values or by the influence of the finite depth of the PC in its 3D configuration. Indeed, the results appear quantitatively very sensitive to this set of parameters. However, the wavelength shift between theory and experiment remains very small (less than 2%) and validates our design procedure.

This demonstration of focusing by a flat PC lens at optical wavelengths is very encouraging even
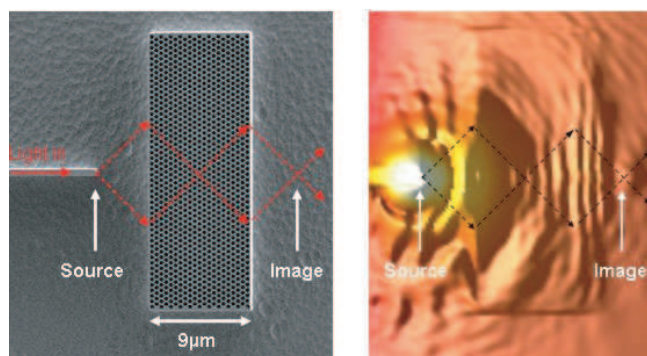
Figure 3: Experimental results for the photonic lens at $\lambda = 1.525\,\mu$m. Left side: SEM top view of the prototype, arrows indicates the light propagation according to Snell's law. Right side: Scanning near-field optical microscope picture of the lens.

if light transmission through the lens has to be optimized. The main difficulty is here to take into account the multiple incidence angles of the light in front of the lens. Indeed, if n can be found isotropic, it is not the case for the surface impedance which strongly varies with incidence. PC surface engineering is commonly proposed to realize this matching over a limited range of incidence but many questions related to the exploitation of additional surface modes to improve transmission and resolution properties of the lens remain open.

3. A PHOTONIC CRYSTAL BASED OPTICAL CLOAKING DEVICE

Negative refraction in PC's is a general concept which can also be used for cloaking purposes. This search for invisibility is very active with metal-dielectric metamaterials in which permittivity and permeability tensors can be "tailored" as will, including negative or smaller than one values of these constitutive parameters. Using conformal mapping techniques, a precise local engineering of ε or μ enables the design of cloaks which bend light in such a way that anything located inside is hidden to the outside, whatever the type of incoming wave [12, 13]. Using PC's such a precise engineering is not easily imaginable and other strategies have to be developed to mimic a similar behavior under restricted operating conditions. Figure 4(a) illustrates schematically such a PC based cloaking device. The basic idea is to rebuild behind the device an incident Gaussian plane wave in amplitude and phase as if the device was absent. To this aim, we make use of two different PC's: PC1 is similar to the one used above and operates with $n = -1$ at 1.55 μm whereas PC2 is opaque to light at the same wavelength. This can be obtained using a higher filling factor and a bigger period to shift the band gap at 1.55 μm. The light motion passing four times a $n = -1/n = 1$ interface is shown in grey. The space region to hide is at the center of PC2. The specific form of PC2 (with the two tips) is used to alleviate straight-through transmission at the singular point where the two PC1 are in contact at input and output. With respect to the discussion of paragraph 2, let us mention that PC1 is always illuminated under single incidence which makes possible the simultaneous optimization of transmission and index matching.

(a)   (b)   (c)
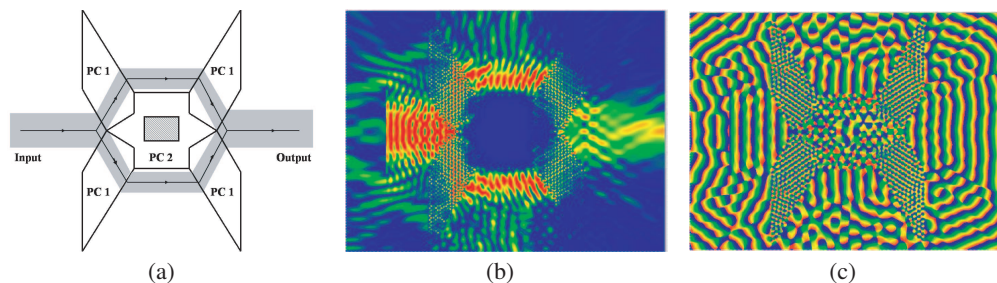
Figure 4: Photonic crystal based cloaking device. (a) Schematic view, PC1 operates in negative refraction regime ($n = -1$), PC2 is opaque and the grey region at the center is hidden for the incident light whose motion is represented by the arrows. (b) Field intensity under cloaking operation at 1.55 μm. (c) Phase plot of the optical field.

Figures 4(b) and 4(c) illustrate the cloaking operating mode whereas Figure 5 shows the figure of merit of the device. It can be seen that the incident wave, first shared in two parts at input is properly reconstructed at output, at the same frequency, showing a regular phase front typical of a plane wave. The central region of the device is completely isolated. The total insertion loss is around $-5/6$ dB, which reveals a good transmission at each $n = -1/n = 1$ interface with losses lower than 1.5 dB. As mentioned before, the device works only for a incident unidirectional plane wave which is a real limitation compared to the metamaterial approach. Conversely, one can note that the hidden area dimension can be chosen as will by enlarging the PC1 extension as long as propagation and insertion losses in PC1 and air remain at reasonable levels. No intrinsic limitation related to the operating wavelength exists with this kind of design. This reasoning can be extended, with the same limitations in mind, to the spatial extension of the incident plane wave.
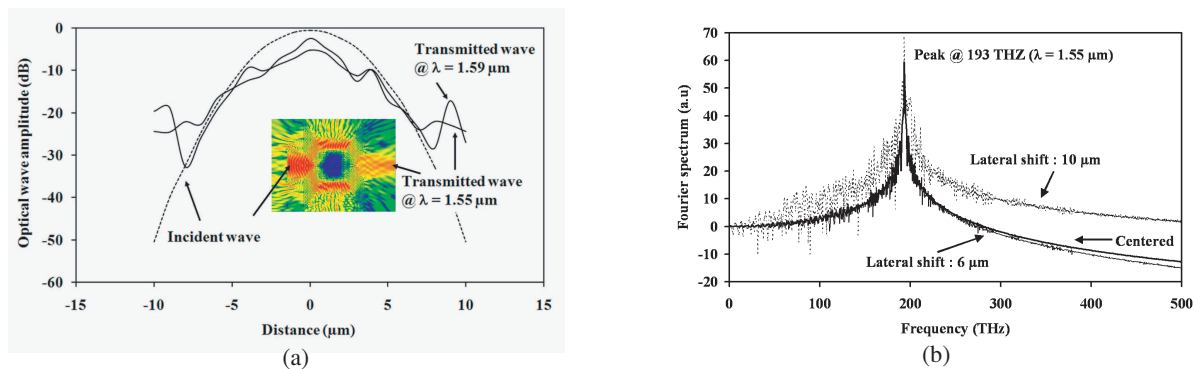
Figure 5: Figures of merit of the cloaking device. (a) Transmission properties at $1.55\,\mu m$ ($193\,THz$). (b) Fourier spectra of the output signal.

4. CONCLUSIONS

In this paper, we have shown the potentialities of 2.5D dielectric photonic crystals for subwavelength imaging or invisibility cloaking operation at optical wavelengths. Subwavelength focusing (0.8λ) at $1.525\,\mu m$ has been obtained with a PC based flat lens. Also, an original design base on the combination of PC's working under different regimes has shown ability to mimic a cloaking operation for an incident plane wave with a high versatility in terms of hidden zone dimension or incident wave spatial extension. It is believed that the exploitation of abnormal propagation regimes in photonic crystals represents a significant step for nanophotonics and that the proposed devices will become building blocks for the future generation of integrated devices.

ACKNOWLEDGMENT

REFERENCES

1. Yablonovitch, E., "Inhibited spontaneous emission in solid state physics and electronics," *Phys. Rev. Lett.*, Vol. 58, No. 20, 2059–2062, 1987.
2. Perrin, M., S. Fasquel, T. Decoopman, X. Mélique, O. Vanbésien, E. Lheurette, and D. Lippens, "Left-handed electromagnetism obtained via nanostructured materials: Comparison with that from microstructures photonic crystals," *J. Opt. A*, Vol. 7, S3–S11, 2005.
3. Veselago, V. G., "The electrodynamics of substance with simultaneously negative values of ε and μ," *Sov. Phys. Usp.*, Vol. 10, No. 4, 509–514, 1968.
4. Pendry, J. B., "Negative refraction makes a perfect lens," *Phys. Rev. Lett.*, Vol. 85, No. 18, 3966–3969, 2000.
5. Berrier, A., M. Mulot, M. Swillo, M. Qiu, L. Thylen, A. Talneau, and S. Anand, "Negative refraction at infrared wavelengths in a two-dimensional photonic crystal," *Phys. Rev. Lett.*, Vol. 93, 073905, 2006.
6. Matsumoto, T., K. Eom, and T. Baba, "Focusing of light by negative refraction in a photonic crystal superlens on silicon-on-insulator substrate," *Opt. Lett.*, Vol. 31, 2786–2788, 2006.

7. Fabre, N., S. Fasquel, C. Legrand, X. Mélique, M. Muller, M. François, O. Vanbésien, and D. Lippens, "Towards focusing using photonic crystal flat lens," *Opto-Electron. Rev.*, Vol. 14, 228–232, 2006.

8. Casse, B. D. F., W. T. Lu, Y. J. Huang, and S. Shridar, "Nano-optical microlens with ultrashort focal length using negative refraction," *Appl. Phys. Lett.*, Vol. 93, 053111, 2008.

9. Fabre, N., L. Lalouat, B. Cluzel, X. Mélique, D. Lippens, F. de Fornel, and O. Vanbésien, "Optical near-field microscopy of light focusing through a photonic crystal flat lens," *Phys. Rev. Lett.*, Vol. 101, 073901, 2008.

10. Croenne, C., N. Fabre, D. P. Gaillot, O. Vanbésien, and D. Lippens, "Bloch impedance in negative index photonic crystals," *Phys. Rev. B*, Vol. 77, 125333, 2008.

11. Vanbésien, O., N. Fabre, X. Mélique, and D. Lippens, "Photonic-crystal-based cloaking device at optical wavelengths," *Appl. Opt.*, Vol. 47, No. 10, 1358–1362, 2008.

12. Cluzel, B., D. Gérard, E. Picard, T. Charvolin, F. de Fornel, and E. Hadji, "Subwavelength imaging of field confinement in a waveguide-integrated photonic crystal cavity," *J. Appl. Phys.*, Vol. 98, 086109, 2005.

13. Pendry, J. B., D. Shurig, and D. R. Smith, "Controlling electromagnetic fields," *Science*, Vol. 312, 1780–1782, 2006.

14. Gaillot, D. P., C. Croenne, and D. Lippens, "An all-dielectric route for terahertz cloaking," *Opt. Exp.*, Vol. 16, No. 6, 3986–3992, 2008.

Broadband Acoustic Cloak with Multilayered Homogeneous Isotropic Materials

Y. Cheng, **J. Y. Xu**, and **X. J. Liu**

Key Lab of Modern Acoustics of MOE, Nanjing University, Nanjing 210093, China

Abstract— In the long wavelength limit, it is shown that the acoustic cloak can be constructed by using concentric multilayered structure with alternating homogeneous isotropic materials. The cloak designs for concealing objects with arbitrary shape and acoustic properties have been proposed. In this paper, we will present an in-depth analysis on the cloaking mechanism of the proposed design. The details of pressure field distribution in each cloak layer are obtained by rigorous acoustic scattering derivations. Calculated pressure fields demonstrate that the cloaking effect is ascribed to a specific multiple scattering process determined by the cloak's microscopic material distribution and structural details.

1. INSTRUCTION

The coordinate transformation based solutions to the Maxwell equations have been demonstrated to be an efficient approach to control the flow of electromagnetic (EM) waves [1]. It enables the design of EM invisibility cloaks which can hide arbitrary objects from exterior EM illumination and suppress all the scattered waves [1–3]. This exciting design was confirmed by full-wave simulations [2] and realized in experiments at microwave frequencies [3]. All these pioneering work further inspire a great deal of research interest in the acoustic cloak due to the similarity between Maxwell equations and acoustic wave equations. In [4], Cummer et al. employed the transformation method to design two-dimensional (2D) cylindrical acoustic cloak. Later on, a three-dimensional (3D) acoustic cloak with the same material specifications was further proposed separately by comparing the acoustic field and dc conductivity equations [5] and by acoustic scattering theory [6]. They demonstrated that the acoustic cloak can bend elastic wave around objects with arbitrary shape and properties. These attractive designs promise potential applications such as hiding submarines from enemy's active sonar.

Analogues to its EM counterpart, the acoustic cloak also require constitutive parameters (mass density and bulk modulus) with anisotropy and space gradient [4–6]. Unfortunately, the material with the required characters can not be found in nature [4, 6]. The construction of acoustic cloak with acoustic metamaterials is challenging for both the considerable absorption of sound wave and complex fabrication process. On the other hand, the acoustic cloak can also be constructed by concentric alternating layered structure with homogeneous isotropic materials. Based on the effective medium theory, the cloaking effect has been demonstrated [7–10]. In this paper, we will present in-depth analysis on how the multilayered cloak conceals the object.

2. FORMULATIONS

Figure 1 shows the configuration of an M-layered acoustic cloak shell with inner radius a and outer radius b. For simplicity, the cloak layers are consecutively numbered as 1, 2, ..., $M-1$, M, while the host medium and the concealed region are denoted as 0 and $M+1$, respectively. Thus, the radius of the inner boundary of the mth layer is $r = R_m$, with $R_0 = b$ and $R_M = a$. The cloak is designed through a two-step procedure [7].

Consider a plane harmonic pressure wave $p_{in} = P_0 e^{i(k_0 x - \omega t)} = P_0 e^{i(k_0 r \cos\varphi - \omega t)}$ incidents on the cloak along the $\varphi = 0$ direction. P_0 is the amplitude of p^{inc}, here i indicates the imaginary component and k_0 is the wave number in the host medium. The pressure fields in all the layers follow the general wave equation since the layered cloak is constructed by homogeneous isotropic materials.
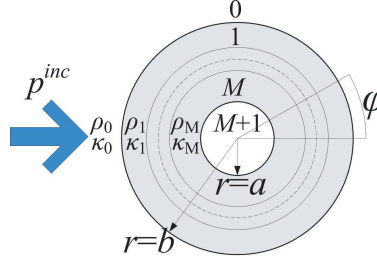
Figure 1: Schematic of the M-layered acoustic cloak in cylindrical coordinates. The inner boundary is at $r = a$ while the outer boundary is at $r = b$. Plane harmonic elastic wave p^{inc} incidents along the $\phi = 0$ direction.

Hence the pressure fields in all regions can be described with the following expressions [11]:

$$p_0 = P_0 \sum_{n=-\infty}^{+\infty} \left[i^n J_n(k_0 r) + C_{0n} H_n^{(1)}(k_0 r) \right] e^{in\varphi}, \ (r > b) \tag{1}$$

$$p_m = P_0 \sum_{n=-\infty}^{+\infty} \left[A_{mn} J_n(k_m r) + C_{mn} H_n^{(1)}(k_m r) \right] e^{in\varphi}, \ (R_m > r > R_{m-1}) \tag{2}$$

$$p_{M+1} = P_0 \sum_{n=-\infty}^{+\infty} \left[A_{(M+1)n} J_n(k_{M+1} r) \right] e^{in\varphi}, \ (r < a) \tag{3}$$

where J_n and H_n are the Bessel and Hankel function of the first kind, respectively; k_m and $k_{(M+1)}$ are the wave number in corresponding layer. C_{0n}, A_{mn}, C_{mn}, and $A_{(M+1)n}$ are the unknown expansion coefficients which can be solved by matching the continuous scalar pressure p and radial velocity $v_r = (i/\omega \rho_r \rho_0)(\partial p/\partial r)$ at all interfaces. By solving the linear equation system, the coefficients are obtained to be

$$
\begin{aligned}
A_{(M+1)n} &= \frac{1}{\beta_M l_M^{11}}, & C_{0n} &= \frac{l_M^{21}}{l_M^{11}} \\
A_{mn} &= \frac{C_{on} l_{m-1}^{12} - l_{m-1}^{22}}{\beta_{m-1}(l_{m-1}^{12} l_{m-1}^{21} - l_{m-1}^{11} l_{m-1}^{22})}, & C_{mn} &= \frac{l_{m-1}^{21} - C_{on} l_{m-1}^{11}}{\beta_{m-1}(l_{m-1}^{12} l_{m-1}^{21} - l_{m-1}^{11} l_{m-1}^{22})} \ (m = 1 \text{ to } M),
\end{aligned}
\tag{4}
$$

where $\beta_M = Q_0 Q_1 Q_2 \ldots Q_M$, $\beta_{m-1} = Q_0 Q_1 Q_2 \ldots Q_{m-1}$ with $Q_m = \frac{-\pi k_m R_m}{2i}$; l_M^{ij} is the entry of the 2×2 matrix $L_M = P_0 P_1 P_2 \ldots P_M$, l_{m-1}^{ij} is the entry of the 2×2 matrix $L_{m-1} = P_0 P_1 P_2 \ldots P_{m-1}$, with the entries of the 2×2 matrix P_m are

$$
\begin{aligned}
p_m^{11} &= \alpha_m J_n'(k_{m+1} R_m) H_n^{(1)}(k_m R_m) - J_n(k_{m+1} R_m) H_n^{(1)'}(k_m R_m), \\
p_m^{12} &= \alpha_m H_n^{(1)'}(k_{m+1} R_m) H_n^{(1)}(k_m R_m) - H_n^{(1)}(k_{m+1} R_m) H_n^{(1)'}(k_m R_m), \\
p_m^{21} &= J_n(k_{m+1} R_m) J_n'(k_m R_m) - \alpha_m J_n'(k_{m+1} R_m) J_n(k_m R_m), \\
p_m^{22} &= H_n^{(1)}(k_{m+1} R_m) J_n'(k_m R_m) - \alpha_m H_n^{(1)'}(k_{m+1} R_m) J_n(k_m R_m).
\end{aligned}
\tag{5}
$$

Here $\alpha_m = k_{m+1} \rho_m / k_m \rho_{m+1}$.

3. RESULTS AND DISCUSSIONS

When all the coefficients are obtained, the pressure field in the each layer can be reconstructed accordingly through Eqs. (1)–(3). As an example, we use the same geometry details and material distributions as that in [7]. The cloak geometry uses the common configuration $b = 2a$, the cloak shell is constructed by 40 thin layers, and the host medium is set as water with $\rho_0 (= 998 \, \text{kg/m}^3)$ and $\kappa_0 (= 2.19 \, \text{GPa})$.

    We first calculate the local velocity in each layer of the cloak to provide a straightforward view on the cloaking process. The x component v_{mx} and y component v_{my} at $k_0 = 2\pi$ are plotted

in Figs. 2(a), (b), respectively. A fraction of the cloak in the second quadrant is depicted for convenience of illustration. Comparing Fig. 2(a) with 2(b), it is noted that v_{mx} exhibits parallel plane wavefront outside the cloak, while v_{my} appears to be near zero. This behavior indicates that the host medium vibrates along x direction uniformly, and the cloak reduces most perturbations to the vibration in host medium. On the contrary to the uniform vibration in host medium, v_{mx} and v_{my} in the cloak shell shows transformed distribution in both radial and angular directions, indicating the specific multiple scattering processes. The occurrence of v_{my} should be ascribed to the tangential scattering between the circular layers. In addition, the amplitude of v_m in group A is remarkable high compared with that in group B because the mass density ρ_A is much lower than the corresponding ρ_B. We further calculate the norm velocity $|v_m| = \sqrt{v_{mx}^2 + v_{my}^2}$, as depicted in Fig. 2(c). $|v_m|$ vanishes as r decreases to about $1.1m$ along the incident direction, implying that the cloak may not reflect the incoming wave back in the incident direction.
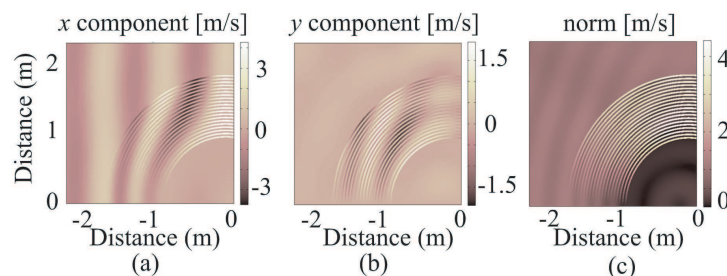
Figure 2: Contour plots of local velocity in each layer of the cloak shell: (a) x component, (b) y component, and (c) norm. The spatial frequency $k_0 = 2\pi$.

We further investigate the scattering mechanism which induces the cloaking process. Contrary to the cloak constructed by metamaterials, multiple scattering phenomena are ubiquitous throughout the cloak due to the distinct contrast in mass density between adjacent layers. Inside the cloak layers, the incident wave is recursively scattered by the interfaces of the cloak. When the inward wave impinges the circular interface $r = R_{m-1}$, partial wave is transmitted into layer m in random directions. Then the transmitted wave is reflected back at the interface $r = R_m$, and reflected again and again between the two interfaces. Acoustic waves in all the layers perform similar walk which leads to the multiple scattering process in the entire cloak. On the other hand, since the layer thickness of the cloak is much smaller than the wavelength, the scattering is prominently prompt. The randomness of the interaction could be averaged out by the large number of scattering events, so that the final scattering process appears to be a deterministic distribution of intensity. Here, we treat the multiple-scattering process as a multiple transmission-reflection process which incorporates full description of the problem in the long wavelength limit and reduces the random process to steady-state averages of the statistics.

4. CONCLUSIONS

With the help of multiple scattering algorithms, we present in-depth study on the cloaking mechanism of the proposed acoustic cloak. By investigating the details of pressure field distribution in each cloak layer, we find that the cloaking effect is induced by the specific multiple scattering process. This specific multiple scattering process can be controlled by the microscopic material distribution and structural details in the multilayered structure, which jointly determines the cloak's macroscopic scattering characteristics. The results may provide us with deeper insights into the cloaking phenomenon as well as a simple path to experimental realization of acoustic cloak.

REFERENCES

1. Pendry, J. B., D. Schurig, and D. R. Smith, "Controlling electromagnetic fields," *Science*, Vol. 312, 1780, 2006.

2. Cummer, S. A., B.-I. Popa, D. Schurig, et al., "Full-wave simulations of electromagnetic cloaking structures," *Phys. Rev. E*, Vol. 74, 036621, 2006.

3. Schurig, D., J. J. Mock, B. J. Justice, et al., "Metamaterial electromagnetic cloak at microwave frequencies," *Science*, Vol. 314, 977, 2006.

4. Cummer, S. A. and D. Schurig, "One path to acoustic cloaking," *New J. Phys.*, Vol. 9, 45, 2007.

5. Chen, H. and C. T. Chan, "Acoustic cloaking in three dimensions using acoustic metamaterials," *Appl. Phys. Lett.*, Vol. 91, 183518, 2007.

6. Cummer, S. A., B.-I. Popa, D. Schurig, et al., "Scattering theory derivation of a 3D acoustic cloaking shell," *Phys. Rev. Lett.*, Vol. 100, 024301, 2008.

7. Cheng, Y., F. Yang, J. Y. Xu, and X. J. Liu, "A multilayer structured acoustic cloak with homogeneous isotropic materials," *Appl. Phys. Lett.*, Vol. 92, 151913, 2008.

8. Torrent, D. and J. Sánchez-Dehesa, "Acoustic cloaking in two dimensions a feasible approach," *New J. Phys.*, Vol. 10, 063015, 2008.

9. Cheng, Y. and X. J. Liu, "Resonance effects in broadband acoustic cloak with multilayered homogeneous isotropic materials," *Appl. Phys. Lett.*, Vol. 93, 071903, 2008.

10. Cheng, Y. and X. J. Liu, "Three dimensional multilayered acoustic cloak with homogeneous isotropic materials," *Appl. Phys. A*, DOI 10.1007/s00339-008-4882-7.

11. Morse, P. M. and K. U. Ingard, *Theoretical Acoustics*, Mcgraw-Hill, New York, 1968.

Pulse Electromagnetic Sounding of the Permafrost Layered Medium

V. L. Mironov and K. V. Muzalevskiy

Kirensky Institute of Physics, SB RAS Krasnoyarsk, Russia

Abstract— In this paper, the reflection of a super wide-band electromagnetic pulse from an active layer of permafrost is theoretically modeled. The dielectric model applied for such a frequency dispersive medium takes into account in thawed and frozen soil the contributions from the organic/mineral fraction of the soil, liquid soil water, bound soil water, and ice particles. The numerical algorithm developed allows to model the amplitude of the GPR pulse, which finally makes possible to simulate GPR images of a layered permafrost medium in the course of its freezing and thawing. The impact of a frozen/thawed boundaries inside the layer on the reflected pulse form were studied. The dynamic range of the GPR allowing to monitor the depth of thawed layer is evaluated.

1. INTRODUCTION

In view of the global climate change, the seasonal variations inside the active permafrost layer has become a crucial factor in modeling the global climate change processes. The ground penetrating radars (GPR) can be used as an effective instrument to detect and retrieve the subsurface structure [1, 2]. At that the adequacy of relevant wave scattering models is an important factor in developing the GPR data processing algorithms. In this research, a model relating to the reflection of a wideband electromagnetic pulse from the active permafrost layer is considered. The pulse is emitted with a magnetic dipole antenna located over the layered half-space and received at the observation point placed above the layer at some distance from the transmitting dipole.

2. METHODOLOGIES APPLIED

On the basis of the spectroscopic soil dielectric model [3], the GPR pulse complex amplitude was calculated using solution of a 3-D wave scattering problem. A point magnetic dipole was taken to model the pulse transmitting antenna. The electric current density of a point like dipole was assigned as follows:

$$\vec{j}^m(x, y, z, t) = I(t) l \delta(x) \delta(y) \delta(z - D) \cdot \vec{e}_z. \tag{1}$$

The pulse current $I(t)$ was taken in a form of the first derivative of the *Blackman-Harris* window function having duration of one nanosecond, $T = 1$ ns. The geometry of the problem is shown in Fig. 1. Given the rectangular coordinates (x, y, z), the plane of contact between the permafrost layer and air is determined by the equation $z = 0$. The lower dielectric half-space, $z < 0$, has an inhomogeneous layered structure. The electromagnetic field can be expressed with the use of time and spatial spectra. The E_x component of the reflected field, $z > 0$, has the form (2) where $w = \sqrt{k_0^2 \varepsilon(\omega) - \chi^2}$-transverse wave number, k_0-free space wave number, $I^m(\omega)$-spectrum of pulse current, $H_1^{(1)}(\chi \Delta L)$-*Hunkel* function of a first kind, l-length of the dipole, $R(\chi)$-*Fresnel* reflection coefficient, which is to be computed via an iterative procedure in case of inhomogeneous medium
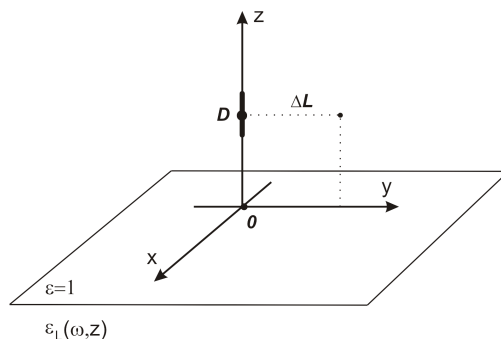

Figure 1: Geometry of the problem.

being considered. Inside the permafrost layer, the dependence of complex permittivity on the depth, z, is given by $\varepsilon_L(\omega, z)$. The complex permittivity profiles $\varepsilon_L(\omega, z)$ were obtained using the spectral dielectric model [3, 4], in conjunction with the temperature and moisture profiles, as measured in the area of Franklin Bluffs Alaska [5]. The temperature profiles measured are shown in Fig. 2.

$$E_x(\omega, x = 0, y = \Delta L, z = D) = -\frac{i}{8\pi} I^m(\omega) l \int\limits_{-\infty}^{\infty} \frac{\chi^2 d\chi}{w} R(\chi) H_1^{(1)}(\chi \Delta L) e^{2iwD} \qquad (2)$$

The respective dielectric constant, ε', and loss factor, ε'', profiles inside the layer are shown in Fig. 3, calculated with the use of temperatures dependences presented in Fig. 2.
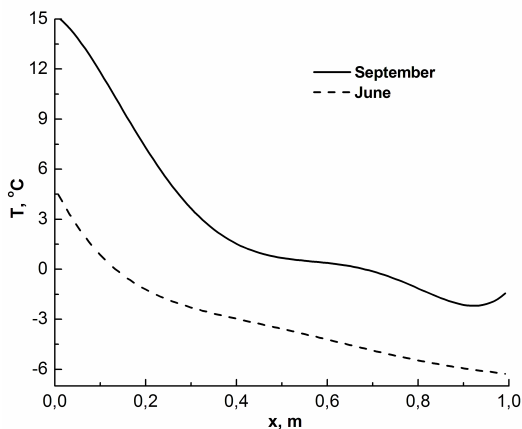
Figure 2: Measured temperature profiles in the permafrost active layer.
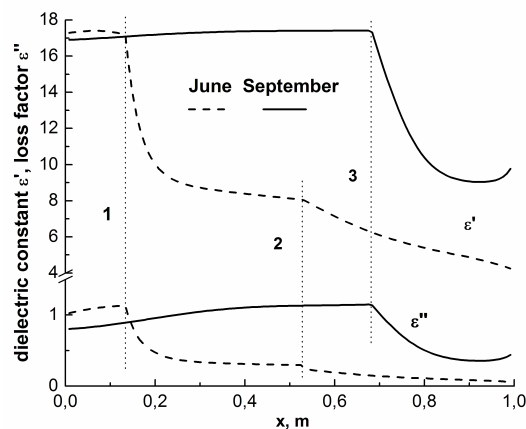
Figure 3: The soil dielectric constant, ε', and loss factor, ε'', as a function of depth in case of the June and September temperature profiles shown in Fig. 2. Wave frequency is equal to 1 GHz.

3. ANALYSIS OF PULSE RESPONCE FROM A PERMAFROST LAYER

Based on the integral representation (2), the reflected field pulse was computed, as shown in Fig. 4. The height of the transmitter and receiver over the permafrost layer boundary was assigned to be of $D = 0.1$ m, with the distance between them being of $\Delta L = 0.2$ m. The duration of the *Blackman-Harris* window was taken to be $T = 1.0$ ns. As seen from Fig. 3, there are noticeable breakdowns in the complex permittivity at the boundaries between the thawed ($0 \leq x \leq 0.12$ m or $0 \leq x \leq 0.7$ m) and frozen (0.12 m $< x$ or 0.7 m $< x$) layers in July and September, respectively. In
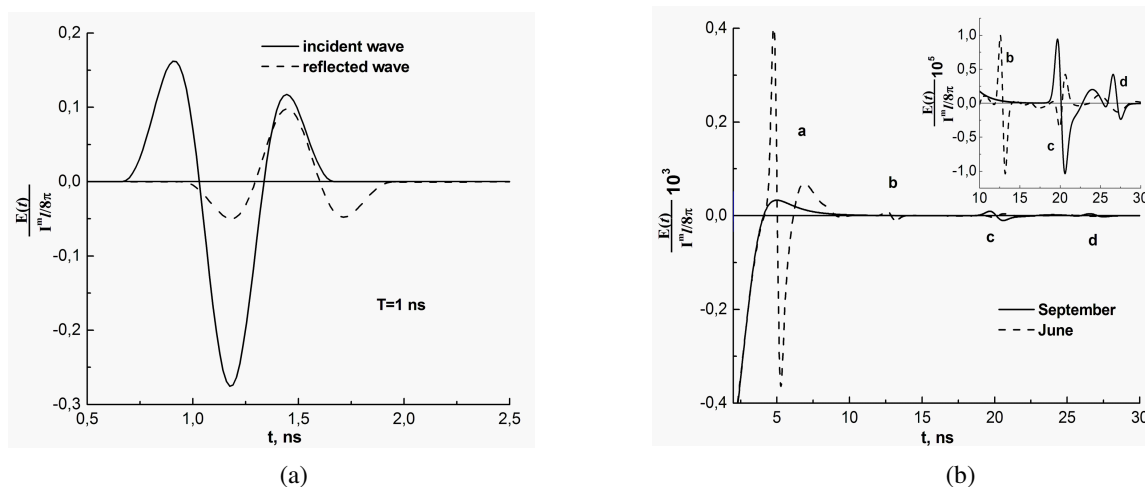

(a)

(b)

Figure 4: (a) The pulse forms for the incident and reflected waves. (b) Reflected pulse forms.

Fig. 4(a), the dash line corresponds to the pulses reflected from the air-soil boundary in June and September. While in Fig. 4(b), the graphs show forms of the pulses reflected at the thawed/frozen boundaries inside the active layer, the dash and solid lines corresponding to June and September, respectively. The pulses marked with (a), (b), (c), and (d) have to be related to the complex permittivity discontinuities at the depths of 0.12 and 0.7 m, as shown in Fig. 3. As can be noticed from Fig. 4(b), a GPR with pulse duration of 1 ns and dynamic range of about 100 dB can ensure detection of the internal structure of thawing permafrost layer, with the instrument being located above the surface of the layer.

4. CONCLUSIONS

In this paper, there was carried out a feasibility study regarding the use of the GPR for monitoring the structure of the thawing/freezing permafrost layer. The electromagnetic field of a 1.0 ns pulse was numerically simulated with the use of a rigorous solution of the *Maxwell* equations in a form of the integral over frequency and spatial wave number domains. The *Blackman-Harris* window pulse waveform in conjunction with a realistic dielectric model of the permafrost soil have been employed to come up with the conclusion that a GPR with the dynamic range of about 100 dB can ensure monitoring of an active permafrost layer in the course of thawing.

REFERENCES

1. Daniels, D., *Surface Penetrating Radar*, Inst. Elect. Eng., London, U.K., 1996.
2. Lu, T., W. Cai, and P. Zhang, "Discontinuous Galerkin time-domain method for GPR simulation in dispersive media," *IEEE Trans. Geosci. Remote Sensing*, Vol. 43, No. 1, 72–80, 2005.
3. Mironov, V. L., M. C. Dobson, V. H. Kaupp, S. A. Komarov, and V. N. Kleshchenko, "Generalized refractive mixing dielectric model for moist soils," *IEEE Trans. Geosci. Remote Sensing*, Vol. 42, No. 4, 773–785, 2004.
4. Mironov, V. L., V. H. Kaupp, S. A. Komarov, and V. N. Kleshchenko, "Frozen soil dielectric model using unfrozen water spectroscopic parameters," *Proc. IGARSS'03*, Vol. 7, 4172–4174, Toulouse, France, 2003.
5. http://www.geobotany.uaf.edu

Evaluation of Scattering in Collision Avoidance Radar Application

Wei-Han Lee and Wen-Jiao Liao

Department of Electrical Engineering, National Taiwan University of Science and Technology, Taiwan

Abstract— This work uses a high frequency electromagnetic numerical tool to evaluate the scattering characteristics of vehicles. The projected application is for use on the collision avoidance radar utilizing frequency modulated continuous wave. Experiment results show that a car is a relatively complicated scattering target, which can be regarded as a collection of several scattering sources. The distributions of scattered signals vary as the geo locations of the radio source and scattering targets change. Effects of critical parameters such as the longitudinal distance to the target, lateral offset of the target and the orientation of the target, are investigated. The findings can be used to develop detection and tracking algorithms for collision avoidance radars.

1. INTRODUCTION

In recent years, the collision avoidance radar (CAR) has been put on luxury cars to assist the drivers for prevention of car accidents. Most of the CAR systems developed use millimeter wave or optical laser techniques. Unlike infrared, which is often affected by the weather conditions, the microwave and millimeter wave radar are immune to atmospheric absorption. The frequency modulated continuous wave (FMCW) technique can be used to determine the target velocity and estimate the distance to the target as well [1, 2]. An alternative approach to implement CAR is using the pulse radar scheme. By calculating the phase delay time, the geometric relationship of the source, scattering targets and the receiver can be determined. However, since the pulse radar requires an impulse transmitter and a broadband antenna, currently, the pulse radar option is less attractive in terms of cost. On the other hand, the components required to implement the FMCW radar is ready and cost effective [3]. Nevertheless, the narrow band nature of FMCW radar is likely to be affected by the multi-path interference [4].

The numerical tool used in this work is Numerical Electromagnetic Code-Basic Scattering Code (NEC-BSC), which is a high frequency electromagnetic code that implements the ray tracing technique and the geometric theory of diffraction to calculate the scattered electric fields from complicated targets. Because the study of target scattering features requires evaluating the radar echo distribution from various target locations and target orientations, the NEC-BSC simulation model needs to be modified very frequently. Therefore, the functions of setting up of the simulation model, calculating scattered fields, retrieving simulation results as well as data processing are incorporated into a MATLAB program. The scattering mechanisms calculated in NEC-BSC include direct propagation, reflection, edge diffraction, vertex diffraction and higher order terms. Since NEC-BSC allows one to calculate different scattering terms separately, the effect of a specific target structure or certain geometric relationships can be investigated individually to reveal their physical insights.

The setup of the simulation model is introduced in the second part of this paper, which also addresses the calibration of direct coupling from source antenna to receiving antennas. Parametric studies associated to the vehicle's longitudinal distance, lateral offset and orientation are conducted in the third part. The analysis of simulation results is also given. Primary findings of this work are presented in the fourth part.

2. CONSTRUCTION OF SIMULATION MODEL

In the simulation model, a four dipole array is served as the transmitting antenna. A total of 101 observation points, which spans for 1 meter, are located on the sides of the antenna. A metal plate is put behind the antenna and observation points since we assumed that they are placed at the front of a car.

Due to the proximity between the antenna and observation points, substantial coupling may take place. Therefore, a calibration measure is needed to eliminate the background fields at the observation point. The major coupling mechanisms include direct transmission from the antenna, reflection from the back plate, and diffraction from the edges of the back plate. Simulation results shown in Figure 2 indicate that edge diffraction is relatively small comparing to the other two factors, and therefore can be neglected.

The car, which serves as the scattering target, is composed by 12 flat conducting plates in the NEC-BSC model as shown in Figure 3. The length, width and height are 3 m, 1.6 m and 1.1 m,
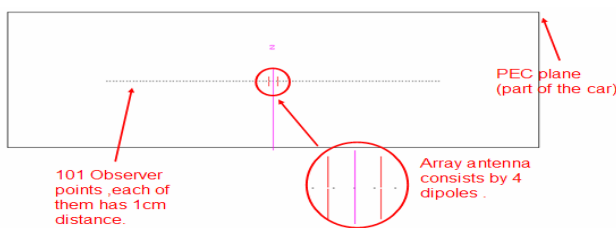
Figure 1: Locations of radar transmitter and observation points.

respectively. A large metal plate is also incorporated to emulate the ground. The vehicle position and orientation can be modified by the MATLAB control program. Electric field at the observation points are recorded for different simulation setups for further data processing.
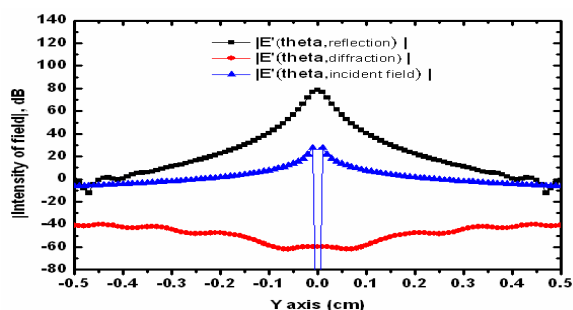

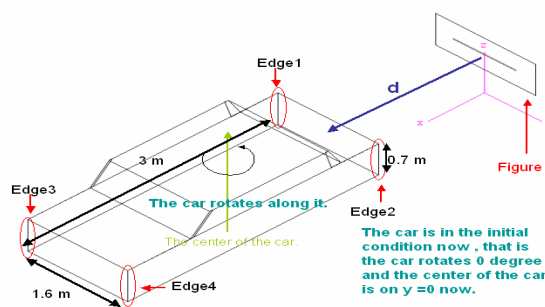Figure 2: Couplings from the radar transmitter to observation points.

Figure 3: The vehicle target model and its geometric relationship to the radar transmitter.

Since both the scattering from the vehicle and the coupling from the transmitting antenna appear at the observation points, we need to perform certain calibration measures to eliminate the background interference. This can be done by first calculating the fields at the observation points without the presence of the vehicle, then the resulting background coupling fields can be subtracted from following simulation results, which have the car in place.

The source array antenna consists of four half wavelength dipoles. The spacing between each dipole is also half wavelength. The frequency of the transmitted signal is 5.8 GHz. A total of 101 observation points are set on the sides of the array antenna, and the spacing between observation points is 1 cm. The vehicle target, which contains several plates and edges, is likely to be a collection of multiple scatters. Therefore, the small spacing between the observation points, which is smaller than the quarter wavelength at 5.8 GHz, is chosen so that constructive and destructive additions of different scattering sources and paths can be revealed. As to the notation, the sum of electric fields is represented by E. The first order scattering terms such as reflection and edge diffraction are denoted as E', while second order terms are denoted as E''. Since the antenna elements are set vertically and aligned with the z axis, most of the backscattering belongs to E_θ. Therefore, the following observations are focused on the changes in E_θ. When one car is trailing behind another one, the reflection from the tail of the front car may be the primary scattering source, while the vertical edges shown in Figure 3 are not as important. However, the dominant specular reflection may not appear all the time, because some lateral offsets may exist between two cars. Therefore, the diffracted fields, which scattered to a larger area, may play important roles in determining the orientation and relative geometric relationship to the target vehicle. In the following experiments, we compare the received scattered fields of different target setups on horizontally aligned observation points to analyze the importance of different scattering mechanisms.

3. ANALYSIS OF SIMULATION RESULTS

The vehicle target is arranged in three different scenarios to perform parametric studies on the scattering mechanisms of collision avoidance radar application. The three sets of experiments include as follows.

3.1. Longitudinal Movement

The car target is placed right in front of the radar transmitting antenna. The longitudinal distances between the car's tail and the antenna are ranged between 1 to 3 meters with 0.4 meter spacing. Since the NEC-BSC code allows one to determine which scattering terms to be computed, we examine not only the overall electric fields but individual scattering component as well. Figure 4 shows the overall fields of six different longitudinal distances. In general, as the longitudinal distance increases, the amount of total field diminishes. However, the field decreasing rate is not consistent and reveals some ripple-like variation at different observation points. Next, we examine the reflected fields in Figure 5. It shows the changes in reflected fields are proportional to the changes in distance. Therefore, we can conclude that the disturbance in the overall fields comes from other scattering terms, which are likely to be the diffractions from vertical edges of the car's tail. Because there are two vertical edges at the back and their distances to the observation points are different, multipath interference may take place. The diffracted fields shown in Figure 6 exhibit ripple-like variations, which are the result of constructive and destructive additions of two diffraction sources. Judging from the averages of diffracted fields of the six experiments, one can roughly estimate the changes in the longitudinal distance. Because the car target is placed in front of the radar transmitter, the specular reflection term is the dominant source of backscattering. The diffracted fields are relatively weak and can only impose some disturbance in the overall field.
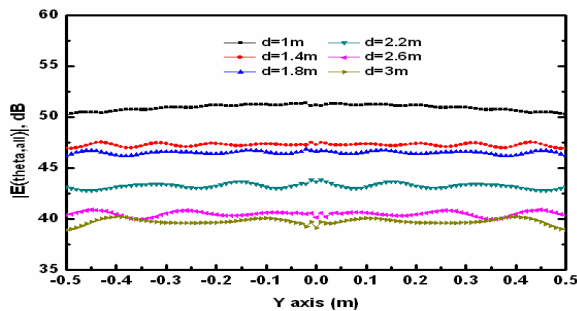

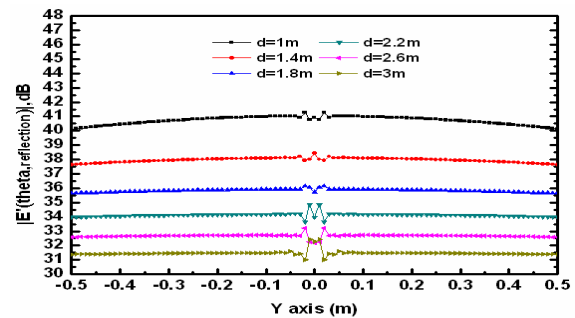Figure 4: Overall fields at various longitudinal distances.

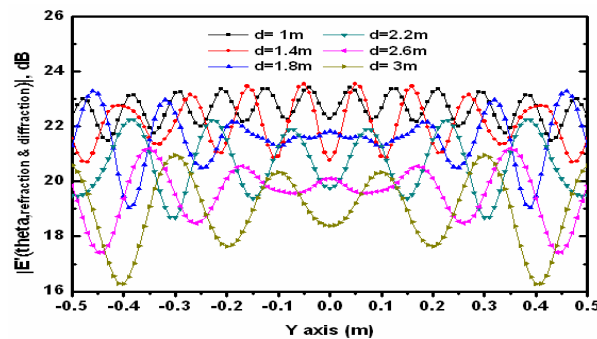Figure 5: Reflected fields at various longitudinal distances.

Figure 6: Diffracted fields at various longitudinal distances.

3.2. Lateral Movement

In this set of experiments, the car's tail was placed 2.5 meter in front of the radar antenna. 10 trials of lateral offsets ranging from 0 to 90 cm were performed. Figure 7 shows the overall fields, the curves indicate that the car is shifted to the right. As the car moves to the right, specular returns only appear on observation points on the right hand side. The reflected fields shown in Figure 8 indicates that when the car center is placed at $Y = 0.6$ m, only observation points located at $Y > -0.4$ receive reflected fields. As the car center is further shifted to $Y = 0.9$ m, only points located at $Y > -0.2$ receive reflected fields. The continuously decreasing fields shown in Figure 7 are contributed by diffracted fields. When the car's position moves, the diffracted field distribution also moves as Figure 9 indicates.
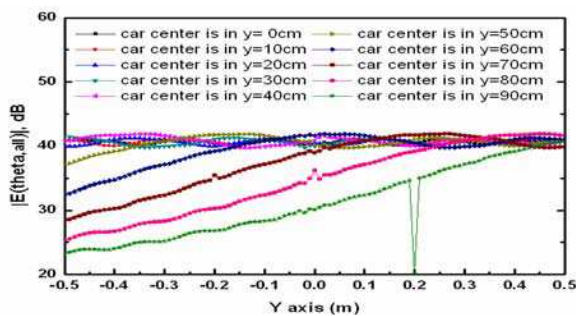
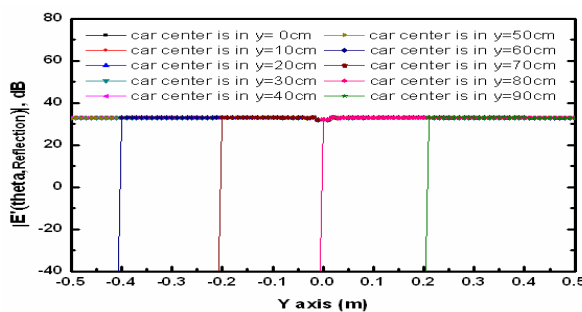Figure 7: Overall fields at various lateral offsets.

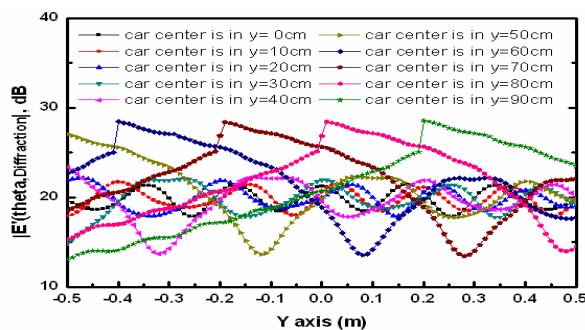Figure 8: Reflected fields at various lateral offsets.

Figure 9: Diffracted fields at various lateral offsets.

3.3. Vehicle Orientation

In this set of experiments, the vehicle's end was placed 2.5 m in front of the radar antenna. The vehicle was rotated counter clockwise 2° a step. A total of 10 trials were performed. Figure 10 shows the overall fields. When the car was rotated more than 12°, the reflected field disappeared, and we only observed the contribution from the diffracted field, which is shown in Figure 11. The peaks in Figure 11 are corresponding to Edge 2 in Figure 3. As the car was rotated counter clockwise, the peak was shifted to the left.
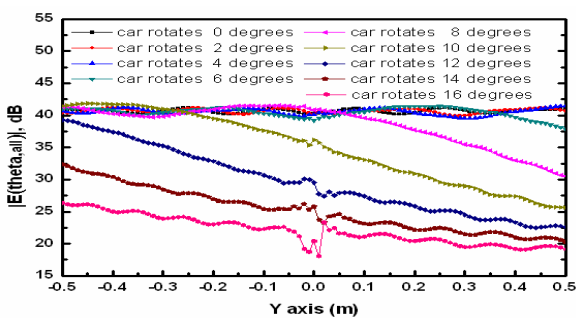

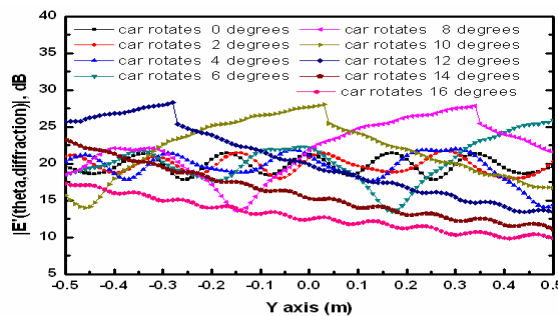Figure 10: Overall fields at various car orientations.

Figure 11: Diffracted fields at various car orientations.

In the first set of experiments, the vehicle distance can be estimated using field strengths at any observation point. For the second and third sets of experiments, it is easier to determine the target location using results collected from multiple observation points.

4. CONCLUSIONS

Through experiments and discussions presented above, we can conclude that when the car is placed in front of the radar antenna, the car's end plate is the primary source for the dominant specular reflection. The strong reflection term can be used to estimate the car distance directly. However, in most cases, the car target may exhibit some lateral offset or orientation variation. Therefore, the specular reflection terms may not appear on the receivers. Nevertheless, the diffracted fields

from the car's vertical edges, though much weaker than reflection, may serves as a more reliable scattering source due to its broad distribution.

According to the experiment results, the car's position or orientation can be determined judging from the field variations in observation points. Research results of this work can be forwarded to the development of the target detection, location and movement prediction algorithms. To improve the emulation of collision avoidance radar operation, future work topics include exploration of other scattering sources and consideration of the multiple cars scenario.

REFERENCES

1. Chan, K. L. and S. R. Judah, "A beam scanning frequency modulated continous wave radar," *IEEE Trans. Instruction and Measurement*, Vol. 47, 1223–1227, 1998.
2. Woods, G. S., D. L. Maskell, and M. V. Mahoney, "A high accuracy microwave ranging system for industrial applications," *IEEE Trans. Instruction and Measurement*, Vol. 42, 812–816, 1993.
3. McGregor, J. A. and E. M. Poulter, "Switching system for single antenna operation of an S-band FMCW radar," *IEEE Proc. — Radar, Sonar Naving*, Vol. 141, No. 4, 241–248, 1994.
4. Chet, K. V., C. B. Kuan, and C. H. Teik, "Development of a ground-based radar for scattering measurements," *IEEE Antennas and Propagation Magazine*, Vol. 45, No. 2, 36–42, 2003.

Monitoring of Satellite Thermal Patterns of Warm Core Ring in Subarctic Sea Surface

S. Nakamura

Kyoto University, Japan

Abstract— This notes a part of monitoring of satellite thermal patterns of warm core ring in subarctic sea surface. A modeling is introduced in a scope of hydrodynamics. In practice, satellite data should be combined to the data obtained by direct measurement of the related factors by the survey ships along the stretched legs.

1. INTRODUCTION

This work is a part of the research project started in Kyoto University. The author has studied satellite thermal patterns on the sea surface found when he had received the satellite signals directly by a simple system settled on the coast facing the ocean. Signals from a couple of the existing satellites with a polar orbital motion are obtained by a processing software for a personal computer. The author aimed to give a model in order to realize a warm water ring found in the subarctic cold ocean water zone.

2. SATELLITE DATA AND HYDROGRAPHIC DATA

In this work, the author introduces a result of monitoring combining the data issued officially by the Hydrographic Office. A special reference is in the subarctic zone of the northwestern Pacific. In order to receive the satellite signals, the author has had agreements by NOAA and JMA (Japan Meteorological Agency).

3. MODEL AND HYDRODYNAMICS

For the author's convenience, a front is introduced as a common boundary of two water masses. This case is that a warm ocean water mass is contacting to a cold water mass. In the ocean, the water masses can not see only looking about the surface areas. The water motion in the ocean is three dimensional, for example, both of the two interested water masses. Problems on an ocean front should be considered in a scope of hydrodynamics. The ocean water covers the earth's surface, so that it is necessary to consider an effect of the earth's rotation even in this work.

Then, a modeling of ocean front evolution should be a problem of hydrodynamics of the ocean water, or, a problem of geophysical hydrodynamics. Nevertheless, it is necessary us to aware of that we have to observe or survey the ocean as a part of the earth. For observations or surveys, there are various instruments and techniques.

Now, in the author's case now, it is essential to know about actual water motion in the ocean. There is found a thermal pattern on the ocean. Then, we use the satellite data and the data in-situ obtained by survey ships on the stretched legs.

With the author's understanding of the satellite thermal pattern on the sea surface under a scope of hydrodynamics, a hypothetical evolution of the ocean front between the subtropical warm water mass and the subarctic cold water mass is a reasonable to give an illustration of a process for formation of a warm core ring in the subarctic sea area.

The interested two water masses around the ocean front is strongly affected by stresses caused by difference of the two water motion. In fact, the front observed is complicated pattern, though geophysical hydrodynamics tells us what is essential pattern.

4. FRONT AND WARM CORE RING

One of convenient illustration series for the interested ocean front may be shown as a following process (Figure 1), for example, a straight ocean front is undulating by the effect of the two water masses to evolve up to form a matured warm core ring. That is to say, as follow:

A) Straight line of ocean front.
B) Sinusoidal undulation of ocean front.
C) Nonlinear process of ocean front to form a kick crest.

D) Nonlinear process of ocean front to form a spit crest.
E) A spit crest evolves to form an initial stage of warm core ring.
F) A growing stage of warm core ring.
G) A matured stage of warm core ring to separate out of the warm water mass.
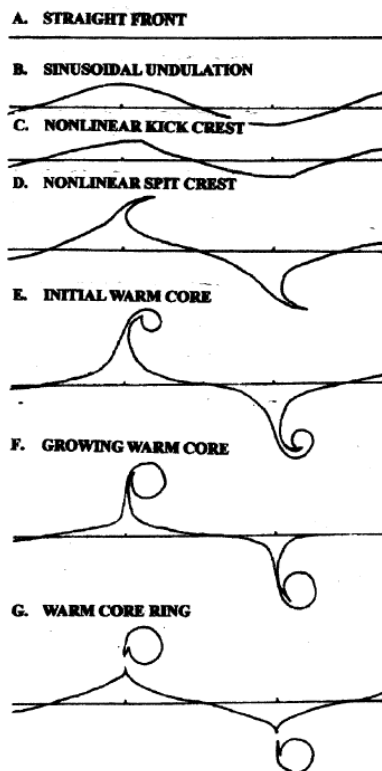

Figure 1: A model for warm core ring formation.

This evolution process of an ocean front can be identified after a combining the data of the satellite thermal pattern of the sea surface and the in-situ data of the survey ships on their leg. The satellite thermal pattern informs only sea surface information though the interested warm core ring formation is three dimensional. Exactly speaking, the ring has to be studied in a scope of hydrodynamics. Then, we could see the actual structure of the interested warm core ring in the subarctic zone of the ocean.

An example of the warm core ring in the subarctic sea area is shown in Figure 2. This warm core has rather long life because the cooled sea surface water of the skin layer tending to be stable
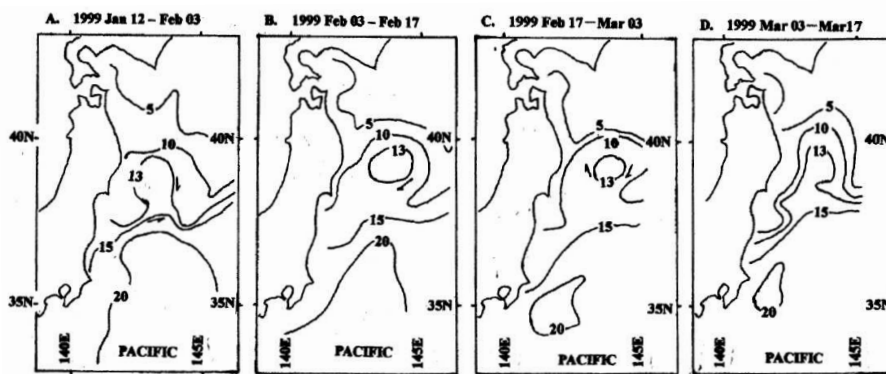

Figure 2: Example of a warm core ring in cold water of the ocean in subarctic zone. (1) Isotherms of 5, 10, 13, 15 and 20°C are shown. (2) Warm core is in the area encircled by the isotherm of 13°C.

by forming a thermal shelter around the warm core water. A more detail of the warm core ring could be understood dynamically when we could have a more advanced mathematical model in a scope of geophysical hydrodynamics.

5. CONCLUSIONS

Referring to the data obtained by satellite monitoring and ocean survey in-situ, typical problem of warm core ring is introduced in order to see what process is found in a process of a ocean front undulation in the actual ocean. For the author's convenience, One of the examples is introduced for a warm core ring found in the cold water of the subarctic ocean zone. The motions of the warm and cold water masses around the ocean front could be a measure to give us a valuable information about what is seen on the sea surface. That could be taken a part of the subsurface or deeper part of the global ocean waters. In this work, the author introduced a concept model process of a warm core ring for his understanding in a scope of hydrodynamics, and, to have a more advanced mathematical model in a scope of geophysical hydrodynamics.

Monitoring of Satellite Thermal Patterns of Ocean Front Evolution in Relation to Ocean Water Stratification

S. Nakamura

Kyoto University, Japan

Abstract— This work concerns on monitoring of satellite thermal patterns of ocean front evolution in relation to ocean water stratification. First, densimetry of the ocean water is determined by the two main factors, i.e., salinity and water temperature. Then, ocean water motions can be understood in relation to the specific stratification patterns. The water motion in the ocean is governed by densimetric condition under the earth's rotation. Several simplified patterns are introduced to see a relation to what is found in the satellite thermal monitoring.

1. INTRODUCTION

The author introduces a brief note about monitoring of satellite thermal patterns of ocean front evolution in relation to ocean water stratification. First of all, the factors of salinity and water temperature are noted how the two factors act to stratify the ocean waters. This stratification is the densimetric problem of the ocean water. The two factors diagram, i.e., ST-diagram with a parameter of density is convenient to see water mass identification. Several cases are noted for helping to see what stratified pattern in the ocean should be in the author's consideration in this work. Several simple models are introduced in order to show typical patterns of vertical cross sections around the interested ocean. This work may help for finding a key to our advanced understanding about a three dimensional process of the ocean water motion found in the ocean.

2. OCEANOGRAPHIC DATA AND SATELLITE DATA

The author has studied satellite thermal patterns on the sea surface found when he had received the satellite signals directly by a simple system settled on the coast facing the ocean. Signals from a couple of the existing satellites with a polar orbital motion are obtained by a system for directly receiving satellite signal in a personal computer. Although, the satellite date informs us only what is seen on the sea surface. In order to see about ocean water motions around an interested ocean front evolution, it is necessary to survey about the factors not only on the sea surface but in the sea as a whole. Referring to the past data, the author introduces the two important factors for the ocean water motions. These factors are salinity and water temperature in-situ.

The author here concentrates his interest to note a way to realize typical patterns of the existing cases of vertical structure of the ocean water motion around the front.

3. SALINITY AND WATER TEMPERATURE

Usually, the water in nature on land have many contents in a dissolved and suspended form. As for the ocean water, it is essentially same to the water contents on land and in the ocean though the important factors for the ocean water motion are salinity and water temperature. The other miner factors in the ocean may be taken as passive factors of the ocean water motion and not be taken as the most important factors for the ocean water motion around the interested front. The density of ocean water is essentially controlled by these two factors, i.e., salinity S and water temperature T. A diagram, so-called "ST-diagram", is a convenient to plot the ocean water masses. In the diagram, a parameter of "sigma-t" (or σ_t), is introduced and defined as an index of the specific densimetric property of the ocean water. The parameter σ_t is defined as a difference of the ocean water density and to the pure water.

4. MODELING OF STRATIFIED OCEAN

With the author's understanding of the satellite thermal pattern on the sea surface under a scope of hydrodynamics, a hypothetical evolution of the ocean front between a warm ocean water mass and a cold ocean water mass is in the author's interest. He believes it necessary here to consider the satellite data in relation to the vertical structure of the ocean waters around the front.

In order to illustrate a typical thermal pattern in a vertical section of ocean water stratification, several models are introduced for convenience (cf. Figure 1).

The ocean scientists has been used to assume a two layered ocean for convenience of their mathematical modeling of ocean water motions. Two layer models have been frequently used for making us it easy to analyze by a mathematical formulation and by a numerical computation under some approximation and assumptions. They have had used to assume several conditions for solving the model problem, even though the solution is simply the solution of their model but the actual ocean.

Then, the author introduces three typical patterns for his interested vertical section of the ocean in a simplified form as follow (refer to Figure 1).
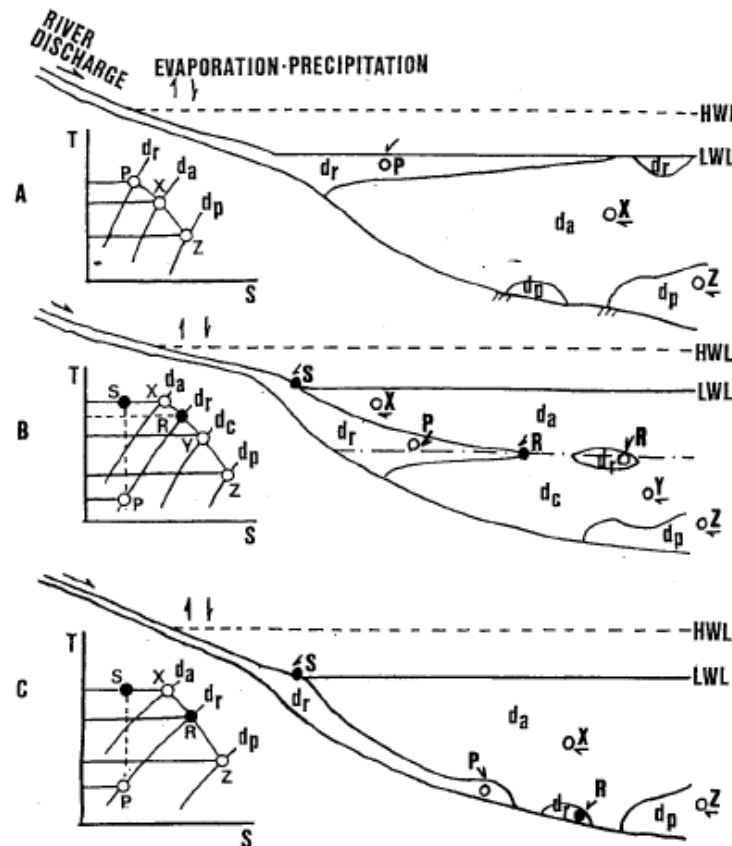

Figure 1: Modeling of ocean waters stratification. A — Warm coastal water covering ocean deep water; B — Cold coastal water intruding to form subsurface layer (with a detoutched subsurface densimetric lens); C — Chili coastal water creeping on ocean floor.

The first pattern is a simple layer model for a coastal warm water forming surface layer and deep layer of low temperature as in case A of Figure 1. The surface layer might be formed by an inflow of river waters (d_r) from the land, or, a coastal layer. This layer is affected by an atmospheric effect and by the solar radiation effect. In fact, the thermocline is found between the surface layer (d_r) and the deep layer (d_p). As for the density of d_r, d_a, and d_p, stable condition should be as $d_r(P) > d_a(X) > d_p(Z)$.

In this case A, the tip of the water d_r could be detected as a ocean front formed between the coastal water and ocean waters.

The second pattern as case B in Figure 1 is a case of a horizontal water intrusion into a slit to form a stable layer between the two layer d_a and d_c. Then, densimetric condition should be satisfied as $d_a(X) > d_r(P) > d_c(Y) > d_p(Z)$ for some condition of the salinity and the water temperature. Balance of buoyancy is interesting factor in this case. When the water at the position S of d_r is transferred to the position P, it is possible that the water density of d_r at P maintains without any affect of the ocean waters of d_a and d_c if there is no double diffusion effect of salinity and water temperature.

When any double diffusion effect is appeared at position P or at R, it may seen a change of the water density. Nevertheless, no signal about the change can be detected by the satellite because

the change is only appeared under the sea surface.

When the double diffusion effect is a buoyant effect, a dispatched water mass in a form of a lens may tend to move upward to find a balancing position. Conditionally, the satellite can detect the lens if the lens get to the sea surface as a thermal patch.

In this case B, the location of S in the figure must be taken as the ocean front.

The third pattern as case C in Figure 1 is a case of chill coastal water creeping on the ocean floor. In this case, densimetric stability is maintained under the condition of $d_a(X) > d_r(R) > d_p(Z)$.

With an understanding about densimetry of the ocean water as what noted above, we could have a chance to have more advanced satellite monitoring system for the thermal pattern on the sea surface as a part of the global ocean in a scope of hydrodynamics, and, in a scope of geophysical hydrodynamics.

5. A KEY TO HYDRODYNAMICS

The author's interest is to see the ocean front evolution by satellite thermal monitoring. Then, he has to see the ocean water motion around the interested ocean front.

Now, what is noted above for the illustration of Figure 1 make us easy to consider problem on water masses, and it looks far from any dynamical understanding.

For solving the water motion in the ocean, it is essential to start the formulation of water motion in a form of the equation of motion first. Processes of double diffusion for the processes of thermal diffusion and of salinity should be formulated in order to construct a set of equations as a simultaneous equation to be solved. Some conditions should be assumed as the initial condition and as the boundary conditions.

Looking at any one of the three patterns, it is hard to get an exact solution, One of the convenient techniques is to apply a numerical modeling under an approximation. A solution tells us about mechanism of the interested ocean water motion if modeling is appropriate as an approximation. Then, we have to be careful to see whether our modeling fit well to found in the actual processes in the ocean.

6. SATELLITE DATA

As noted above already, the satellite monitoring is possible for finding an interesting process not only the thermal process but other processes on the ocean surface. Then, it is necessary to see several other observed data in-situ obtained by the survey ships. With some understanding after joining the satellite data on the ocean surface covering the ocean front and the other data referring to the knowledge on hydrodynamic theory or the related subjects, it might be easily found what processes are found around the ocean front on the sea surface in relation to what ocean water motion found under the sea surface.

7. DISCUSSIONS

After the above noted considerations, various kinds of the existing processes could be well realized.

For example, the surface layer in the pattern A in Figure 1 might be seem in the mid-latitudinal zone, subtropical zone and tropical zone. Range of a scaling might be for a local area of several kilometers square or for a wide area of several hundred kilometers square or a global scale. One of the interesting processes might be found on the coastal zone in the Arctic Sea.

The pattern B in Figure 1 might be found in mid-latitudinal zone, for example, in a small bay in an area of several kilometers square. A possible exciting problem is to consider about the Mediterranean lenses found in the area of the subsurface zone in the northeastern Atlantic.

There is an interesting problem about what is "meddy" which is taken as an eddy of a specific water mass flown out of the Mediterranean. This meddy might be a key to see the water motion in the northeastern Atlantic as stated it by Professor Ana Martins of the University of Azores (personal information to the author).

The pattern C in Figure 1 might be found, for example, in a small bay as well as in the coastal zones of the Antarctic Continent.

It should be considered another phases in the other areas for the aim in this work. Nevertheless, the author would not give any notice concerning these phases. Simply, a brief note for the essential problems is introduced here.

The author has noted several problems which should be considered at utilizing the satellite thermal monitoring of ocean front evolution in relation to the structure of the ocean water stratification. Essential problem might be solved under a simple condition though any model for actual ocean

should be constructed under some conditions for the locality of the interested area. Topographic and bathymetric conditions must be the two of the most important factors to be considered. There are many problems left to be solved as soon as possible for a more fruitful human activity in the world.

8. CONCLUSIONS

The author introduced his note for what should be considered at utilizing the satellite thermal monitoring system of ocean front evolution in relation to the structure of the ocean water stratification. Essential problem might be solved under some conditions, and, any model for actual ocean should be constructed under some conditions fit for the specific local conditions, i.e., topography, bathymetry and other related factors.

In this work, salinity S and water temperature T are the mast important factors for the densimetric problem of the ocean water motion. A ST-diagram is noted to be widely used in practice at finding the density of the ocean water sample. Referring to the ST-diagram with a parameter of sigma-t, several typical patterns are introduced for illustrating the vertical structure of the ocean water. Topography, bathymetry and the related factors of the specific locality in the interested area should be considered at satellite thermal monitoring of the ocean front evolution found on the ocean surface.

Referring to what introduced in this work, it could be obtained a key to promote a more advanced research for finding to the details of the ocean front evolution.

Coupling the CUPID and TRGM Models to Study the Temporal Variations of Thermal Emission Directionality of Crop Canopies

H. G. Huang[1, 2, 3], **Q.-H. Liu**[1, 2], **and W. H. Qin**[1, 2, 4]

[1]State Key Laboratory of Remote Sensing Science
Institute of Remote Sensing Applications, Chinese Academy of Sciences
Beijing 100101, China
[2]Beijing Normal University, China
[3]Key Laboratory for Silviculture and Conservation of Ministry of Education
Beijing Forestry University, China
[4]NASA Goddard Space Flight Center, Greenbelt, Maryland, USA

Abstract— Many field experiments have observed the significant temporal variation of TIR (Thermal Infrared) emission directionality, so it is necessary to quantitatively explain this phenomenon to improve the application of the directional remotely sensed TIR observation. In this paper, the 3D Thermal Radiosity Graphics Model (TRGM) and an extended energy balance model CUPID are coupled to simulate the seasonal and daily variation of directional brightness temperature (DBT) of wheat canopies. The improvement of CUPID model mainly focused on simultaneously retrieving the shaded temperature and sunlit temperature of soil surface, which is very important for coupling with TRGM Vegetation structure, soil water content and micrometeorology data are the main input parameters that affect the component temperature distribution. Driven by the input data set *in situ*, the half-hourly DBT dataset of wheat canopy are simulated for 22 days. Several new variables are defined to analyze the DBT temporal variation, which include nadir DBT, "hot spot" DBT, azimuthally averaged DBT at zenith angle of 55°, standard deviation of DBT at zenith angle of 55°, etc. The validation results show that the simulated directional temperature differences have consistent trend with field measurements. Based on the sensitivity analysis, it is found that the soil water content and air temperature are the two most significant factors that affect the DBT distributions, while the wind speed and solar radiation affect DBT much more than air temperature.

1. INTRODUCTION

Surface radiative temperature measured by remote sensors may significantly vary with different view angles, which has been discovered as early as in the mid 1960s [5, 9]. According to the experimental results presented by Paw [18], the directional radiative temperature of row crop canopies can vary up to 13 K, thus it is not possible to accurately evaluate surface fluxes by using a single directional surface radiative temperature measured by remote sensors. This issue becomes more and more important as the satellite sensors with large swath angles are coming into use, such as MODIS, AVHRR, ATSR and AATSR. In recent years, researchers have developed a series of directional thermal radiation models for homogeneous canopies [4, 13, 20], row crop canopies [2, 11, 19] and three-dimensional surfaces [6, 14]. All these models require component temperatures as input to predict the *directional brightness temperature*, denoted as $T_B(\theta)$.

Both $T_B(\theta)$ and the distribution of component temperatures in a canopy vary with time [8]. Chehbouni et al. [1] discussed seasonal variations of $T_B(\theta)$ through analyzing radiative temperature differences between nadir and off-nadir directions during the growing season of a semiarid grassland. The results of his study indicate that differences between nadir $T_B(0)$ and off-nadir $T_B(55)$ could be up to 5 K. Such differences are highly correlated with surface soil moisture and vegetation biomass, and also have seasonal cycles. Guillevic et al. [6] demonstrated about 2 to 3 K daily $T_B(\theta)$ variation from 6 h to 13 h with the *directional brightness temperature* data collected by Kimes and Kirchner [10] in a cotton canopy. But little work has been done to model the temporal variation of $T_B(\theta)$ under different environmental conditions, which is the purpose of this study.

2. METHODOLOGY

The extended CUPID model [7] based on the 1990 version [16] is chosen to simulate the canopy component temperature distribution because it accounts for leaf inclined angle effect on leaf temperature distribution. 3D Thermal Radiosity-Graphics Model TRGM [14] is employed to simulate the $T_B(\theta)$ distribution because of its high accuracy for realistic, inhomogeneous canopies.

These two models are both originally developed for crop canopies so it is feasible to couple them together. To characterize the $T_B(\theta)$'s anisotropy, six parameters are defined here, including $T_B(0)$, $T_B(\theta_s, \phi_s)$, $T_B(55)$, $T_{B,std}(55)$, $\Delta T_{B,hot}$ and ΔT_B). The 55 degree is selected to represent the off-nadir brightness temperature based on the geometry of ATSR and MODIS.

The $\Delta T_{B,hot}$ is the difference between $T_B(\theta_s, \phi_s)$ and the mean $T_B(\theta)$ value at the two sides of the W_h (view zenith width) around θ_s (see Fig. 1 and Eq. (1)). The W_h is derived from the $T_B(\theta)$ profiles in solar principle plane by marking off the "hot" area around θ_s.

$$\Delta T_{B,hot} = T_B(\theta_s, \phi_s) - 0.5[T_B(\theta_s - 0.5W_h) + T_B(\theta_s + 0.5W_h)] \tag{1}$$

ΔT_B and $\Delta T_{B,hot}$ represent the view zenith effect and hot spot effect separately. Maximum ΔT_B of each day, $\Delta T_{B,\max}$, is defined to characterize the seasonal variation pattern.
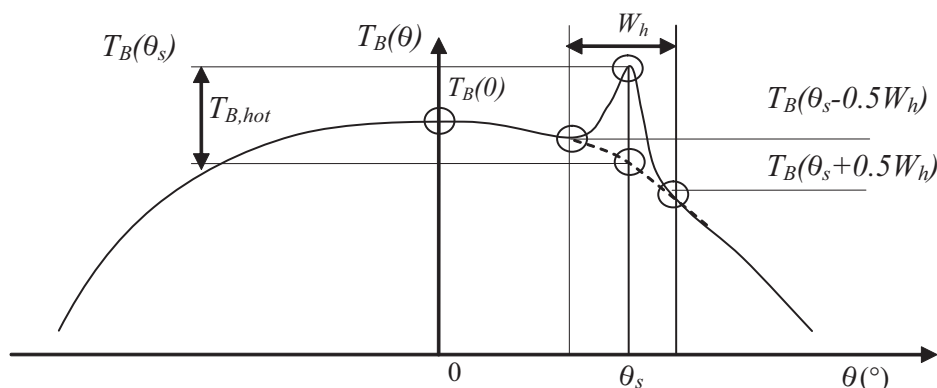

Figure 1: Illustration of hot spot intensity calculation from brightness temperature profile in solar principle plane.

Based on simulated half-hourly $T_B(\theta)$ day by day, these characteristic parameters are calculated to demonstrate the temporal variation trend of directional brightness temperature. And the measured data can be used to partly validate the simulated results.

3. FIELD EXPERIMENT AND DATA PREPROCESSING

The field experiment over winter wheat canopies was conducted for 22 days, from April 1 to April 21 as well as March 25 in 2001, in Beijing (116°34'33", 40°11'40"), China. The winter wheat was planted in the previous fall. The row orientation was 6 degrees Northwest with row spacing of 0.14 m. The average canopy height increased from 0.08 m to 0.32 m with LAI from 0.3 to 2.3, as well as the clumping index (estimated crop row width/row spacing) from 0.3 to 0.7. The mean leaf angle is about 60°. The micrometeorological data at 2 m height from ground (R_{sun}, T_a and h_a, u) and soil water content were measured *in situ*. The soil type is *Aquic Brown Soil* (a kind of salty loam in China) with a bulk density of 1.3 g/cm^3 at surface. Field moisture capacity and wilting coefficient are 22% and 10% at the soil surface, respectively. There was a rainfall (about 2–3 cm) on April 5, and it was irrigated (about 0.6 cm) on April 14.

The field measured data on April 20 and 21 in 2001 are used to validate the coupled model. The canopy height is 0.32 m with LAI of 2.3. Row width is about 0.1 m (70% of the row spacing). At the two days, atmosphere visibility is 12 km, and daily averaged wind speed is 1.2 m/s. T_a was between 21 and 4°C respectively. The four-component (sunlit and shaded leaf or soil) temperatures are measured by a thermocouple thermometer (JM424 digital thermometer). The emissivities of leaf and soil are 0.98 and 0.95 respectively. A two-channel thermal radiometer system (with FOV of 8.4°) was employed to measure $T_B(\theta)$ in spectrum of 8 ∼ 11 μm and 10.6 ∼ 14 μm respectively. A linear correction method is applied to remove the temporal effect [12].

By changing the measurement values of wind speed, solar radiation, air temperature and humidity, soil water content and LAI to all possible values within their physical ranges, additional $T_B(0)$, ΔT_B and $\Delta T_{B,hot}$ data are computed to build a complete dataset under all possible conditions. This dataset is used to test how the five variables affect the $T_B(\theta)$ distribution.

4. RESULTS ANALYSIS

4.1. Validation of the Coupled Model

The ΔT_B measurements in two days (April 20, 21) are used to partly verify the coupled model. On April 20, the RMSE of $T_B(0)$, $T_{B,mean}(55)$, $T_B(\theta_s, \phi_s)$ and ΔT_B simulation are about 0.8 K, 1.1 K, 2.3 K and 2.0 K respectively. The results prove that the diurnal trend of simulated $T_B(\theta)$ is reasonable.

4.2. Temporal Variation of $T_B(\theta)$

Figure 2 shows the temporal variation of ΔT_B, $T_{B,std}(55)$ and $\Delta T_{B,hot}$ from April 1 to April 21, 2001 based on simulation datasets. In most days, the ΔT_B behaves a "W" shape. The middle peak is at noon. $T_{B,std}(55)$ has a similar trend with ΔT_B except that the absolute value is about 1/3 of ΔT_B. That means the azimuth angular effect is symmetrically smaller that that of zenith angular effect. The maximum $\Delta T_{B,hot}$ appears at noon with less than 0.5 K. Moreover, the hot spot zenith width $\Delta\theta$ is estimated as 4°. The small hot spot size is quite different form hot spot effect in VIS/NIR range.
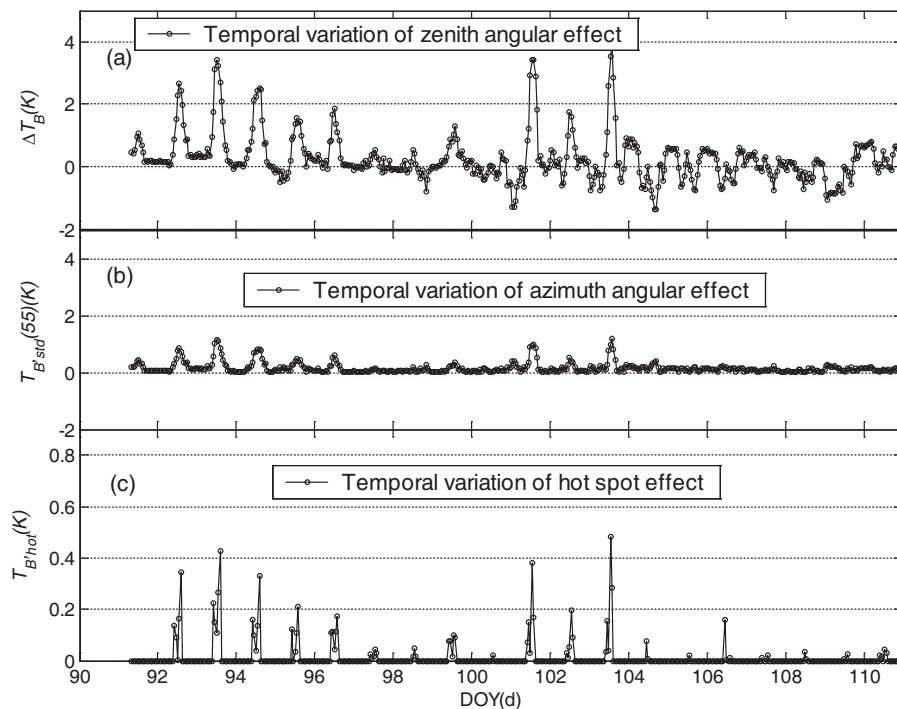

Figure 2: The temporal variations of (a) ΔT_B, and (b) $T_{B,std}(55)$ and $\Delta T_{B,hot}$ for the wheat canopy from April 1 to April 21.
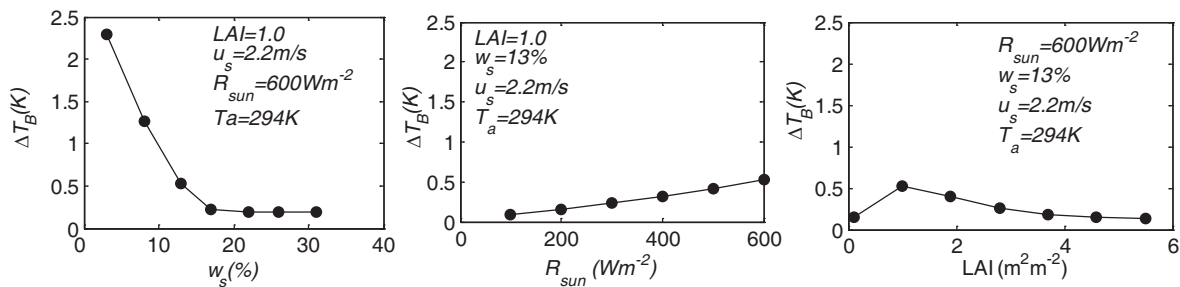

Figure 3: Sensitivity analysis of the three major parameters that affect on ΔT_B: $w_s R_{sun}$ and LAI. The other input parameters are the same at 10:30 in April 17, 2001.

Several days has stronger directional effect, such as the highest ΔT_B occurred in DOY 93 and DOY 104. On the contrary, there exist two periods without significant directional effect, including those days between DOY 94 and 99, and days after DOY 105. These variations are well correlated to the daily total solar radiation (R_{tot}) and soil surface moisture (w_s).

4.3. Factors Affecting $T_B(\theta)$ Temporal Variation

There are five sensitive parameters that affect the $T_B(\theta)$ temporal variation, including w_s, T_a, LAI, R_{sun}, and u (Fig. 3).

5. CONCLUSIONS

By coupling the CUPID model and the TRGM model, the temporal variations of thermal emission directionality are studied in this paper. We found that the $T_B(\theta)$ distributions have significant temporal variation related to the different land surface parameters. The daily $\Delta T_{B,hot}$ or ΔT_B have a peak value around 12:00. ΔT_B is mainly controlled by the soil moisture, air temperature, LAI, wind speed and solar radiation, but it is not sensitive to air humidity. The maximum ΔT_B reached 4 K when the LAI of the row structured canopy is about 0.8. The $\Delta T_{B,hot}$ is generally small (< 0.3 K), which is quite different from VIS/NIR range.

REFERENCES

1. Chehbouni, A., Y. Nouvellon, Y. H. Kerr, M. S. Moran, C. Watts, L. Prevot, D. C. Goodrich, and S. Rambal, "Directional effect on radiative surface temperature measurements over a semiarid grassland site," *Remote Sens. Environ.*, Vol. 76, 360–372, 2001.

2. Du, Y., Q. H. Liu, L. Chen, Q. Liu, and T. Yu, "Modeling directional brightness temperature of the winter wheat canopy at the ear stage," *IEEE Trans. Geosci. Remote Sensing*, Vol. 45, 3721–3739, 2007.

3. Flerchinger, G. N., W. P. Kustas, and M. A. Weltz, "Simulating surface energy fluxes and radiometric surface temperatures for two arid vegetation communities using the SHAW model," *J. Applied Meterol.*, Vol. 37, 449–460, 1998.

4. Francois, C., C. Ottlé, and L. Prévot, "Analytical parameterization of canopy directional emissivity and directional radiance in the thermal infrared. Application on the retrieval of soil and foliage temperatures using two directional measurements," *Int. J. Remote Sens.*, Vol. 18, 2587–2621, 1997.

5. Fuchs, M., E. T. Kanemasu, J. P. Kerr, and C. B. Tanner, "Effect of viewing angle on canopy temperature measurements with infrared thermometers," *Agron. J.*, Vol. 59, 494–496, 1967.

6. Guillevic, P., J. P. Gastellu-Etchegorry, J. Demarty, and L. Prévot, "Thermal infrared radiative transfer within three-dimensional vegetation covers," *Journal of Geophysical Research (Atmospheres)*, Vol. 108, ACL 6-1, 2004.

7. Huang, H., X. Xin, Q. Liu, Q. Liu, C. Liangfu, and X. Li, "Modeling soil component temperature distribution by extending CUPID model," *IGARSS'2006*, IEEE, Denver, USA, 2006.

8. Humes, K. S., W. P. Kustas, M. S. Moran, W. D. Nichols, and M. A. Weltz, "Variability of emissivity and surface temperature over a sparsely vegetated surface," *Water Resour. Res.*, Vol. 30, 1299–1310, 1994.

9. Monteith, J. L. and G. Szeicz, "Radiative temperature in the heat balance of natural surfaces," *Quart. J. Roy. Meteorol. Soc.*, Vol. 88, 496–507, 1962.

10. Kimes, D. S. and J. A. Kirchner, "Directional radiometric measurements of row crop temperatures," *Int. J. Remote Sens.*, Vol. 4, 299–311, 1983.

11. Kimes, D. S., S. B. Idso, J. P. J. Pinter, R. J. Reginato, and R. D. Jackson, "View angle effects in the radiometric measurement of plant canopy temperatures," *Remote Sens. Environ.*, Vol. 10, 273–284, 1980.

12. Liu, Q. H., Q. Liu, X. Z. Xin, R. R. Deng, G. L. Tian, L. F. Chen, J. D. Wang, and X. W. Li, "Experimental study on directionality in thermal infrared observations of corn canopy," *IGARSS'2001*, Sydney, Australia, 2001.

13. Liu, Q., L. F. Chen, Q. H. Liu, and Q. Xiao, "A radiation transfer model to predict canopy radiation in thermal infrared band," *Chinese Journal of Remote Sensing*, Vol. 7, 161–167, 2003.

14. Liu, Q. H., H. G. Huang, W. H. Qin, K. H. Fu, and X. W. Li, "An extended 3D radiosity-graphics combined model for studying thermal emission directionality of crop canopy," *IEEE Trans. Geosci. Remote Sensing*, Vol. 45, 2900–2918, 2007.

15. Luquet, D., A. Vidal, J. Dauzat, A. Bégué, A. Olioso, and P. Clouvel, "Using directional TIR measurements and 3D simulations to assess the limitations and opportunities of water stress indices," *Remote Sens. Environ.*, Vol. 90, 53–62, 2004.

16. Norman, J. M., J.-L. Chen, and N. Goel, "Thermal emissivity and infrared temperature dependence on plant canopy architecture and view angle," *IGARSS'90*, 1747–1750, IEEE, Piscataway, NJ, 1990.

17. Paw, U. K. T., "Anisotropy of thermal infrared exitance above and within plant canopies," *ESA-SP-319*, 369–374, Courchevel, France, 1991.

18. Paw, K. T., "Development of models for thermal infrared radiation above and within plant canopies," *ISPRS J. Photogrammetry Remote Sensing*, Vol. 47, 189–203, 1992.

19. Sobrino, J. and V. Caselles, "Thermal infrared radiance model for interpreting the directional radiometric temperature of a vegetative surface," *Remote Sens. Environ.*, Vol. 33, 193, 1990.

20. Verhoef, W., L. Jia, Q. Xiao, and Z. Su, "Unified optical-thermal four-stream radiative transfer theory for homogeneous vegetation canopies," *IEEE Trans. Geosci. Remote Sensing*, Vol. 45, 1808–1822, 2007.

Ray Tracing of CMP Antenna Array GPR System

Xuan Feng[1], Motoyuki Sato[2], and Cai Liu[1]

[1]College of Geo-Exploration Science and Technology, Jilin University
No. 6 Xi Minzhu Street, Changchun 130026, China
[2]Center for Northeast Asian Studies, Tohoku University
41 Kawauchi, Aob-ku, Sendai 980-8576, Japan

Abstract— There is much interest in vehicle mounted array GPR systems for landmine detection and work is being carried out on various national, because high speed array GPR systems can sweep a large area in a relatively short time to improve performance and efficiency. We developed the vehicle mounted stepped-frequency (SF) CMP antenna array GPR system. The system has three pairs of Vivaldi antennas that configure a symmetrical CMP antenna array. Based on this configuration, multi-offset CMP data gather can be acquired at every measurement position. Migration is an important signal processing method that can improve signal-clutter ratio and reconstruct subsurface image. Diffraction stacking migration and Kirchhoff migration sum amplitudes along the migration trajectory that generally is hyperbolic. But when the ground surface varies acutely, we have to modify the migration trajectory by the technique of ray tracing. We compute the travel time between transmitter, receiver and each subsurface scattering point, and search the propagation ray depending on the Fermat's principle. The method is tested by an experiment data acquired by the SF CMP antenna array GPR system. The target is a metal ball that is buried under a sand mound. A nice result of ray tracing is shown in the case.

1. INTRODUCTION

Accidents caused by antipersonnel (AP) landmines remaining from past conflicts have received considerable public exposure in the last several years. The global land mine crisis is creating immense social and economic problems worldwide [1]. Because ground-penetrating radar (GPR) is sensitive to changes in all three electromagnetic characteristics of a medium, electric permittivity, electric conductivity, and magnetic permeability, GPR is one of a number of technologies that has been extensively researched as a means of improving mine detection efficiency. In this case, the GPR antenna(s) must be elevated above the ground [2, 3]. This requirement results in heavy surface clutter, especially when the ground is rough [4, 5].

A GPR array is usually employed to sweep a large area in a relatively short time to improve performance and efficiency [6]. We proposed a robust GPR common midpoint (CMP) array technique, including CMP antenna array and CMP processing technique, to enhance signal-to-clutter ratio (SCR). The CMP multi-fold coverage technique, which is widely applied in seismic exploration as a method of enhancing useful signals, was employed into the configuration of antenna array for collecting CMP GPR data over the region to be probed. Then, the CMP processing techniques, including velocity spectrum and multi-fold stacking, were combined to increase the SCR [7]. Based on the velocity, the migration technique can be used to reconstruct subsurface imaging and dramatically enhance the SCR by summing the diffraction wave [8]. The migration technique [9] is now commonly used to process GPR data, and has been in use for almost five decades in seismic reflection surveys.

Generally the diffraction stacking migration or Kirchhoff migration need compute the travel time. But the travel time surface generally is smooth spherical surface in the case of zero-offset data and smooth ellipsoidal surface in the case of nonzero-offset data whose curvature is governed by the velocity function. When the height of ground surface varies largely in the very rough ground area, for example mound, the travel time surface will be affected by the ground surface. In this paper we will consider the effect of the ground surface into the travel time for the CMP Antenna Array GPR System. We will compute the ray path of the electromagnetic wave among the transmitters and diffraction points and receivers and get the diffraction travel time surface.

2. RAY TRACING AND TRAVEL TIME

The basic idea is to find the first-arrival traveltimes by using Fermat's principles in a velocity model. According to Fermat's principle, the path with the smallest traveltime is the one best approximating the ray trajectory. It is possible to determine the traveltime between two arbitrary

points. Fig. 1 shows a 2D velocity model. The velocity is C in the air layer and V in the subsurface layer. B is the arbitrary subsurface scattering point and A is the arbitrary transmitter or receiver. The ray path between A and B is AR_iB. R_i is the arbitrary point in the ground surface. So if the distance between A and R_i is d_{il} and the distance between R_i and B is d_{i2}, the travel time between A and B is:

$$T_i = \frac{d_{i1}}{C} + \frac{d_{i2}}{V}, \quad i = 1, \ldots, n \tag{1}$$
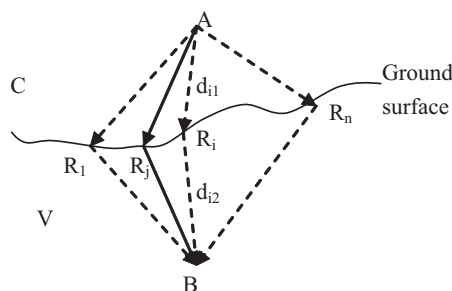
Figure 1: Sketch figure of ray path.

Then we can define the T_j in the ground surface, depending on the Fermat's principle,

$$T_j = \min(T_i) \tag{2}$$

So, the ray path between A and B is AR_jB and the first-arrival travel time is T_j. For the 3D case, we just need to extend the R_i from line to surface. Using the approach, we calculate the ray path for the experiment data of mound case.

3. CMP ANTENNA ARRAY GPR SYSTEM

The GPR system is a stepped frequency radar system with an operational frequency range of 30 KHz to 6 GHz. A vector network analyzer was chosen for its flexibility in selection of frequencies. To employ CMP technique, we developed an antenna array, shown in Fig. 2, which is constructed of 6 antipodal Vivaldi antennas. These antennas are used to form 3 pairs of transmitting-receiving antennas. In each pair of antennas, one is used to transmit the signal while the other receives. Separation between the antennas is 6 cm. Depending on the antenna array configuration of GPR system, we can acquire CMP multi-offset data directly. The antipodal Vivaldi antenna was chosen because it is a wide frequency range antenna and can easily be used to construct an antenna array due to its flat shape. A coaxial switch is used to connect the vector network analyzer to the transmitting-receiving antenna array, because while the antenna array has five transmitter and receiver antennas, the vector network analyzer has only one transmitting port and one receiving port. A position controller is used to move the antenna array in two dimensions (the X and Y direction) with precision. Based on the vector network analyzer, position controller, coaxial switch and transmitting-receiving antenna array, we developed the stepped-frequency continuous-wave array antenna system, shown in Fig. 3.
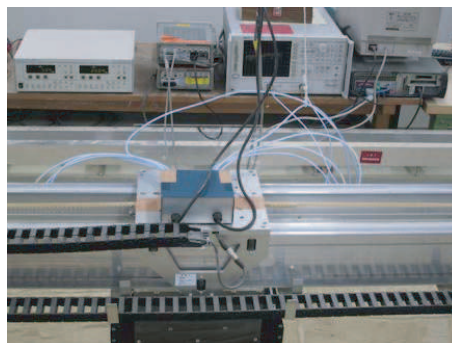

Figure 2: Antenna array.

Figure 3: SAR-GPR system.

4. EXPERIMENT

We buried a small metal ball whose radius is about 6 cm in the homogenous soil and the ground surface is mound, shown in Fig. 4, in laboratory. Then we measured it in C-scan model using GPR. The operational frequency range is from 300 MHz to 6 GHz. The number of frequency points is 401. The distance between two antennas is 6 cm. The height of antenna is 8 cm and the depth of metal ball is 10 cm. The x interval and y interval are 1 cm.
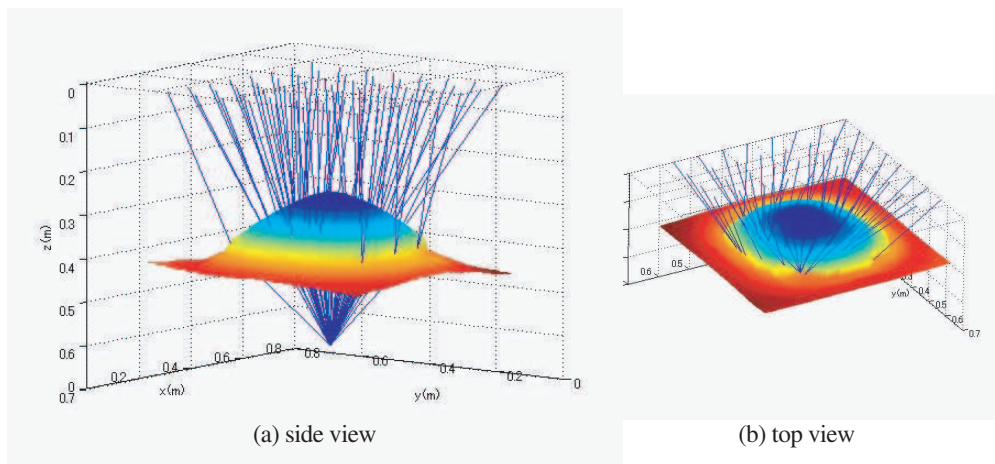

Figure 4: Metal ball.

(a) side view   (b) top view

Figure 5: The ray path in the mound.
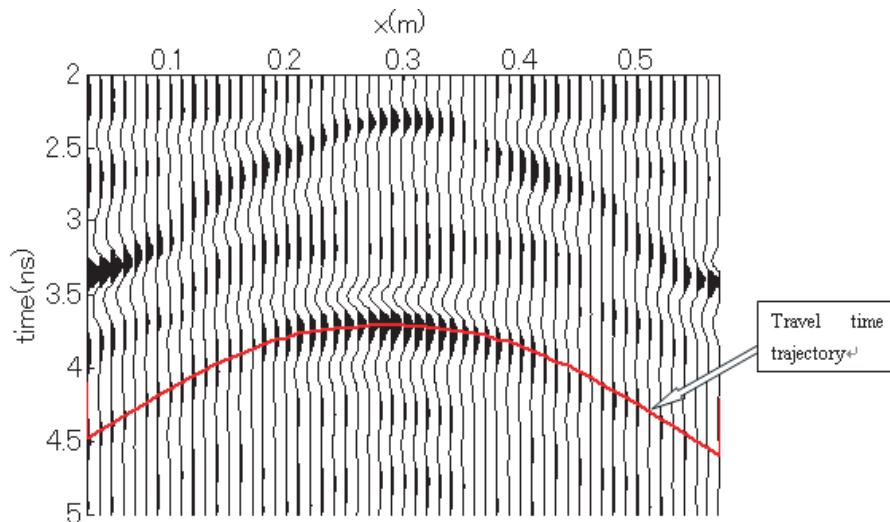
Figure 6: A profile through metal ball buried in the mound and travel time trajectory.

Using the ray tracing, we calculate the ray path for the experiment data. Fig. 5 shows the ray path among one subsurface scattering point and all antenna position in C-scan model. From the figure, we can find the ray path distributed irregularly. Fig. 6 shows one vertical profile of the survey line above the metal ball and the travel time trajectory. From the figure, we can find the travel time trajectory and the diffraction signal from the metal ball correspond each other very well. So in the situation, summing amplitudes along the travel time trajectory in stead of hyperbolic trajectory, migration can improve signal-clutter ratio and reconstruct subsurface image.

5. CONCLUSIONS

The height of ground surface varies largely in the very rough ground area, for example mound, the hyperbolic trajectory can not correspond with the diffraction signal very well. In the case, the ray tracing can offer a accurate travel time trajectory.

ACKNOWLEDGMENT

REFERENCES

1. Merwe, A. and I. J. Gupta, "A novel signal processing technique for clutter reduction in GPR measurements of small, shallow land mines," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 38, No. 6, 2627–2637, 2000.
2. Sai, B. and L. P. Ligthart, "GPR phase-based techniques for profiling rough surfaces and detecting small, low-contrast landmines under flat ground," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 42, No. 2, 318–326, 2004.
3. Walker, P. D. and M. R. Bell, "Noniterative techniques for GPR imaging through a nonplanar air-ground interface," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 40, No. 10, 2213–2223, 2002.
4. Carin, L., J. Sichina, and J. F. Harvey, "Microwave underground propagation and detection," *IEEE Trans. Microwave Theory Tech.*, Vol. 50, No. 3, 945–952, 2002.
5. El-Shenawee, M., C. Rappaport, E. L. Miller, and M. B. Silevitch, "Three-dimensional subsurface analysis of electromagnetic scattering from penetrable / PEC objects buried under rough surfaces: Use of the steepest descent fast multipole method," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 39, No. 6, 1174–1182, 2001.
6. Xu, X., E. L. Miller, C. M. Rappaport, and G. D. Sower, "Statistical method to detect subsurface objects using array ground-penetrating radar data," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 40, No. 4, 963–976, 2002.
7. Sato, M., Y. Hamada, X. Feng, F. N. Kong, Z. Zeng, and G. Fang, "GPR using an array antenna for landmine detection," *Near Surface Geophysics*, Vol. 2, No. 1, 7–13, 2004.
8. Feng, X. and M. Sato, "Pre-stack migration applied to GPR for landmine detection," *Inverse Problems*, Vol. 20, No. 6, S99–S115, 2004.
9. Yilmaz, Ö., *Seismic Data Analysis*, Society of Exploration Geophysicists, Tulsa, 2001.

FDTD Analysis of Spatial Filtering of Scattered Waves for Optical CT of Medical Diagnosis

Yasumitsu Miyazaki and Kouhei Kouno

Department of Media Informatics, Aichi University of Technology
50-2 Manori, Nishihazama-cho, Gamagori 443-0047, Japan

Abstract— Medical image diagnosis and computer aided diagnosis are modern important medical techniques developed with computer technology. Particularly, medical image diagnosis using optical waves of lasers is very important technical tools for physiological examination of human body. Transmission properties mainly depend on optical absorption effects due to biological characteristics consisting of atomic and molecular structure. Image responses of optical transmitted projection include optical scattering characteristics that disturb transmission properties through biological structures. Computer numerical simulations of spatial filtering for optical scattering superposed on transmitted and attenuated waves are discussed to improve image diagnosis. By numerical simulation of FDTD method, statistical characteristics of optical waves are discussed for optical propagation, attenuation and scattering in random inhomogeneous biological media that consist of random particle models given by random numbers programming. Spatial filtering characteristics of grid structure are shown for exact image optical projection excluding scattering effects through physiological media. Spatial filtering characteristics for off-axial scattering optical waves are graphically shown by propagation properties with large attenuation in lossy waveguide-type grids by FDTD method.

1. INTRODUCTION

Medical image diagnosis and computer aided diagnosis are modern important medical techniques developed with computer technology [1]. Particularly, medical image diagnosis using optical waves of lasers is very important technical tools for physiological examination of human body. Optical CT is effective for body tissues such as eyes, hands, children heads, and womens breasts, where transmitted signals of propagating optical waves can be received after propagation. Image responses of optical transmitted projection include optical scattering characteristics that disturb transmission properties through biological structures depending on optical absorption effects due to biological characteristics consisting of atomic and molecular structure [2–4]. In this paper, computer simulation by FDTD method of spatial filtering for optical scattering superposed on transmitted and attenuated waves are discussed to improve image diagnosis. Statistical theory of optical waves is numerically described for optical propagation, attenuation and scattering in random inhomogeneous biological media. Scattered and attenuated fields are studied by numerical computer simulation methods of FDTD [5, 6]. Spatial filtering characteristics of grid structure are shown for exact image optical projection excluding scattering effects through physiological media by FDTD method. Spatial filtering characteristics for off-axial scattering optical waves are discussed in lossy waveguide-type grids. Scattered fields of large scattering angles have large attenuations in waveguide grids and are filtered. Transmitted and scattered fields of small scattering angles have small attenuations and can pass through waveguide grids.

Statistical characteristics of optical propagation in random media consisting of biological tissues are discussed by numerical difference equation method and FDTD method, using particle models in random media. Biological tissues are consisted of random particles with several complex dielectric constants. Scattering characteristics of incident optical Gaussian beam in random physiological media consisting of biological materials such as biological cells are discussed. Electromagnetic filtering properties by optical waveguides array grids with lossy clad for off-axial scattered fields are discussed using optical waveguides consisting of lossy metal clad by FDTD method. Image recognition of biological objects surrounded by random bio-medical media, using optical lossy waveguide grid filter are discussed. Based on this computer simulation of scattered fields and waveguide grids, the optimum design of CT system may be accomplished.

2. OPTICAL CT DIAGNOSIS AND LASER FIELD IN BIOLOGICAL RANDOM MEDIA

The computer simulation by the FDTD method is studied by using the model of biological tissues in Fig. 2. The scattering characteristic I_s of the laser beam in random biological medium is given

by correlation length ρ_0, wavenumber, and index deviation $\overline{\Delta\eta^2}$. Parallel computation is required, dividing analytical area into small areas for study of optical scattering in large random media and the spatial filter characteristic by the FDTD method.
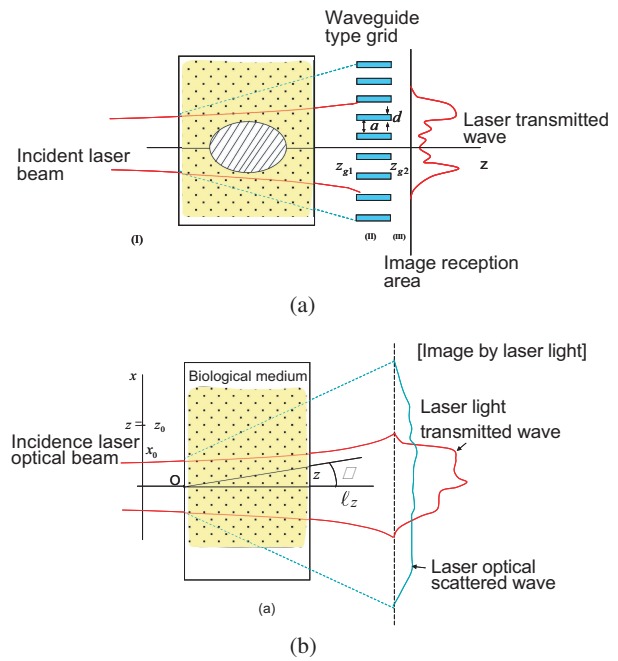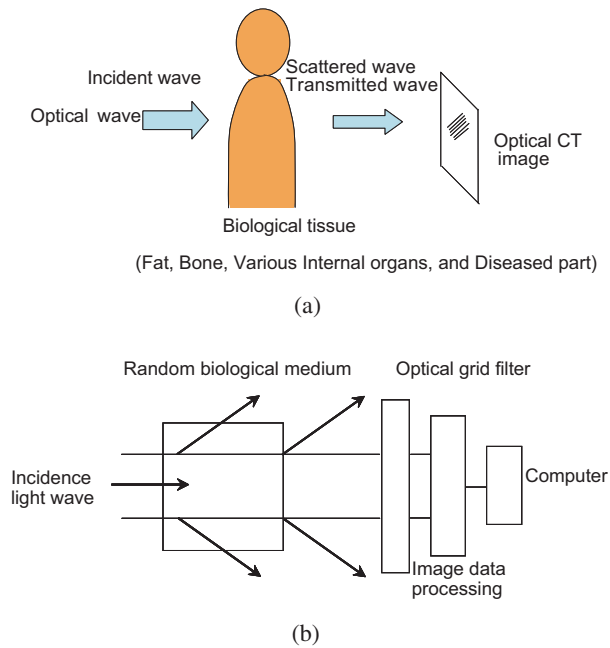

Figure 1: Optical CT system. (a) Diagnostic imaging by optical CT. (b) Computer-aided diagnosis and optical scattered wave filter system.

Figure 2: Scattering of laser beam and grid filter. (a) Laser optical beam and waveguide type grid filter in biological tissue. (b) Laser beam and scattered waves in random biological media.
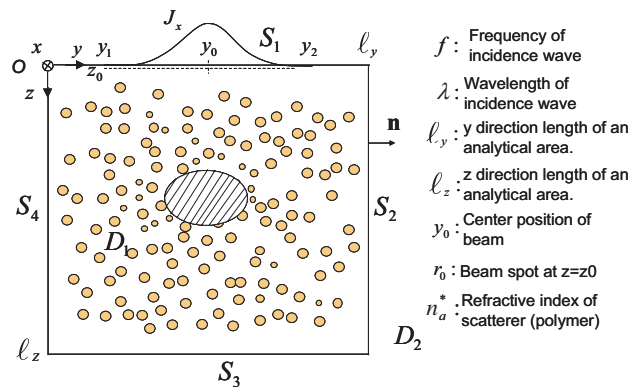

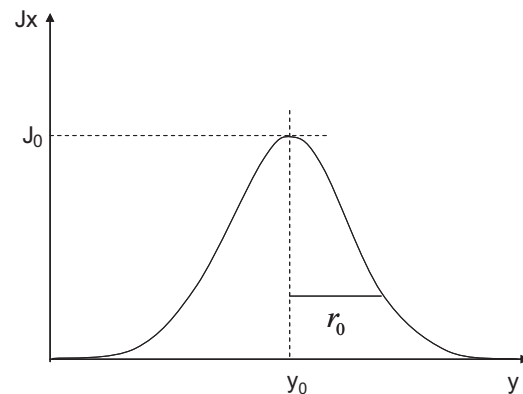Figure 3: Calculation area of FDTD method.

Figure 4: Current distribution in $z = z_0$.

3. FDTD CALCULATION OF SCATTERED FIELDS IN RANDOM MEDIA

The model of the FDTD calculation is random medium of biological tissues as shown in Fig. 3. Here, two dimensional model is used for simplicity. The electric field is assumed to be x polarized electromagnetic wave. The difference $\mathbf{E}, \mathbf{H}$ of the FDTD method for the electric and magnetic fields $\mathbf{E}, \mathbf{H}$ (1). Here, the size of the cell is assumed to be Δs and coordinates are assumed to be $y = i\Delta s$ and $z = j\Delta s$. Moreover, the time t is assumed to be $t = n\Delta t$. An analytical area is given by $0 \leq y \leq \ell_y = (Ny - 1)\Delta s$ and $0 \leq z \leq \ell_z = (Nz - 1)\Delta s$. Material constants are $\varepsilon(i, j)$ and

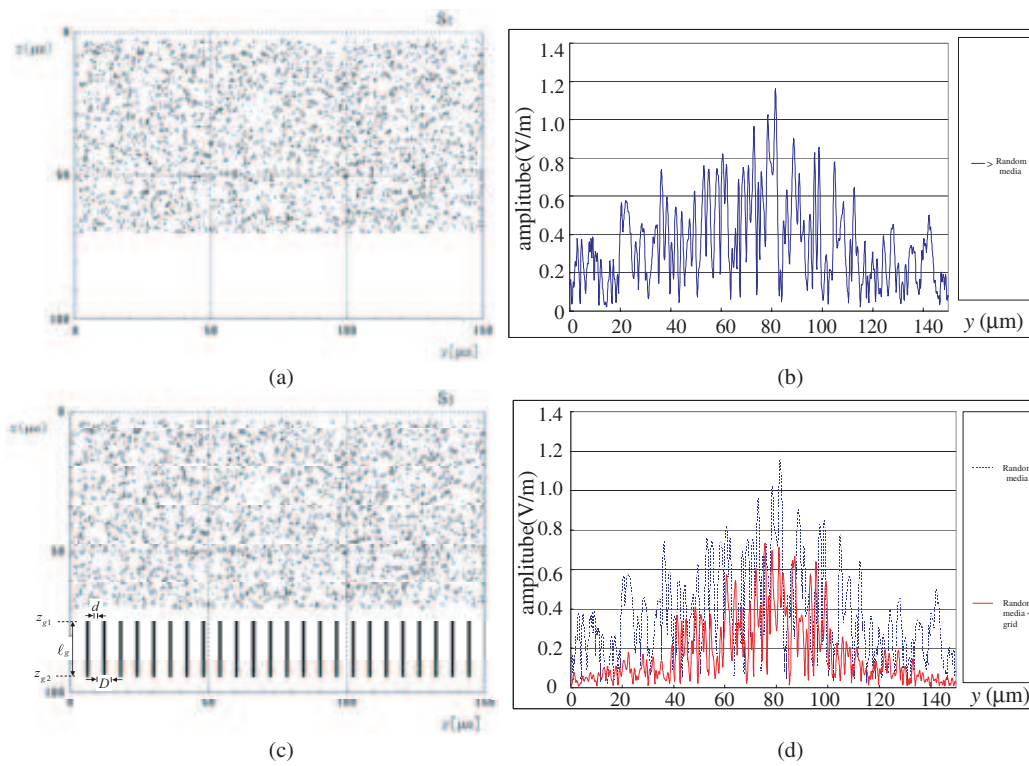Figure 5: Random media and filtering (Case I). (a) Random medium model. (b) Electric field amplitude in $z = \ell_z = 100\,\mu$m (no grid filter). (c) Random medium model and waveguide type grid. (d) Electric field amplitude in $z = \ell_z = 100\,\mu$m (with grid filter).
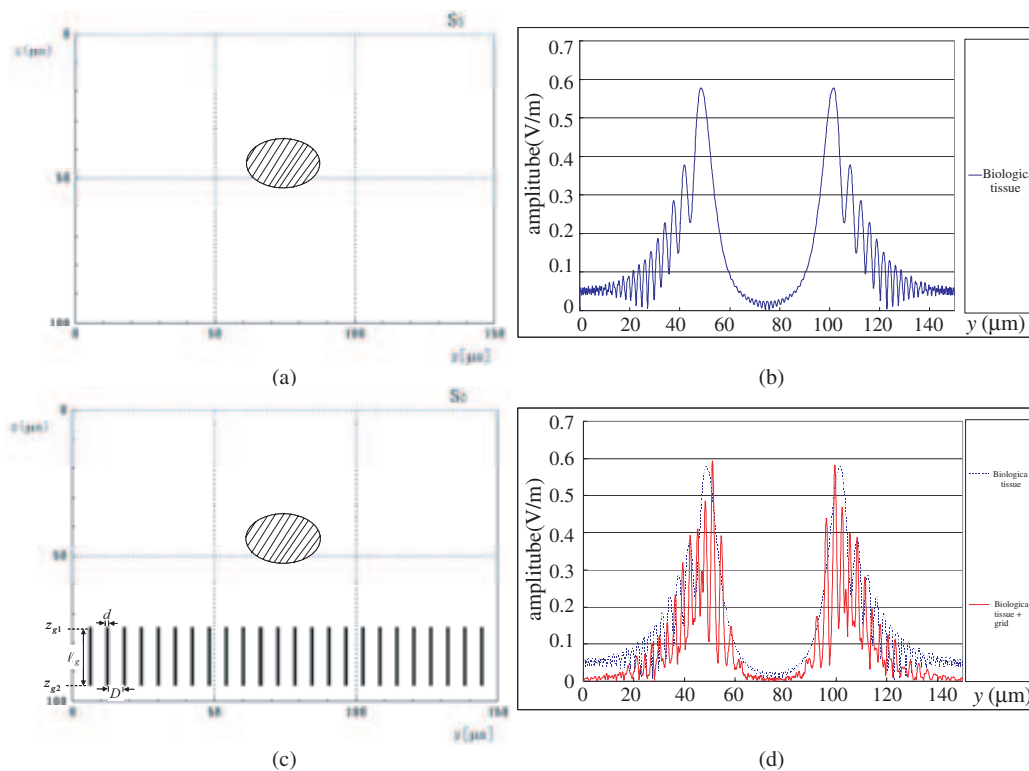

Figure 6: Biological tissues and filtering (Case II). (a) Biological tissue model. (b) Electric field amplitude in $z = \ell_z = 100\,\mu$m (no grid filter). (c) Biological tissue model and waveguide type grid. (d) Electric field amplitude in $z = \ell_z = 100\,\mu$m (with grid filter).

$\sigma(i, j)$.

$$E_x^n(i, j) = C_a(i, j)E_x^{n-1}(i, j) - \frac{\Delta t}{\varepsilon(i, j)}J_x^{n-1}(i, j) + C_b(i, j)\left\{H_z^{n-1}(i, j) - H_z^{n-1}(i-1, j)\right.$$

$$\left. -H_y^{n-1}(i, j) + H_y^{n-1}(i, j-1)\right\} \qquad (1 \leq i \leq Ny - 1, \quad 1 \leq j \leq Nz - 1)$$

$$H_y^n(i, j) = H_y^{n-1}(i, j) - D\left\{E_x^n(i, j+1) - E_x^n(i, j)\right\} \qquad (1 \leq i \leq Ny - 1, \quad 0 \leq j \leq Nz - 1) \quad (1)$$

$$H_z^n(i, j) = H_z^{n-1}(i, j) - D\left\{E_x^n(i+1, j) - E_x^n(i, j)\right\} \qquad (0 \leq i \leq Ny - 1, \quad 1 \leq j \leq Nz - 1)$$

$$C_a(i, j) = \frac{1 - \sigma(i, j)\Delta t/(2\varepsilon(i, j))}{1 + \sigma(i, j)\Delta t/(2\varepsilon(i, j))} \qquad C_b(i, j) = \frac{\Delta t/\varepsilon(i, j)}{1 + \sigma(i, j)\Delta t/(2\varepsilon(i, j))}\frac{1}{\Delta s}$$
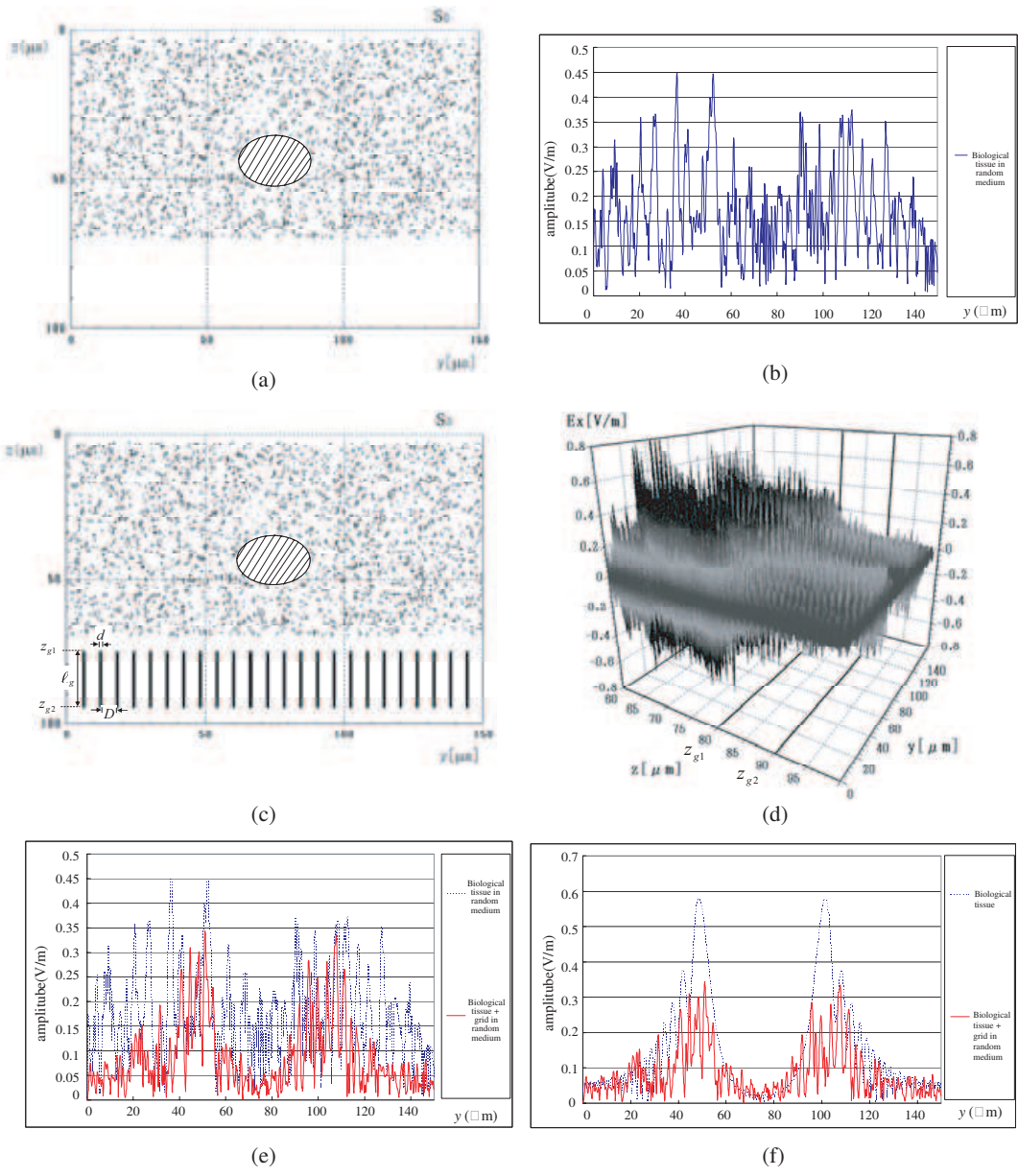

(a)

(b)

(c)

(d)

(e)

(f)

Figure 7: Biological tissues in random media and filtering (Case III). (a) Biological tissue model in random medium (no grid filter). (b) Electric field amplitude in $z = \ell_z = 100\,\mu\text{m}$ (no grid filter none). (c) Random media, the biological tissue model and the waveguide type grids. (d) Electric field amplitude. (e) Electric field amplitude in $z = \ell_z = 100\,\mu\text{m}$ (with grid filter). (f) Electric field amplitude (with grid filter) in $z = \ell_z = 100\,\mu\text{m}$. The comparison of Case II (no grid) and Case III (with grid).

where

$$D = \frac{\Delta t}{\mu_0 \Delta s}$$

Δs: cell size, Δt: time step.

A boundary initial condition of the incidence wave is equivalently given by the current J_x, in $z = z_0 = \Delta s$ $(j = 1)$ shown in Fig. 4.

$$J_x^n(i, 1) = J_0 \exp\left\{ -\left(\frac{(i - i_0)\Delta s}{r_0} \right)^2 \right\} H(n\Delta t) \sin(2\pi f n \Delta t) \tag{2}$$

Here, τ is a pulse width, and

$$H(n\Delta t) = \begin{cases} 1, & n\Delta t \le \tau \\ 0, & n\Delta t < 0, \quad n\Delta t > \tau \end{cases}$$

f and λ are frequency and wavenumber, and further, the beam center is $y_0 = i_0 \Delta s = \ell_y/2$, and the beam spot is $r_0 = 30\lambda$.

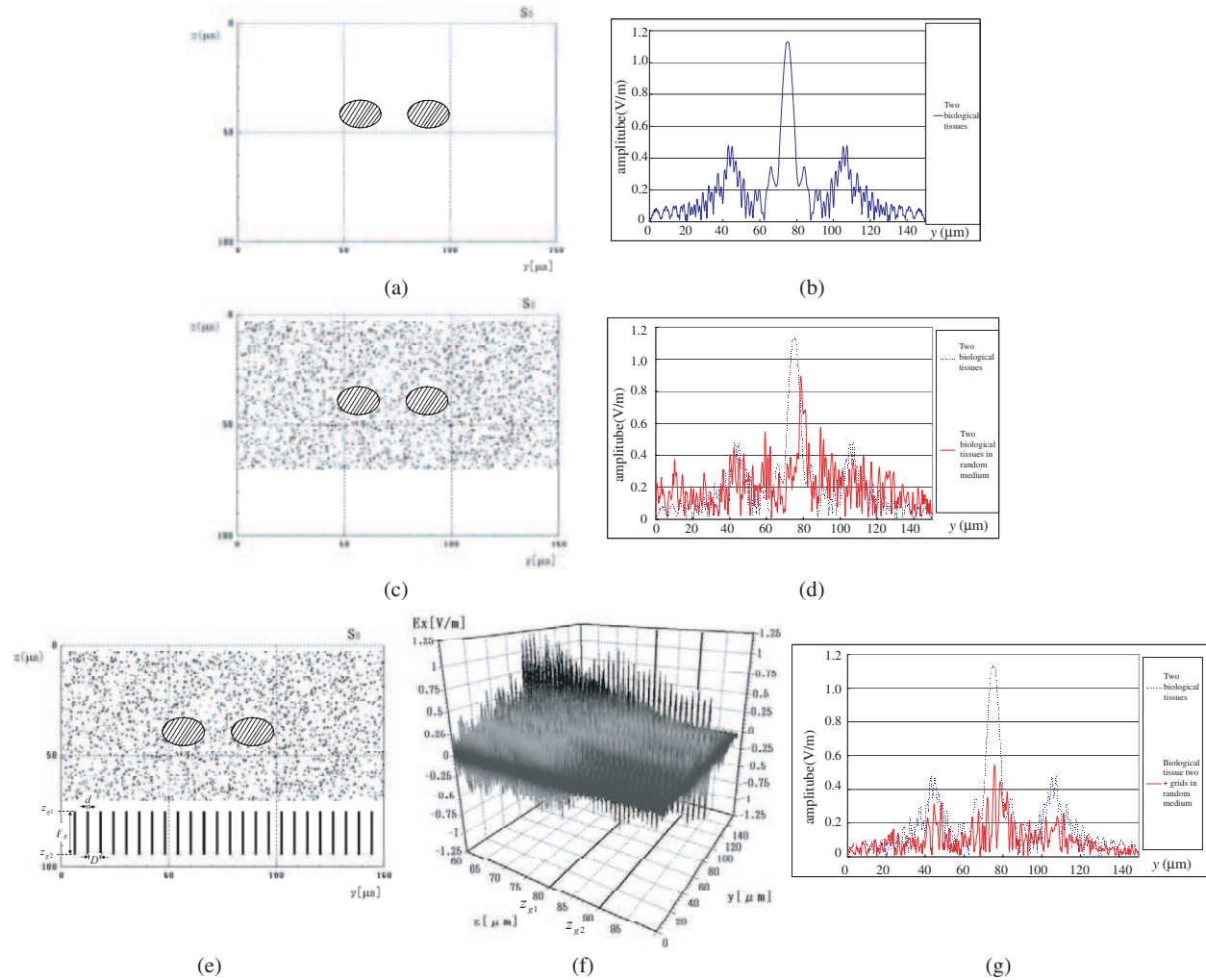Parameters of FDTD calculation are shown in Table 1.

Figure 8: Two biological tissues in random media and filtering (Case IV). (a) Two models of biological tissue. (b) Electric field amplitude in $z = \ell_z = 100\,\mu m$ when two biological tissues exist. (c) Two models of biological tissue in random medium (no grid filter). (d) Electric field amplitude in $z = \ell_z = 100\,\mu m$ when two biological tissues in random media exist (grid filter none). (e) Two biological tissue in random medium model and waveguide type grids. (f) Electric field amplitude. (g) Electric field amplitude and random media in $z = \ell_z = 100\,mm$ when two biological tissues existed (with grid filter).

4. NUMERIC CALCULATION RESULTS OF SCATTERING AND SPATIAL FILTER CHARACTERISTICS

Figures 5–8 show scattering fields by random biological medium and spatial filtering characteristics. From computer simulation results by FDTD methods shown in Figs. 5–8, it is shown that scattering fields that disturb biological diagnosis information given by transmitted fields due to absorption effects in biological cells are filtered by optical grids with lossy waveguide arrays.

Table 1: Numerical values of scattering and attenuation characteristic of laser beam wave.

Parameter	Numerical value
f: Frequency of incidence wave	$300\,\text{THz}$
λ_0: Wavelength of incidence wave	$1\,\mu\text{m}$ (c/f)
ℓ_y: y direction length of an analytical area.	$150\,\mu\text{m}$ $(150\lambda_0)$
ℓ_z: z direction length of an analytical area.	$100\,\mu\text{m}$ $(100\lambda_0)$
y_0: Center position of beam (y coordinates)	$75\,\mu\text{m}$ $(\ell_y/2 = 75\lambda_0)$
r_0: Beam spot at $z = z_0$	$30\,\mu\text{m}$ $(r_0/\lambda_0 = 30)$
Δs: Size of space division cell	$0.05\,\mu\text{m}$ $(\Delta s/\lambda_0 = 1/20)$
Δt: Time increment	$1.15 \times 10^{-16}\,\text{s}$
N: Number of scatterers (polymer) in analytical area	2300
a: Length of about one of scatterer (polymer)	$0.5 \sim 1.0\,\mu\text{m}$ $(a/\lambda_0 = 0.5 \sim 1.0)$
n^*: Refractive index of scatterer (polymer)	1.1
n_g^*: Refractive index of grid wall	$2.0 - j0.5$
d: Width of grid wall	$1\,\mu\text{m}$ $(d/\lambda_0 = 1)$
D: Interval of grid wall	$5\,\mu\text{m}$ $(D/\lambda_0 = 5)$
ℓ_g: Length of grid wall	$20\,\mu\text{m}$ $(\ell_g/\lambda_0 = 20)$

5. CONCLUSIONS

For optical CT in medical diagnosis using laser, spatial filtering of scattered waves by waveguides with lossy cladding is very useful to obtain precise image processing. Scattering characteristics in random bio-medical media consisting of bio-molecules, are studied by computer simulation by FDTD method for electric field. Attenuation properties in the spatial filter grids of scattered fields of large scattering angles are investigated and spatial filtering characteristics of scattered fields by computer FDTD numerical calculations are discussed for improvement of precise diagnosis.

REFERENCES

1. Miyazaki, Y., "Spatial filtering of optical scattered waves in bio-medical media by inhomogeneous waveguide grids for optical CT," *Proc. of APMC 2007*, 527–530, Bangkok, 2007.
2. Miyazaki, Y., "Light scattering of laser beams by random micro-inhomogeneities in glasses and polymers," *Jpn. Appl. Phys.*, Vol. 13, 1238–1248, 1974.
3. Miyazaki, Y., "Partially coherent optical waves in random gradient fibers," *Optical and Quantum Electronics*, Vol. 9, 153–165, 1977.
4. Miyazaki, Y., "Electromagnetic characteristics of grid structures for scattering fields of nanometer electromagnetic waves and X-rays," *Proceedings of PIERS*, 543–547, Tokyo, 2006.
5. Miyazaki, Y. and K. Takahashi, "Computer simulation of X-ray scattering characteristics for medical image diagnosis," *Trans. IEE Japan*, Vol. 126-C, No. 12, 1431–1440, 2006.
6. Takahashi, K., Y. Miyazaki, and N. Goto, "FDTD parallel computational analysis of grid-type scattering filter characteristics for medical X-ray image diagnosis," *IEE J. Trans. EIS*, Vol. 127, No. 12, 1973–1981, 2007.

FDTD Analysis of Microwave Propagation and Scattering Characteristics over Forests for WiMAX Wireless Communications

Yasumitsu Miyazaki and Tatsutoshi Ikeda

Department of Media Informatics, Aichi University of Technology
50-2 Manori, Nishihasama-cho, Gamagori 443-0047, Japan

Abstract— WiMAX wireless communication has been rapidly developed for broadband mobile communication including image transmission. To design excellent high performance mobile communication systems of very high speed transmission, accurate environment evaluation of communication systems is indispensable. In the past, building and street effects on high speed signal propagation have been studied by FDTD method. However, wave propagation and scattering characteristics by presence of forest and trees have not been so much studied. Using several material consants and structure, the effects of multiple scattering and attenuation of microwaves over forests in WiMAX wireless communications are analyzed numerically using FDTD method. These results may yield important factors for the optimum design of high performance and reliable WiMAX communication systems.

1. INTRODUCTION

In recent years, WiMAX wireless communication has been rapidly developed for broadband mobile communication of image and TV transmission. Mobile WiMAX communication system uses microwave carrier of 2.5 GHz frequency band and modulation system is mainly OFDM for transmission of signals. By using OFDM technique, WiMAX provides high speed and reliable communication against the multi pass interference due to the presence of obstacles in communication channels. To design excellent high performance wireless communication systems, accurate evaluation of communication channels is indispensable. By using computer numerical simulation of parallel FDTD method, we studied fundamental microwave propagation and scattering characteristics in urban area [1, 2]. In these analyses, building and street effects on high speed signal propagation have been investigated by computer simulation of FDTD method. However, wave propagation and scattering characteristics by presence of forest and trees are not so much studied [3–5]. The effects of multiple scattering and attenuation of microwave by forest are severe factors of high speed wireless communications. Size of branches and leaves of trees are comparable with microwave wavelength and yield strong interaction phenomena between microwaves and trees of broadband propagation.

In this paper, propagation and scattering characteristics of microwave over forests in WiMAX wireless communication is numerically analyzed using FDTD method. FDTD method can be applied for signal and noise analysis about several different complex models and inhomogeneous materials such as forests in communication channels. By applying FDTD method for the analysis of communication channels of propagation over trees and forests, numerical simulations of signal propagation for various tree structures at different frequencies and digital signal bit rates are demonstrated. In FDTD simulations, the incident wave is assumed to be a traveling wave from transmission antenna station at a far distance. Numerical results in this paper show and analyze the effects of various structures of forests and beam spot size of incident waves. The shape of forest is assumed to be constructed by random surface and distribution of branches of trees are inhomogeneous. Forest and trees in radio communication channel is considered to be as random
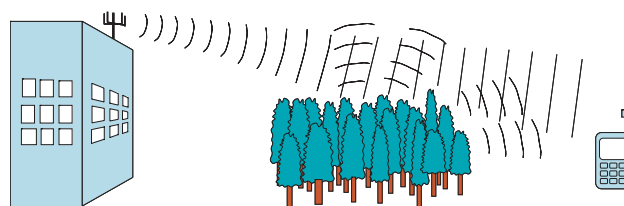

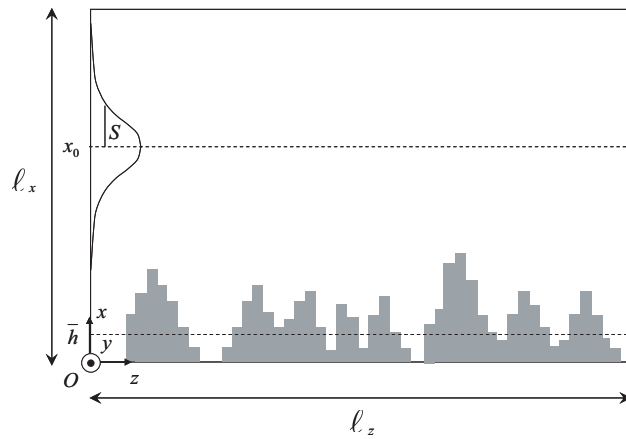Figure 1: Microwave propagation in mobile WiMAX wireless communication.

Figure 2: Two dimensional analysis model of FDTD method.

surface and inhomogeneous materials. Fig. 1 shows the situation for microwave propagation in mobile WiMAX wireless communiction over forests.

The effects of multiple scattering and attenuation due to forests are evaluated using statistical functions, such as avrage, variance and correlation of conductivities and permittivities for trees. These results may yield important factors for design of high performance and more reliable WiMAX communication systems.
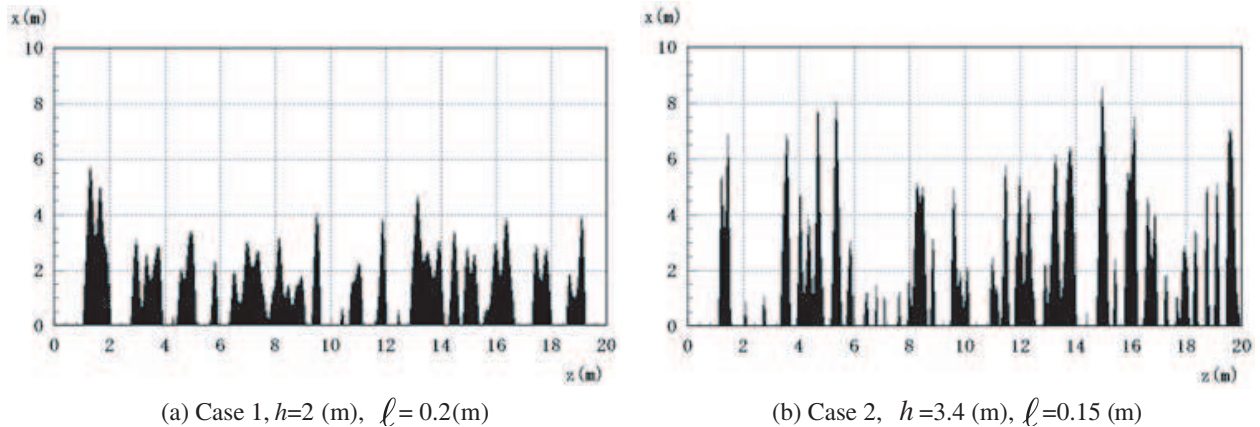

(a) Case 1, h=2 (m), ℓ= 0.2(m)

(b) Case 2, h=3.4 (m), ℓ=0.15 (m)

Figure 3: Forest models by random surface.

(a) Case 1-1-(a), S =1.6 (m)

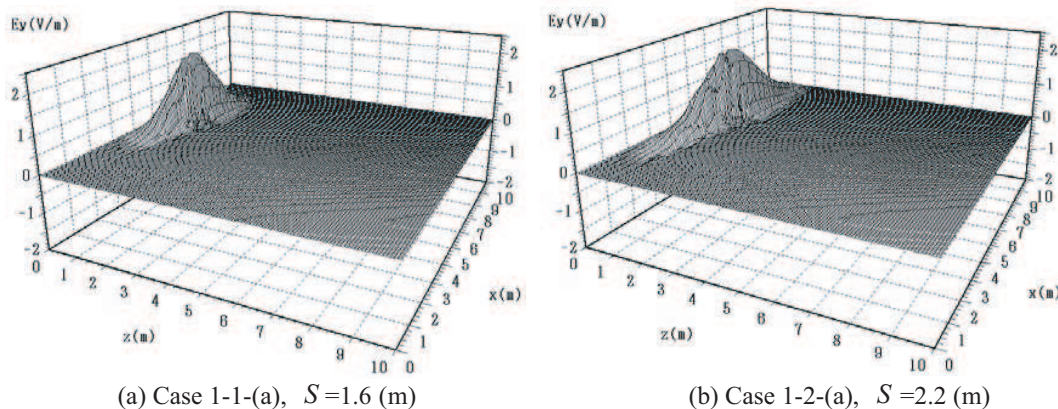(b) Case 1-2-(a), S =2.2 (m)

Figure 4: Electric field of the incident wave at $t = 400\Delta t = 8$ (ns).

2. FDTD ANALYSIS OF MICROWAVE PROPAGATION AND SCATTERING

Microwave propagation and scattering by forests are very important phenomena for wireless broadband communication such as mobile WiMAX communication. Computer simulation using FDTD method is very useful to evaluate these characteristics numerically. Two-dimensional analysis model for microwave propagation and scattering characteristics over forests is shown in Fig. 2. Analysis region is defined as $\ell_x \times \ell_z$. In FDTD simulations, the incident wave is assumed to be a traveling wave from transmission antenna station at a far distance. The incident wave is y-polarized Gaussian beam wave with angular frequency $\omega = 2\pi f$, beam waist $z = z_0$ and beam spot size S. In the simulation model, the electromagnetic fields are point (i, j) at time $n\Delta t$ are calculated by difference equations, as Eq. (1).
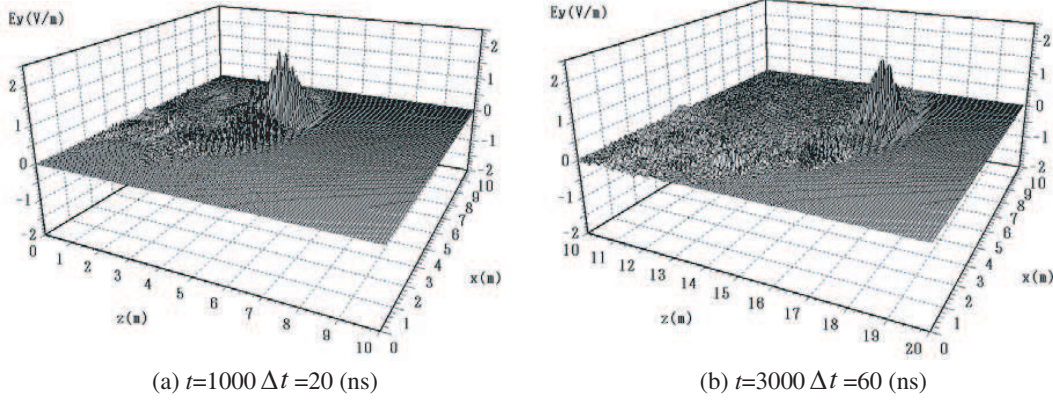

(a) $t=1000\,\Delta t = 20$ (ns)   (b) $t=3000\,\Delta t = 60$ (ns)

Figure 5: Electric field E_y in Case 1-1-(a), $S = 1.6$ (m).

(a) $t=1000\,\Delta t = 20$ (ns)   (b) $t=3000\,\Delta t = 60$ (ns)
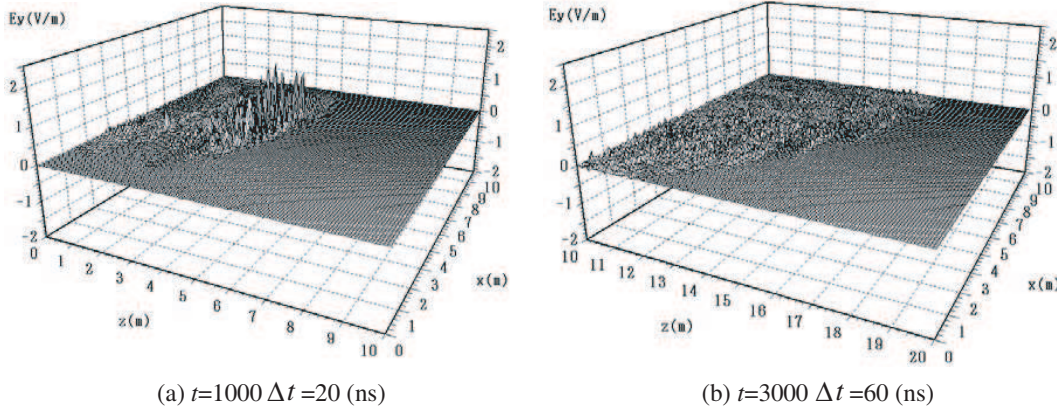
Figure 6: Electric field E_y in Case 2-1-(a), $S = 1.6$ (m).

Parallel computation of FDTD using grid computer can be proceeded for these analysis of large area propagation. The electromagnetic fields in each divided subdomain $D_{u,v}$ $(1 \leq v \leq My)$ are calculated by parallel processing. To proceed the parallel processing, data transfer between adjacent subdomains is carried out using MPI.

$$E_y^n(i,j) = c_1 E_y^{n-1}(i,j) + c_2 J_y^{n-1}(i,j) + c_3 \left\{ H_x^{n-1}(i,j) - H_x^{n-1}(i,j-1) - H_z^{n-1}(i,j) + H_x^{n-1}(i-1,j) \right\}$$

$$H_x^n(i,j) = H_x^{n-1}(i,j) + c_4 \left\{ E_y^n(i,j+1) - E_y^n(i,j) \right\}, \quad H_z^n(i,j) = H_z^{n-1}(i,j) - c_4 \left\{ E_y^n(i+1,j) - E_y^n(i,j) \right\} \quad (1)$$

$$J_y^n(i,j) = \begin{cases} \neq 0 & (j=1) \\ = 0 & (j \neq 1) \end{cases}, \quad c_1 = \frac{1 - \sigma\Delta t/(2\varepsilon)}{1 + \sigma\Delta t/(2\varepsilon)}, \quad c_2 = \frac{\Delta t/\varepsilon}{1 + \sigma\Delta t/(2\varepsilon)}, \quad c_3 = \frac{c_2}{\Delta s}, \quad c_4 = \frac{\Delta t}{\mu_0 \Delta s}$$

Here, x, z and t are discretized by $x = i\Delta s$, $z = j\Delta s$ and $t = n\Delta t$ and $0 \leq i \leq N_x$, $0 \leq j \leq N_z$.
    The incident wave is generated by current density

$$J_y^n(i,1) = J_0 \exp\left\{ -\left(\frac{i\Delta s - x_0}{S}\right)^2 \right\} \exp\left\{ -\left(\frac{n\Delta t - t_0}{T}\right)^2 \right\} \cdot \sin(2\pi f n \Delta t) \quad (2)$$

where f is the frequency of incident wave, x_0 is the center point of incident beam, S is the beam spot size at $z = z_0$ $(j = 1)$, T is the parameter for transmission pulse width. Mur's boundary condition is applied to obtain the electromagnetic fields at the boundary of analysis space.

In the analysis model, the shape of trees and forests are realized by random surfaces. Random surface is generated by giving rms height $\sqrt{\overline{\Delta h^2}}$ and correlation length ℓ. In Fig. 2, $\overline{h}$ is the average height of random surfaces.

| (a) Case 1-2-(a) | (b) Case 1-2-(b) |
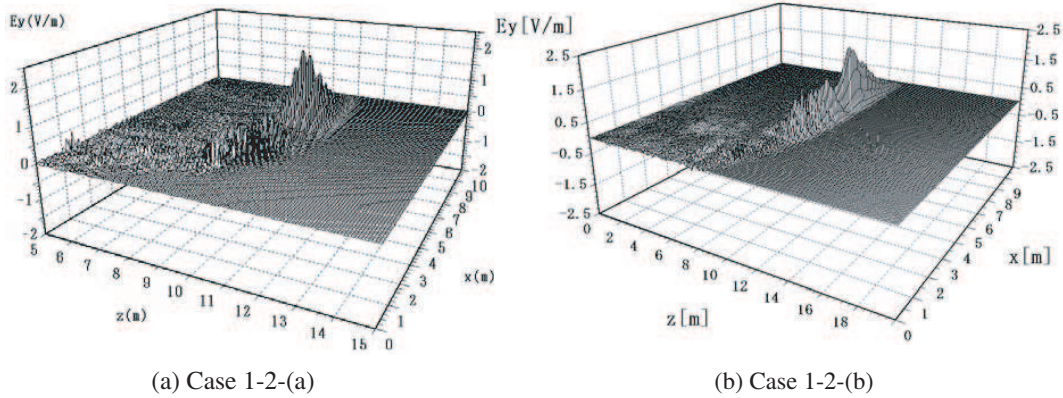
Figure 7: Electric field E_y, $S = 2.2$ (m), $t = 2000\Delta t = 40$ ns.

| (a) Case 2-2-(a) | (b) Case 2-2-(b) |
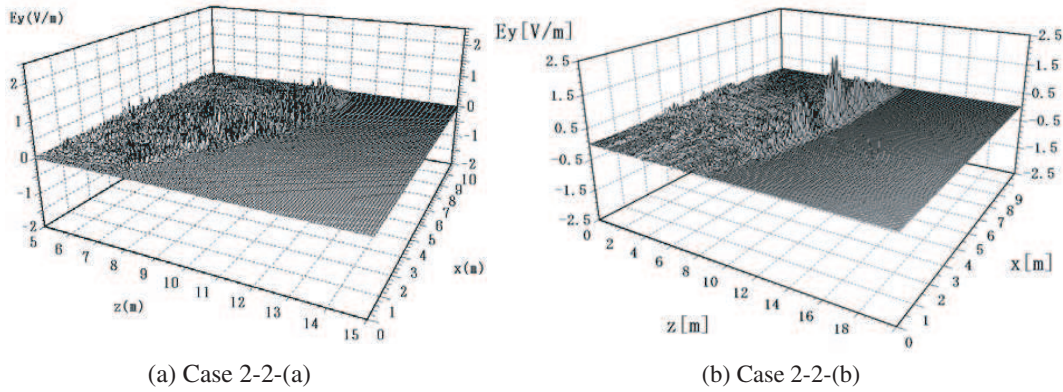
Figure 8: Electric field E_y, $S = 2.2$ (m), $t = 2000\Delta t = 40$ ns.
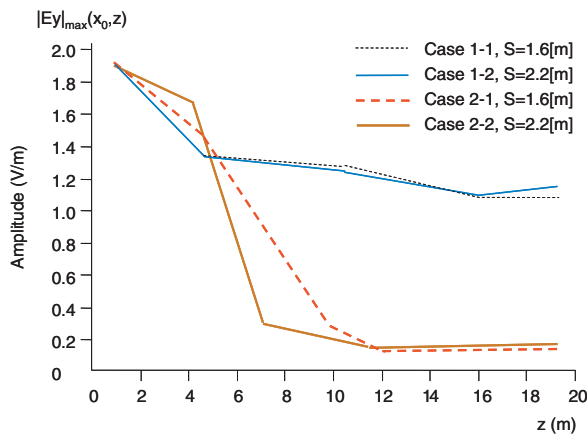
Model		$\overline{h}$ (m)	$\sqrt{\overline{\Delta h^2}}$ (m)	ℓ (m)	S (m)	α (dB/m)
Case	1-1	1.0	1.0	0.20	1.6	0.25
	1-2				2.2	0.22
Case	2-1	1.0	2.4	0.15	1.6	1.12
	2-2				2.2	1.05

$$\alpha / z = -20 \log \frac{|E_y(z)|}{|E_y(0)|} \quad \text{(dB/m)}$$

$$z = \ell_z \text{ (m)}$$

Figure 9: Amplitude attenuation characteristics $x = x_0 = 6$ (m) (Case 1, 2) (a) $\varepsilon_r = 4.0$, $\sigma = 10^{-3}$ S/m.

3. PROPAGATION AND SCATTERING CHARACTERISTICS OF MICOWAVES OVER FORESTS

Propagation and scattering characteristics of Gaussian beam waves with microwave frequency are evaluated using FDTD method. In FDTD simulation, $\ell_x = 1000\Delta s = 10\,(\mathrm{m})$, $\ell_z = 2000\Delta s = 20\,(\mathrm{m})$, $x_0 = 6\,(\mathrm{m})$ and $\overline{h} = 1\,(\mathrm{m})$ are used. For incident wave, $f = 2.5\,\mathrm{GHz}$, $T = 2\,(\mathrm{ns})$, $t_0 = 5\,(\mathrm{ns})$ and $J_0 = 1\,(\mathrm{A/m^2})$ are used. Here, wavelength λ is $0.12\,(\mathrm{m})$ and size of analysis space ℓ_x and ℓ_z become 83.3λ and 166.7λ, respectively. Numerical parameters for FDTD analysis are shown in Table 1. Fig. 3 shows the forest model generated by giving rms height $\sqrt{\overline{\Delta h^2}}$ and correlation length ℓ of random surface as shown in Table 1. Fig. 4 shows incident fields of Gaussian beams in these models.

Table 1: Numerical parameters for FDTD analysis.

Parameters	Values		
f : Frequency of incident wave	2.5 GHz		
λ : Wavelength of incident wave	0.12 m		
ℓ_x: Length of analysis space (x)	10 m (83.3 λ)		
ℓ_z: Length of analysis space (z)	20 m (166.7λ)		
x_0 : Center point of the beam (x)	6 m		
S : Beam spot size	1.6 m (13.3λ), 2.2 m (18.3λ)		
t_0 : Peak time of incident pulse	5 ns		
T : Parameter for pulse width	2 ns		
Δs: Length of a cell	0.01 m		
Δt : Time increment	0.02 ns		
ε_r : Relative dielectric constant of trees	(a) 4		(b) 2
σ : Conductivity of forests	10^{-3} S/m		10^{-5} S/m
$\overline{h}$: Average height of forests	1.0 m		
$\sqrt{\overline{\Delta h^2}}$: rms height of forest	2.0 m (Case 1) 3.4 m (Case 2)		
ℓ : Correlation length of forests	0.20 m (Case 1) 0.15 m (Case 2)		

In Case 1 shown in Figs. 5 and 7, the transmitted signal keeps strong intensity above the height of 6 (m). However, the transmitted signal in Case 2 shown in Fig. 6 is highly attenuated due to the scattering by forests. Particularly, Fig. 6 shows the strong multiple scattering because the correlation length of forest model is comparable to the incident wavelength. Figs. 7 and 8 are the results with different beam spot size of the incident waves. In Figs. 8(a) shows that the attenuation of transmitted signal is still strong and the field distribution is disturbed due to the presence of tall trees. Therefore, in this case, it seems difficult to accomplish reliable com munication between the transmitter and receivers. By using statistical function, such as average, variance and correlation of random surface and time waveform of received electric field at receiving point, signal and noise characteristics in broadband mobile communication can be evaluated, as shown in Fig. 9.

4. CONCLUSIONS

Microwave propagation and scattering by forests are very important phenomena for wireless broadband communication. In this paper, propagation and scattering characteristics of microwave over forests in WiMAX wireless communication is numerically analyzed using FDTD method. Using statistical characteristics, FDTD method can be applied for signal and noise analysis about several different complex models and inhomogeneous materials such as forests in communication channels. By applying FDTD method for the analysis of communication channels of propagation over trees and forests, numerical simulations of signal propagation for various tree structures at different frequencies and digital signal bit rates are shown, using statistical shape parameters ad material constants.

REFERENCES

1. Selormey, P. and Y. Miyazaki, "Electromagnetic compativility characteristics of buildings in mobile radio waves propagation channel," *Trans. IEE Japan*, Vol. 119-C, No. 1, 97–104, 1999.

2. Rodriguez, G., Y. Miyazaki, and N. Goto, "Matrix-based FDTD parallel algorithm for big areas and its applications to high-speed wireless communications," *IEEE Trans. Antennas & Propagat.*, Vol. 54, No. 3, 785–796, 2006.

3. Masuda, T., Y. Miyazaki, and Y. Kashiwagi, "Analysis of electromagnetic wave propagation in out-door active RFID system using FD-TD method," *PIERS Online*, Vol. 3, No. 6, 937–939, 2007.

4. Miyazaki, Y., T. Takada, and K. Takahashi, "Propagation and scattering characteristics of microwaves over forests in WiMAX wireless communications using FDTD method," *PIERS Proceedings*, 504–508, Cambridge, USA, July 2–6, 2008.

5. Miyazaki, Y., T. Takada, and H. Tomita, "Propagation and scattering characteristics of microwaves over forests in WiMAX wireless communications using FDTD method," *IEICE Tech. Rept. MW2008-31*, 1–6, 2008.

An Accelerated Frequency Domain Ray-tracing Simulator for Ultra-Wideband Communications

J. Diskin[1], **A. Alomainy**[2], **and C. Brennan**[1]

[1]Research Institute for Networks and Communications Engineering (RINCE)
School of Electronic Engineering, Dublin City University, Ireland
[2]Antennas & Electromagnetics Research Group
Queen Mary University London, United Kingdom

Abstract— Ray-tracing is an established technique for modelling wireless propagation over a large area in an efficient manner. Existing frequency domain ray tracing simulators incur severe computational burden when extended to model an Ultra-Wideband (UWB) channel of several Giga-Hertz. We propose an acceleration algorithm that improves the efficiency of such simulators by approximating the frequency domain response of individual ray at a reduced set of frequency points. Result show time savings up to 75% with a maximum relative error in the region of 0.8%.

1. INTRODUCTION

The sanctioning of Ultra-Wideband (UWB) as a wireless communications medium by the Federal Communication Commission (FCC), constitutes a potential for development of novel capabilities in wireless applications [1]. In particular the large bandwidth utilized by UWB offers unprecedented data-rates and enhancements to applications such as wireless geo-location [2]. With the arrival of wireless schemes such as IEEE $802.11b/g$, it became apparent that propagation modelling was required for purposes of optimising performance and planning infrastructure in varied environments. Ray tracing is one such modelling method that has came to prominence [3]. While ray-tracing is an asymptotic method and may not be as computationally accurate as full wave methods, it offers a deterministic but efficient means of propagation modelling in large-scale environments at high frequencies.

Ray-tracing can be extended to model UWB wireless systems. However there are computation penalties incurred when using ray-tracing simulators based on frequency domain computations [4], due to the large bandwidths involved. Some ray tracing techniques rely on time-domain based methods to avoid having to calculate channel responses at every frequency [5]. These techniques do not consider the variations of material properties over the extensive UWB bandwidth. We present an acceleration algorithm that reduces the computation associated with frequency domain ray-tracing computation. This algorithm accounts for the effects of frequency dependant material properties. In addition, a comparison of a ray tracing simulation with empirical measurements for an indoor UWB channel is presented.

2. FREQUENCY DOMAIN RAY TRACING

Electric fields at specified receiver points are computed by a Ray-Tracing simulator. During such computations dominant paths by which energy propagates between the receiver and transmitter are identified. Subsequently the electromagnetic fields associated with these paths are calculated and summed to obtain the total electric field. At a point **r** the electric field can be written as,

$$\mathbf{E}\left(\mathbf{r}, \omega\right) = \sum_{i=1}^{N} \mathbf{R}_i\left(\mathbf{r}, \omega\right), \tag{1}$$

where N is the total number of rays and depends on the highest order of scattering event that is allowed. The variation of the electric field with frequency in Equation (1) is due to the frequency dependency of each ray, which in turn can be classed into two categories.

- An substantial contribution to frequency dependence is due to the propagation distance of a ray. Electric lengths of a ray will differ at each individual frequency. Although this results in a phase component that varies rapidly with frequency, the effect can be easily computed.

- Additional frequency dependent effects are attributed to interactions of ray with object in the propagation environment. Transmission and reflection coefficients of such object's surfaces undergo variations due to the frequency dependance of the constituent material's dielectric properties. Consequently, individual rays also exhibit frequency dependant variations. Such effects, while significant, vary slowly with frequency.

Our technique relies on separating these widely varying frequency dependencies for each ray. Fast variations are extracted and the remaining slowly varying effects are approximated using a low order polynomial. Using this method a channel response can be completely specified by analysing the ray at a reduced number of frequency sample points.

3. EFFICIENT FREQUENCY SWEEP

A channel response at a point $\mathbf{r}$ is given by,

$$H\left(\mathbf{r}, \omega\right) = \sum_{i=1}^{N} \mathbf{R}_i\left(\mathbf{r}, \omega\right), \tag{2}$$

where $H(\mathbf{r}, \omega)$ is the component of electric field at the point $\mathbf{r}$ in the direction determined by the polarisation of the receiver antenna. Typically this quantity must be computed for many frequency points ω_γ for $\gamma = 1, \ldots, Q$, choosing Q large enough to ensure a satisfactory sampling rate. Hence a computation burden occurs when the huge bandwidths of UWB channels are simulated. The proposed acceleration technique reduces the amount of frequency samples that must be explicitly computed by separating the fast and slow varying components of each ray. For each point $\mathbf{r}$ we rewrite Equation (2) as (suppressing the $\mathbf{r}$ dependence for notational convenience)

$$H(\omega) = \sum_{i=1}^{N} A_i\left(\omega\right) e^{-\jmath\phi_i(\omega)}, \tag{3}$$

where we have made explicit the amplitude and phase frequency dependence of the ray contribution. This can be factorized as follows,

$$H(\omega) = \sum_{i=1}^{N} A_i\left(\omega\right) e^{-\jmath\tilde{\phi}_i(\omega)} e^{-\jmath\phi_i^{(d)}(\omega)}, \tag{4}$$

where $\phi_i^{(d)}$ is the phase behaviour associated with the free-space propagation of the ray and $\tilde{\phi}_i$ is the residual phase behaviour associated with its interactions with materials.

The phase term $\phi_i^{(d)}$ has the form,

$$\phi_i^{(d)}\left(\omega\right) = \frac{d_i\omega}{c}, \tag{5}$$

where d_i is the distance ray i travels in free space and c is the speed of light in a vacuum. By dividing out this linear phase, a slowly varying term $A_i\left(\omega\right)\tilde{\phi}_i\left(\omega\right)$ remains. Consequently this residual can be approximated by P_i, a low order polynomial of order m,

$$A_i\left(\omega\right) e^{-\jmath\tilde{\phi}_i^{(d)}(\omega)} \simeq P_i(\omega), \tag{6}$$

yielding,

$$H\left(\omega\right) \simeq \sum_{i=1}^{N} e^{-\jmath\phi_i^{(d)}(\omega)} P_i\left(\omega\right). \tag{7}$$

P_i can be completely specified for each ray by computing exactly the channel response at $m+1 \ll Q$ frequency points.

A received signal $r(t)$ can be obtained by multiplying the channel response by $S(\omega)$, the Fourier transform of the transmitted signal $s(t)$, and taking an inverse Fourier transform,

$$r(t) = \mathcal{F}^{-1}\left(S\left(\omega\right) H\left(\omega\right)\right). \tag{8}$$

However, we present an alternative analytical approach that uses the polynomial approximation P_i and the ϕ_i^d free-space phase term to rapidly compute $H(\omega_\gamma)$ at each of the Q frequency points ω_γ. To achieve this we exploit the following identities of Fourier transforms,

$$\mathcal{F}^{-1}\{f^g S(\omega)\} = \frac{d_g}{dt^g} s(t). \tag{9}$$

$$s\left(t - \phi_i^{(d)}\right) = \mathcal{F}^{-1}\left\{S(\omega)e^{-\jmath\phi_i^{(d)}(\omega)}\right\}. \tag{10}$$

The received time domain signal $r(t)$ can be considered as a sum of individual ray contribution $r_i(t)$ where,

$$r_i(t) = \mathcal{F}^{-1}\left(S(\omega)\,P_i e^{-\jmath\phi_i^{(d)}(\omega)}\right). \tag{11}$$

By utilizing the identities of Equations (9) and (5) this time domain ray can be expressed as follows,

$$
\begin{aligned}
r_i(t) = {} & \alpha_{0i}s_i\left(t - \frac{d_i}{c}\right) - \beta_{0i}\mathcal{H}\left\{s_i\left(t - \frac{d_i}{c}\right)\right\} \\
& + \sum_{g=1,3,5\,\ldots}^{m-1} (\sqrt{-1})^{g-1}\left(\alpha_{gi}\frac{d_g}{dt^g}\mathcal{H}\left\{s_i\left(t - \frac{d_i}{c}\right)\right\} + \beta_{gi}\frac{d_g}{dt^g}s_i\left(t - \frac{d_i}{c}\right)\right) \\
& + \sum_{g=2,4,6\,\ldots}^{m} (\sqrt{-1})^{g}\left(\alpha_{gi}\frac{d_g}{dt^g}s_i\left(t - \frac{d_i}{c}\right) - \beta_{gi}\frac{d_g}{dt^g}\mathcal{H}\left\{s_i\left(t - \frac{d_i}{c}\right)\right\}\right),
\end{aligned}
\tag{12}
$$

where $\mathcal{H}$ is the Hilbert transform, c is the speed of light and d_i is the propagation length of the ray. α_{gi} and β_{gi} are the respective real and imaginary parts of the gth order coefficient of the polynomial P_i of Equation (11). A distinct advantage of using the form in Equation (11) is that $s(t)$, $\mathcal{H}s(t)$ and their corresponding derivatives can be pre-computed prior to any simulations. This leaves the polynomial coefficients α_{gi} and β_{gi} and the ray path delays d_i/c to be computed during runtime. Assuming a prior known set of transmitted waveforms for $s(t)$, this offers an efficient means of computing the time domain received signal $r(t)$.

3.1. Propagation through Walls

Transmissions through object with complex varying permittivities results in a ray undergoing an additional frequency dependant effect. Propagation through such objects results in an addition phase component ϕ^{trans} being added to the frequency domain response of a ray. This component is expressed as,

$$\phi^{trans} = d_l\beta_l, \tag{13}$$

where d_l is the propagation distance of the ray inside the wall and β_l is the phase constant due to the material given by,

$$\beta_l = \omega\sqrt{\mu\varepsilon'}\left\{\frac{1}{2}\left[\sqrt{1 + (p_e)} + 1\right]\right\}^{\frac{1}{2}}. \tag{14}$$

It should be noted that we assume a relative complex permittivity of the form,

$$\varepsilon_{re}(\omega) = \varepsilon_r(\omega) - j\left(p_e\varepsilon_r(\omega)\right), \tag{15}$$

where μ is the permeability, ε' is relative permittivity and p_e is the loss tangent of the wall's constituent material.

The total phase of a ray that propagates through an air medium and a wall can be express as

$$\phi_i^{tot}(\omega) = \frac{d_i\omega}{c} + d_l\beta_l \tag{16}$$

where d_i is the free space propagation length. While $\phi_i^{tot}(\omega)$ is not a linear phase response it can be approximated by an effective linear phase $\phi_i^e \approx \phi_i^{tot}$. This is possible as the portion of a typical ray that passes through object in an indoor environment will be significantly shorter than the free-space propagation length such that $d_i \ll d_l$. The effective linear phase ϕ_i^e can be used in Equation (4) instead of ϕ_i^d as a means of isolating the slowly varying residuals.

4. RESULTS

In order to evaluate the accuracy and efficiency of the proposed algorithm a ray tracing simulation was executed using a simple building environment based on the layout of the Sensor Laboratory of the Antennas and Electromagnetic Research Group at Queen Mary University London (Figure 1(b)). For the purposes of our simulations, we use the values of the relative complex permittivity for concrete as measured by Muqaibel et al. in [6] (Figure 1(a)). Rays were limited to two orders of reflection and the transmitted signal $s(t)$ consisted of a Gaussian doublet pulse waveform of Equation (17), as shown in Figures 2(a) and 2(b).

$$s(t) = 1 - \left(4\pi \left(\frac{t}{t_n}\right)^2\right) \exp\left(-2\pi \left(\frac{t}{t_n}\right)^2\right), \tag{17}$$

where $t_n = 0.780$ nanoseconds and thus determines the bandwidth of the pulse.

(a)   (b)
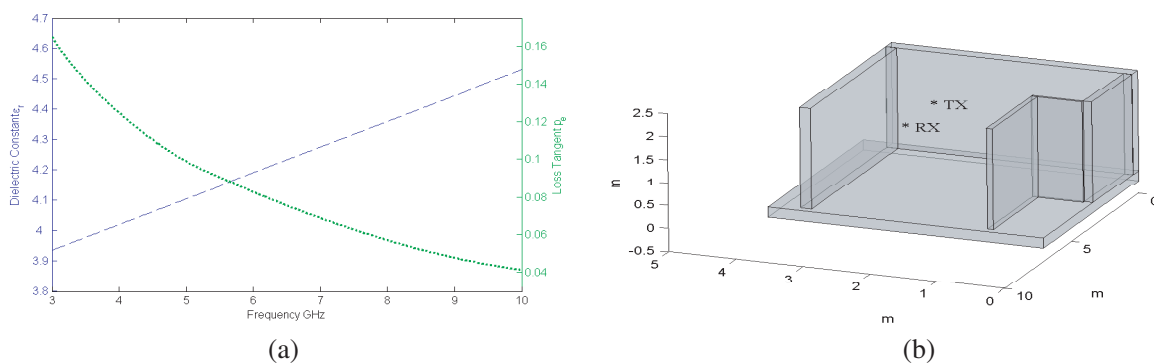
Figure 1: (a) Variation of permittivity of concrete dashed line: Dielectric constant dotted line loss tangent; (b) Building environment.

(a) Frequency response of Gaussian doublet pulse   (b) Gaussian doublet pulse in time domain
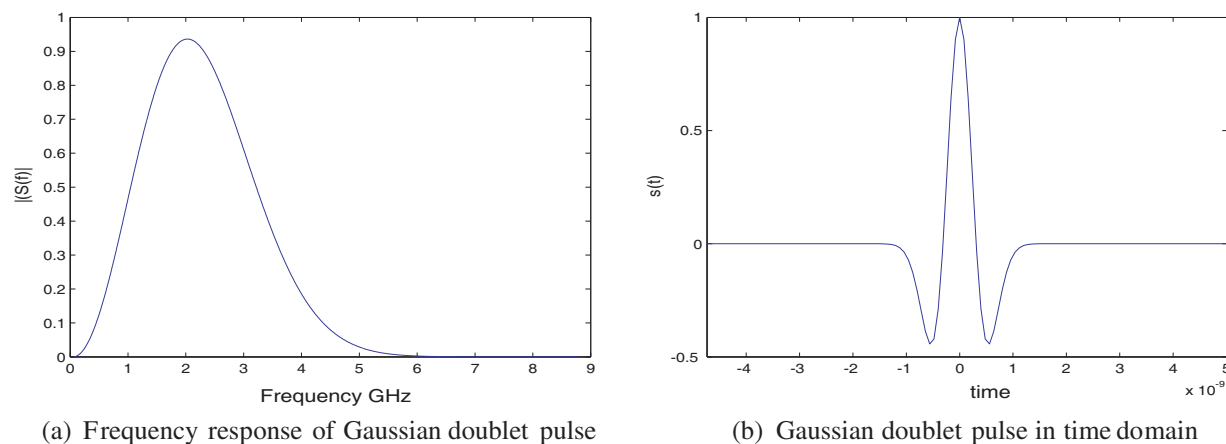
Figure 2: Gaussian doublet transmitted UWB pulse.

A ray trace simulation was carried out over a full set of 800 frequency samples from 0.5 MHz to 10 GHz. Subsequently a simulation using a reduced frequency sweep of 24 samples was executed and the proposed algorithm was applied. The reduced frequency sweep consisted of 71 samples and a 11th order polynomials was used to approximate ray responses in the frequency domain. Figure 3(a) shows a section of the time domain received signal $r(t)$ obtained using the acceleration technique along side the received signal computed using the original full frequency sweep ray-trace. From Figure 3(a) it is clear that the full frequency sweep simulation results match those of the accelerated reduced frequency sweep in the time domain. This is confirmed in Figure 3(b) where a less than 0.8% relative error is observed. By applying the acceleration algorithm a 75% time saving was achieved. In order to access the accuracy of our simulator it was necessary to compare simulation results with measured data for UWB channels. This involved measuring the channel impulse response of a UWB channel in the building environment shown in Figure 1(b). A channel bandwidth from

3 GHz to 10 GHz was used in conjunction with a pair of Planar Inverted Cone Antenna (PICA) developed at Queen Mary University London [7]. This measured impulse response (dotted line) is shown in Figure 3 along with simulated response (dash line). The transmitter-receiver antenna pair was positioned 2.1 meters apart as in Figure 1(b).
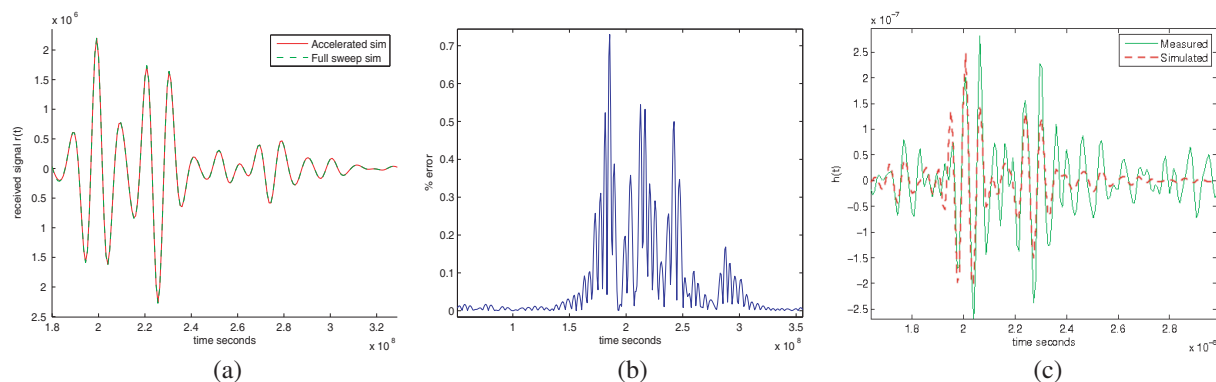

Figure 3: (a) Simulated received signal $r(t)$; (b) Relative error expressed as % of maximum received signal; (c) Measured and simulated results for channel response.

5. CONCLUSION

We have presented a frequency domain based ray tracing acceleration technique that reduced number of frequency samples at which a ray trace has to be computed when simulating UWB propagation. Deploying this accelerated simulator in a indoor building scenario yields a time saving of 75%, with a maximum relative error of less than 0.8%. Additionally, validation of simulation result with empirical measurements has been performed.

ACKNOWLEDGMENT

REFERENCES

1. Porcino, P. and W. Hirt, "Ultra-Wideband radio technology: Potential and challenges ahead," *IEEE Communications Magazine*, Vol. 41, No. 7, 66–74, 2003.

2. Tian, Z., G. B. Giannakis, H. Kobayashi, A. F. Molisch, V. H. Poor, and Z. Sahinoglu, "Localization via ultra-wideband radios: A look at positioning aspects of future sensor networks," *IEEE Signal Processing Magazine*, Vol. 22, No. 4, 70–84, July 2005.

3. Kim, S. C., B. J. Guarino, Jr., T. M. Willis III, E. Vinko, S. J. Fourtune, R. A. Valenzuela, et al., "Radio propagation measurements and prediction using three-dimensional ray tracing in urban environments at 908 MHz and 1.9 GHz," *IEEE Transactions on Vehicular Technology*, Vol. 48, No. 3, 931–945, 1999.

4. Diskin, J. and C. Brennan, "A 3D ray tracing tool for indoor UWB signal propagation modelling," *Proceedings of the IEE Irish Signals and Systems Conference 2005*, 289–93, Dublin, Ireland, September 2005.

5. Yao, R., W. Zhu, and Z. Chen, "An efficient time-domain ray model for UWB indoor multipath propagation channel," *Proceedings of the 2003 IEEE 58th Vehicular Technology Conference VTC 2003-Fall*, Pt. 2, Vol. 2, 1293–7, Orlando, United Stated, 2003.

6. Muqaibel, A., A. Safaai-Jazi, A. Bayram, A. M. Attiya, and S. M. Riad, "Ultrawide-band through-the-wall propagation," *IEE Proceeding Microwave and Antennas Propagation*, Vol. 152, No. 6, December 2005.

7. Alomainy, A., Y. Hao, C. G. Parini, and P. S. Hall, "Comparison between two different antennas for UWB on-body propagation measurements," *IEEE Antennas and Wireless Propagation Letters*, Vol. 4, 2005.

Experimental Study of Atmospheric Turbulence Effects on RoFSO Communication Systems

W. Ni[1]**, Y. Miyamoto**[1]**, K. Wakamori**[2]**, K. Kazaura**[2]**,**
M. Matsumoto[2]**, T. Higashino**[3]**, K. Tsukamoto**[3]**, and S. Komaki**[3]

[1]Hamamatsu Photonics K. K., 5000 Hirakuchi, Hamakita-ku
Hamamatsu-shi, 434-8601, Japan
[2]GITS/GITI, Waseda University, 1011 Okuboyama, Nishitomida
Honjo-shi, 367-0035, Japan
[3]Division of EEIE, Graduate School of Engineering, Osaka University
2-1 Yamada-oka, Suita-shi, 565-0871, Japan

Abstract— Many existing and emerging broadband wireless services benefit from the Radio on Fiber (RoF) technology offering high speed and large bandwidth over long distance at a low attenuation. Although RoF approach works well in many cases, in fiber not feasible area, free-space optics (FSO) communication techniques have received renewed interest as alternative for Radio-Frequency (RF) signal transmission-concept referred to as Radio on FSO (RoFSO)-because they can offer similar capacity while being cost effective and secure access without extensive installation of copper or fiber infrastructure. We are currently in the development stage of the advanced RoFSO system. On the other hand, atmospheric turbulence manifested as beam wander, intensity fluctuation and beam spreading has significant influence on the performance of conventional FSO as well as RoFSO systems. In order to measure, characterize and quantify the influence of atmospheric turbulence in our system deployment environment, we described in this paper a setup Radio Frequency-FSO (RF-FSO) link to test the basic characteristics of RF signal transmission under various atmospheric and weather conditions e.g., clear weather, clouds, rain, fog and snow. We have attempted to quantify the strength of atmospheric turbulence by calculating the refractive-index structure parameter C_n^2 from scintillation index indicating intensity fluctuation caused by atmospheric turbulence. Correlation analysis on carrier-to-noise ratio (CNR) and C_n^2 was also made and presented. In addition, a comparative measurement to characterize intensity scintillation and Angle-of-Arrival (AoA) variance has also been presented in this paper. The experimentally derived data and results will be available for antenna design, link margin prediction and evaluation of performance criterion of future RoFSO systems in different deployment environments.

1. INTRODUCTION

In the future ubiquitous society, consumers will be provided with reliable and cost effective communication systems that can support anytime, anywhere, any media they want. As a result, wideband radio links will become more prevalent in communication systems. To satisfy this increasing demand, the high capacity optical networks should be integrated with the flexibility of radio networks. The aforementioned wideband demands can be specifically met with fiber based wireless access schemes such as RoF [1] by which high-speed multimedia can be easily supported. In areas where optical fiber or other broadband networking technologies deployment are not feasible, FSO communication links can be utilized as an alternative technology because they can offer capacity similar to optical fiber. RoFSO [2] link can be used to reliably transport multiple RF signals comprising of various kinds of communication and broadcasting signals.

We are developing an advanced dense wavelength division multiplexing (DWDM) RoFSO link whereby the basic antenna design concept will be similar to the full-optical FSO antenna [3]. Full-optical FSO links [4] in Figure 1(b) utilize seamless connection of free-space beam to single-mode fiber (SMF) therefore eliminating the necessity of converting the transmitted signal from optical-to-electrical (O/E) or vice versa in conventional FSO as in. Figure 1(a). The use of 1550 nm as the transmission wavelength for FSO systems is attractive for DWDM RoFSO system as it will be compatible with long-haul fiber optic technologies such as wavelength division multiplexing and erbium doped fiber amplifiers [5, 6]. Atmospheric turbulence, i.e., refraction-index fluctuations in free space, will destroy light beam spatial coherence, cause phase front distortion, alter and redirect the flow of energy in the beam, and finally lead irradiance and AoA fluctuations in the plane of the receiver aperture. This can be characterized statistically by the variance of optical intensity as depicted in Figure 1(c). Moreover, in full-optical FSO links, the free-space transmitted optical

beam will be received and coupled directly onto the SMF interface with core diameter of only $10\,\mu\mathrm{m}$, that is more difficult to ensure the stability of the RoFSO communication [7].
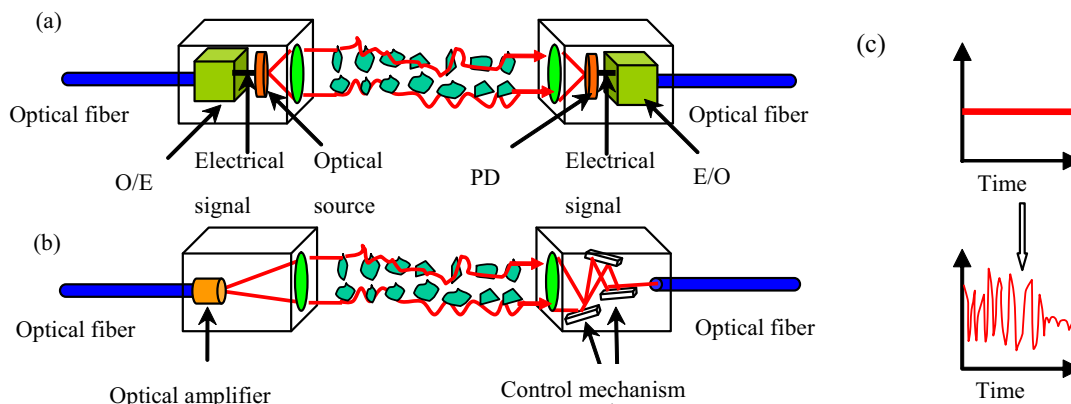

Figure 1: (a) Conventional FSO link; (b) full-optical FSO link; (c) intensity variance attributed to atmospheric turbulence.

2. MEASUREMENT AND ANALYSIS OF ATMOSPHERIC TURBULENCE

2.1. Measurement Setup

In order to study atmospheric turbulence impact on the quality of the envisaged RoFSO, RF-FSO antennas (commercial available Canobeam DT-170 antennas) have been used to set up an experimental link over 1km distance between two campuses of Waseda University in Tokyo. As depicted in Figure 2, an RF signal generator (Agilent E4438C) is used to generate a W-CDMA signal which is propagated through the atmosphere using the RF-FSO antenna installed on the rooftop of Bldg. 14 at the Nishi-waseda campus. On the opposite side, the corresponding RF-FSO installed on the rooftop of Bldg. 55 at the Okubo campus along with a signal analyzer is used to measure the received optical power and calculate the carrier-to-noise ratio (CNR) in 1 minute interval. An optical FSO antenna has been set to measure optical attenuation and intensity fluctuation simultaneously. And a weather meter was used to record temperature, rain rate, and other weather condition data. Signals from both of them are sampled and processed by the monitor personal-computer (PC).
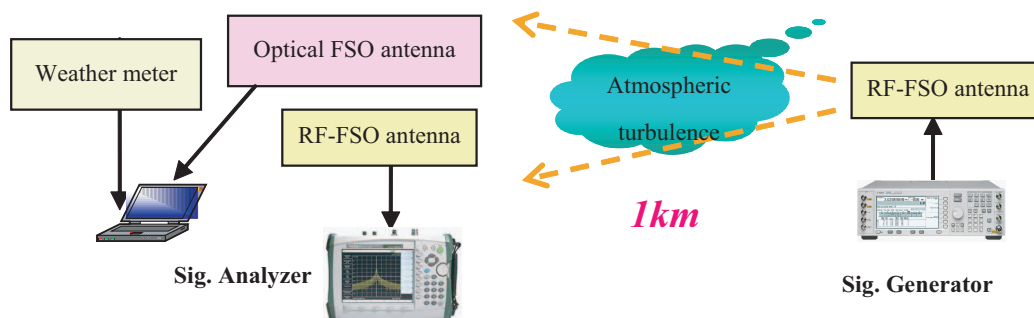

Figure 2: Schematic of RF-FSO test link.

2.2. The Refractive-index Structure Parameter C_n^2

The refractive-index structure parameter C_n^2 $(\mathrm{m}^{-2/3})$ is the parameter most commonly used to describe the strength of atmospheric turbulence. According to Andrews et al. [8], C_n^2 can be approximately calculated as following:

$$C_n^2 \approx \frac{\sigma_1^2}{1.23 k^{7/6} L^{11/6}} \tag{1}$$

where $k = 2\pi/\lambda$ is wavenumber, L is link distance and σ_1^2 is called the Rytov variance [8] which can be obtained by

$$\sigma_1^2 \approx \sigma_I^2 = \frac{< I^2 > - < I >^2}{< I >^2} \qquad (2)$$

where the quantity I denotes the measured irradiance of the optical wave and the angle brackets $<>$ denote an ensemble average or, equivalently, a long-time average. In the RF-FSO test link, optical power received by the optical FSO antenna was sampled as in Figure 3(a). It can be seen that intensity denoted by the monitor voltage V caused by atmospheric turbulence fluctuates more strongly on clear day than rainy day.
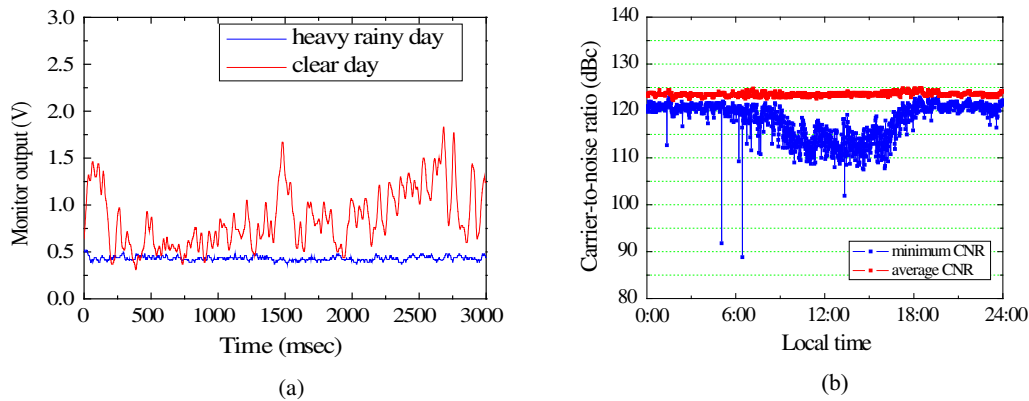
Figure 3: (a) Measured intensity fluctuations; (b) minimum and average CNR (August 6, 2008).

Calculated C_n^2 values nearly show the same diurnal cycle on all clear days due to similar conditions, cloudless sky when daily temperatures have also been relatively higher as shown in Figure 4(a). Around noon, maximum C_n^2 value of about $2 \cdot 10^{-13}$ m$^{-2/3}$ can be found as the refraction-index fluctuations are caused almost exclusively by temperature variations. When the air temperature is close to ground temperature, for example at sunrise and during the night, C_n^2 values are lower and even reach the minimum (about $3 \cdot 10^{-15}$ m$^{-2/3}$). For the month of July in 2008, it rained frequently even though the temperature was still higher. As presented in Figure 4(b), the calculated C_n^2 values were almost the same during a whole day without any characteristic diurnal pattern identified The phenomenon was the similar in other three seasons: on rainy days, C_n^2 values are lower than the ones on clear days in the same season. The main reason for that is on rainy day, temperature stratification of the atmosphere was stable. Figure 4(c) gives the C_n^2 values on clear days in December 2007 and July 2008. The typical diurnal cycle of C_n^2 was both found except that the C_n^2 values were generally higher in summer than those in winter. That is because solar radiation is usually weaker in winter season, thus the temperature is relatively lower as well as temperature stratification. The phenomenon can be more clearly seen on rainy days in Figure 4(d). Therefore, when we make consideration for effect of atmospheric turbulence on the future RoFSO communication link, we should pay more attention in summer season with clear weather condition, e.g., system link margin should be designed to be able to tolerate the strongest strength of atmospheric turbulence at that time with C_n^2 value of $2 \cdot 10^{-13}$ m$^{-2/3}$ obtained from our experimental setup in the same deployment environment.

2.3. Correlation Analysis on CNR

One-day result of the recorded CNR values from the RF-FSO test setup has been displayed in above Figure 3(b). Both minimum and average CNR were generally above the desired minimum threshold value of 110 dBc for our present W-CDMA signal transmission. The maximum variance of the minimum CNR was found at noon period when the strength of atmospheric turbulence was also larger as stated in Section 2.2. We have made correlation analysis on C_n^2 values and CNR deviation which is defined as the difference between average and minimum CNR at the same time point. On clear day in summer season, CNR deviations increased with the increasing C_n^2 values as shown in Figure 5(a). On rainy day, CNR deviations caused by the atmospheric turbulence were thought to be mostly not larger than 6 dB as seen in Figure 5(b) while other larger deviations maybe caused by heavy rain drops. This phenomenon can also be found in other three seasons. So, from
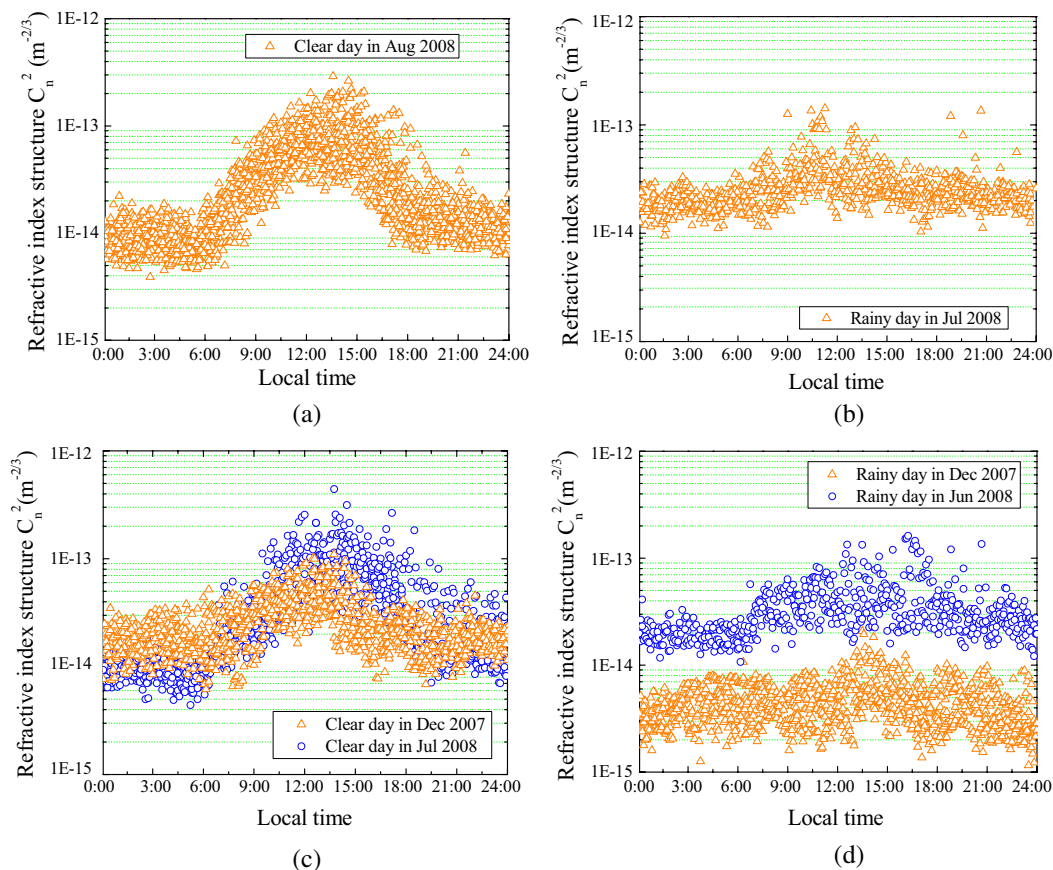
Figure 4: Calculation results of C_n^2 values on: (a) clear day in summer; (b) rainy day in summer; (c) clear day in winter and summer; (d) rainy day in winter and summer.

the relationship between CNR deviation and C_n^2 values at corresponding time point, we can get an estimation of redundancy for CNR deviation to mitigate the effect of atmospheric turbulence, e.g., threshold of CNR deviation should be not less than 10 dB in the setup RF-FSO link system to withstand strongest atmospheric turbulence.
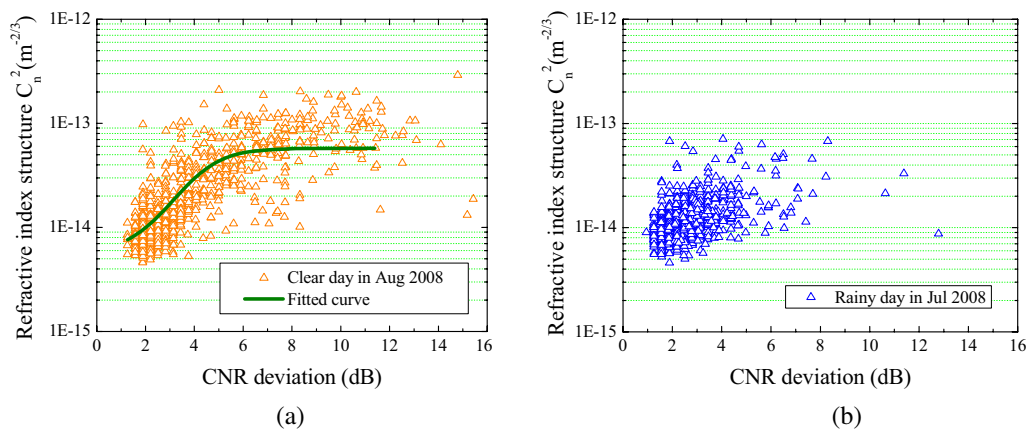

Figure 5: C_n^2 values vs. CNR deviation: (a) on clear day; (b) on rainy day in summer of 2008.

2.4. AoA Fluctuations

We have also set up a FSO link, depicted in Feigure 6, for studying turbulence-induced phase fluctuations on the propagating optical wave. The link distance is 300 m between two buildings in Hamamatsu Photonics K. K. On the receiver side, optical beam propagating through atmosphere

will be focused onto a position sensitive detector (PSD) inside one FSO antenna with aperture diameter of 0.04 m. A/D conversion of signals from the two receiving antennas will be made by Peripheral Interface Controller (PIC). The interested data including received optical power, scintillation index, average arrival angles and variance of AoA were sampled at 10 kHz and recoded in 15 seconds interval in PC, as shown in Figure 6. The AoA variance is a result of phase change and can relate to C_n^2 according to [8, 9]:

$$< \beta^2 > = 2.91 C_n^2 L D_R^{-1/3} \tag{3}$$

where L is link distance, here is 300 m and D is the receiver aperture of 0.04 m. The C_n^2 values can be calculated from the received intensity data of the FSO antenna with PD by applying (1) and (2). Comparison was made between the simulated AoA variance from (3) and experimentally recorded one from PSD. On clear day, AoA variances in both X and Y directions agreed well with the calculated ones. And all of those variances fluctuated with the C_n^2 values. On rainy day as in Figure 7(b), the experimental AoA variances also agreed with the calculated ones and changed with C_n^2 except for some large deviation points caused by large rain drops. And AoA variance changes more severely with clear sunny weather than cloudy or rainy ones as shown in Figure 7. The largest deviation happens mostly at noon time on clear day. Therefore, as expected, AoA variance fluctuates well with the strength of atmospheric turbulence. Moreover, it is helpful that we can make a better estimation of AoA variance as long as we know the strength of atmospheric turbulence, i.e., C_n^2 values in the deployment environment in preparation for RoFSO system design, e.g., setting suitable tracking characteristics.
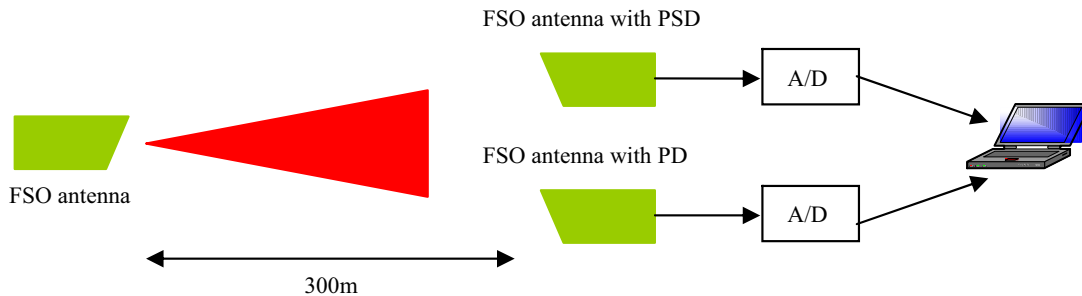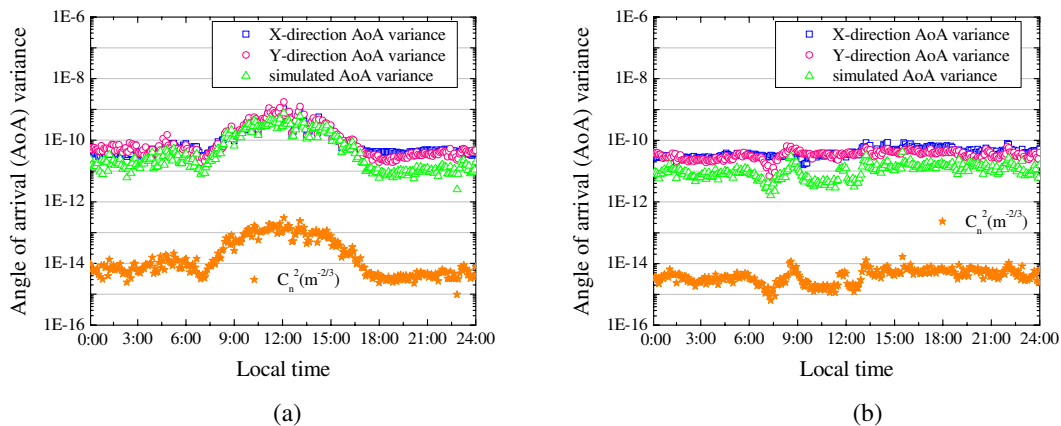

Figure 6: Angle of arrival measurement setup.

Figure 7: Simulated and experimental AoA variances: (a) on clear day (Mar 18th, 2008); (b) on rainy day (Mar 19th, 2008).

3. CONCLUSIONS

We have presented some preliminary experimental results with more emphasis on the effect of atmospheric turbulence from an existing RF-FSO test bed which is significant in the design and

evaluation of the advanced RoFSO system. The results will be available for antenna design, link margin prediction and evaluation of performance criterion of RoFSO systems in different deployment environments.

ACKNOWLEDGMENT

REFERENCES

1. Al-Raweshidy, H. and S. Komaki, eds., *Radio over Fiber Technologies for Mobile Communication Networks*, First ed., Artech House Publishers, 2002.
2. Tsukamoto, K., K. Nakaduka, et al., "Development of DWDM radio on free space optic link system for ubiquitous wireless," *Asia-Pacific Microwave Photonics Conference (AP-MWP 2007)*, Apr. 2007.
3. Arimoto, Y., "Multi-gigabit free-space optical communication system with bidirectional beacon tracking," *IEEJ Trans. FM*, Vol. 127, No. 7, 385–390, 2007.
4. Marioni, F., Z. Sodnik, and F. E. Zocchi, "2.5-Gb/s free-space optics link over 1.1 km with direct fiber coupling to commercial devices," *Free-space Laser Communications IV*, Vol. 5550, 60–69, Oct. 2004.
5. Nykolak, G., P. Szajowski, D. Romain, G. Tourgee, H. Presby, and J. Auborn, "Update on 4 × 2.5 Gb/s, 4.4 km free-space optical communications link: Availability and scintillation performance," *Optical Wireless Communications II*, Vol. 3850, 11–19, SPIE, Boston, Massachusetts, Sept. 1999.
6. Kazaura, K., K. Omae, et al., "Performance evaluation of next generation free-space optical communication system," *IEICE Transactions on Electronics Special Section on Evolution of Microwave and Millimeter-wave Photonics Technology*, Vol. E90-C, No. 2, 381–388, Feb. 2007.
7. Aburakawa, Y. and T. Otsu, "Dense wavelength division multiplexed optical wireless link towards terabit transmission," *Proc. MWP'03*, 135–138, 2003.
8. Andrews, L. C. and R. L. Phillips, *Laser Beam Propagation through Random Media*, SPIE Press, Washington, 1998.
9. Andrews, L. C., R. L. Phillips, and C. Y. Hopen, *Laser Beam Scintillation with Applications*, SPIE Press, Bellingham, Washington, 2001.

Soft-lithography-based Inter-chip Optical Interconnects

Wei Ni[1], **Rubing Shao**[1], **Jing Wu**[2], **and X. Wu**[1]

[1]State Key Laboratory of Modern Optical Instrumentation, Department of Optical Engineering
Zhejiang University, Hangzhou 310027, China
[2]University of California at San Diego, La Jolla, CA 92093, USA

Abstract— The increasing performance of microprocessors leads to higher bandwidth requirements for the data flow to and from the processor. Today, all signaling on a PCB is performed electrically, using copper lines that are integrated in the board. However, issues such as propagation loss and inter-channel crosstalk limit the scalability of electrical interconnects to ever higher bandwidth densities. Optical interconnects feature a higher bandwidth × length product, are more power-efficient and enable a higher bandwidth density than electrical interconnects do. This paper describes a kind of two-dimensional monolayer optical interconnects providing interconnections between chips on conventional PCB. We have designed a soft-lithography-based, versatile coupling structure with a 45° total internal reflector (TIR), a beam duct, and a polymer waveguide in order to vertically couple light beams between transmitter (or receiver) and the waveguide layer. This proposed integrated architecture of a polymeric optical interconnection has been demonstrated to be advantageous in the aspects of misalignment tolerance, ease and low cost of fabrication, as well as relative simplicity in deployment. We also investigated the characteristics of in-plane connections including cross-over and branching nodes in the optical interconnects with experimental and theoretical analysis. The theoretical crosstalk, as calculated by a function of crossing angle, was determined for a set of interconnect pairs with varying cross-sections, and was compared with experimental measurements. Furthermore, a suitable branching angle was found for branching node and the effects of short-distance mode scrambling in highly multimode polymer waveguides were studied in detail in this paper too.

1. INTRODUCTION

Photonic technologies have been widely accepted as a way to alleviate bottlenecks in platform-to-platform, machine-to-machine and board-to-board interconnections [1]. Recent breakthroughs in the fabrication of spatial arrays of optoelectronic emitters and detectors and their heterogeneous integration with Si-CMOS electronic chips now encourage the use of optics as an electronic wire replacement technology also at the chip-to-chip and on-chip interconnection level as in Figure 1. The main objective for introducing two dimensional photonic pin-outs at this level of the interconnection hierarchy aims at relaxing the bandwidth limitations between these electronic processing modules primarily imposed by fundamental electrical signal propagation issues and the limited number of electrical chip pin-outs. With the debut of 25 Gb/s board-level interconnects, optical interconnects have demonstrated their ability to provide communications infrastructure for next-generation computing [2, 3]. In the domain of very-high-bandwidth short-range communications, light-based waveguides have consistently demonstrated higher placement density, more packaging flexibility, and superior alignment reliability than their electrical counterparts [4, 5].

A kind of two-dimensional monolayer optical interconnects providing interconnections between chips on conventional PCB are proposed in this paper. Detailed theoretical analysis along with the experimental measurements of the interconnection circuit performance is also presented.

2. VERTICAL COUPLING

Most prior designs of optical interconnections usually call for a high-precision Vertical Cavity Surface Emitting Laser (VCSEL)/Photodiode (PD) alignment, often with an alignment-error requirement of less than a few micrometers; this exhibits an immediate difficulty in the assembly of electro-optical PCBs. In order to increase the tolerance of the interconnection waveguide to alignment errors, we have designed coupling structure with a 45° TIR, a beam duct, and a polymer waveguide as in Figure 2(a) in order to provide high-speed optical communications within a board; the driving electrical pulses modulate the VCSEL, and the light received at the photodiode through the waveguide demodulates back as electric signals on the surface of the PCB. We utilized Zemax® to simulate the efficiency of the beam duct by varying the length of duct section while output light in the waveguide was monitored, and portion of simulation is displayed in Figure 2(b). The coupling efficiency peaked at around 6 mm for TIR reflector's triangular prism structure at the proposed 0.5×0.5 mm with waveguide cross section of 0.3×0.3 mm.
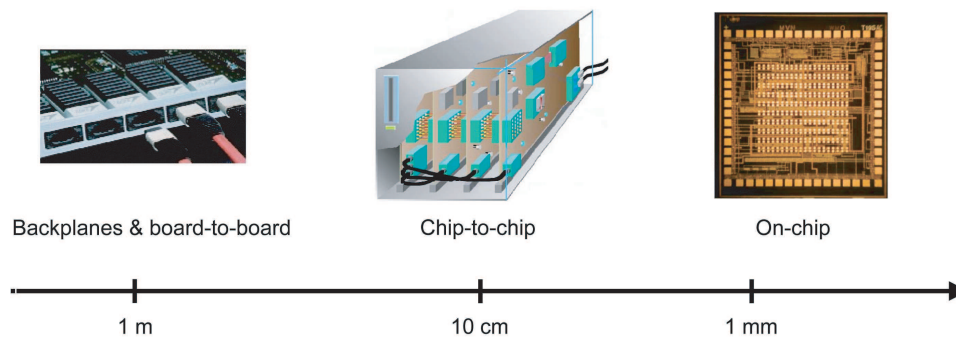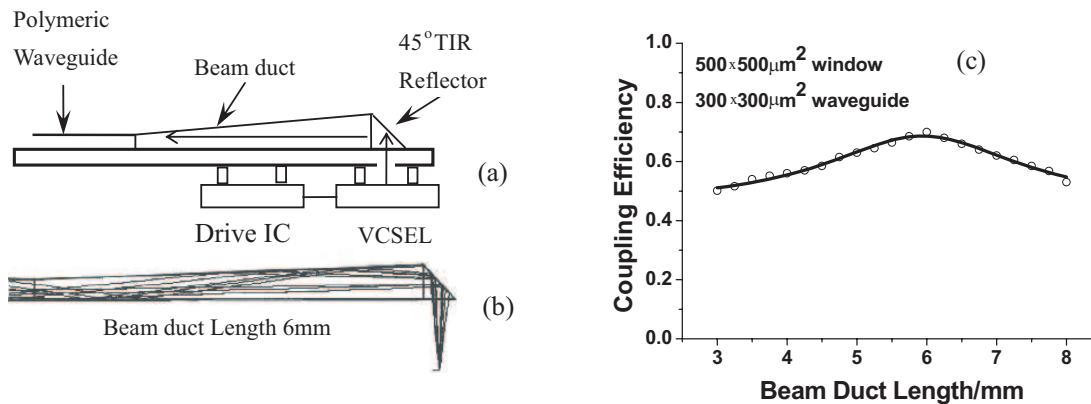
Figure 1: Interconnect distance.

Figure 2: (a) Schematic of the coupling portion of the photonic circuits on PCB, (b) ray tracing result with optimal beam duct length, (c) coupling efficiency as a function of beam duct length.

3. CROSSING AND BRANCHING OF LIGHT GUIDES IN IN-PLANE INTERCONNECTS

However, as optical interconnects will inevitably cross in-plane when used heavily within circuits and PCBs, cross-over or branching nodes are necessary for signal crossing, splitting, or isolation, and the performance of these nodes in the circuit becomes critical in determining the overall quality of optical signal transmission.

3.1. Crossing Node

For the cross-over node, as shown in Figure 3(a), crosstalk is usually required to be as low as 20–30 dB for a reliable data communications. Considering the multimode nature of the waveguides for on-board optical interconnections, we utilized both wide-angle BPM and ZEMAX® to simulate the efficiency of polymer rectangular waveguide by varying cross angle, while the output from the waveguide was monitored.

The results are shown in Figure 4(a) and (b) for both strong-confinement and weak-confinement core/clad assemblies, respectively. As shown in Figure 4, crosstalk as a function of cross angle was calculated and compared to experimental measurement for four selected cross sections of $50 \times 50\,\mu m^2$, $100 \times 100\,\mu m^2$, $200 \times 200\,\mu m^2$, and $300 \times 300\,\mu m^2$, with a cross angle ranging from $10°$ to $55°$ for strong-confinement assembly (core 1.50, cladding 1.00) and $5° \sim 12.5°$ for weak-confinement (core 1.50, cladding 1.48). For weak confinement, it is found that the crosstalk decreases linearly with crossing angle from $5°$ to $9°$, and then exhibits a faster-than-exponential attenuation when cross angle increases from $9°$ to $12°$; when above $12°$, a crosstalk of less than $-30\,dB$ is obtained. For strong confinement, however, a greater crossing angle of about $52°$ is needed to achieve a $-30\,dB$ crosstalk. It is also noted in Figure 4 that the $-20\,dB$ cross angle, as denoted with dotted blue lines, increases slightly with waveguide cross-section for both strong and weak confinement conditions. It is understood that as cross-section increases, the size of the crossing joint also increases, and so does the window of leakage to adjoining waveguide, therefore it takes a greater cross angle to achieve the same crosstalk.
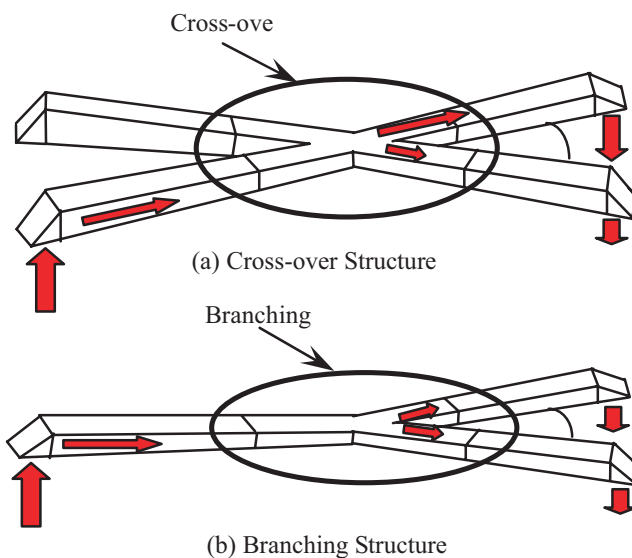
Figure 3: (a) Schematic of a cross-over interconnect pair, (b) schematic of a branching interconnect pair;
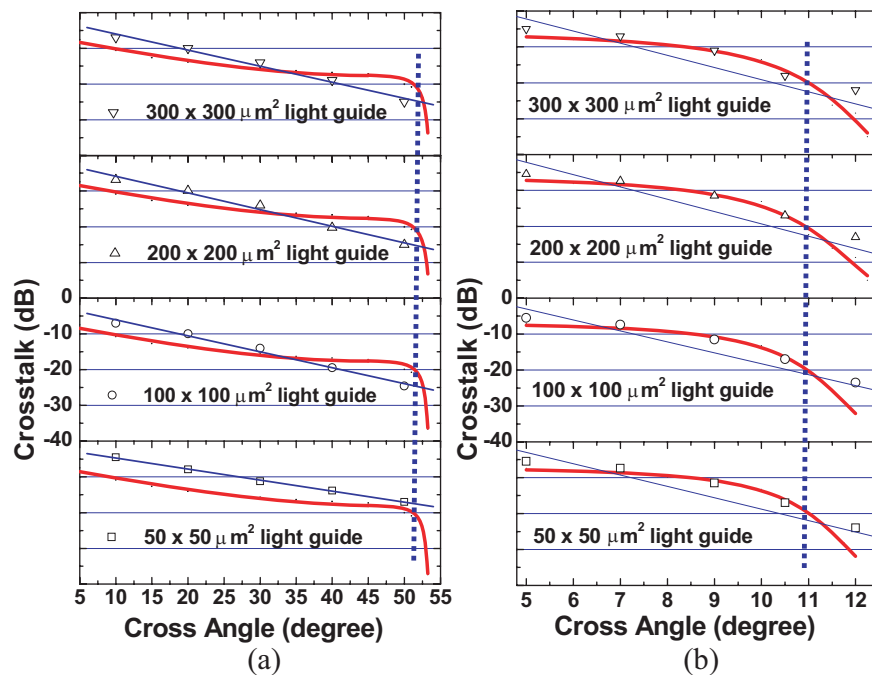

Figure 4: Dotted blue line shows that crossing angle changes with cross-section at −20 dB crosstalk. Open symbols and blue fitted curves stand for experimental measurements. Refractive indices of core/cladding were at (a) 1.50/1.00 and (b) 1.50/1.48.

3.2. Branching Node

Highly multimode Y-branching deserves special attention for its possible applications in polymer interconnects, particularly due to the special phenomenon of beam center shift in post-branching waveguides, which significantly affects their crossing characteristics. We took the case of weak-confinement assemblies and calculated power leakage in BPM simulation for both $50 \times 50\,\mu\mathrm{m}^2$ and $100 \times 100\,\mu\mathrm{m}^2$ cross-sections as a function of branching angles at the Y-junction, as shown in Figure 5(left), for a branching angle varying from $1°$ to $18°$. By using BPM simulation we found that until a branching angle of $7°$ is reached, the leakage for $50 \times 50\,\mu\mathrm{m}^2$ light guides is negligible. In the range of $7° \sim 12°$, the leakage suffers a linear increase from 0 to 100%, and remains at a constant for angles beyond $12°$. The phenomenon is partly explained by Figure 5(right), which shows part of leaking waves in Y-branching at $3°$, $8°$, and $15°$.
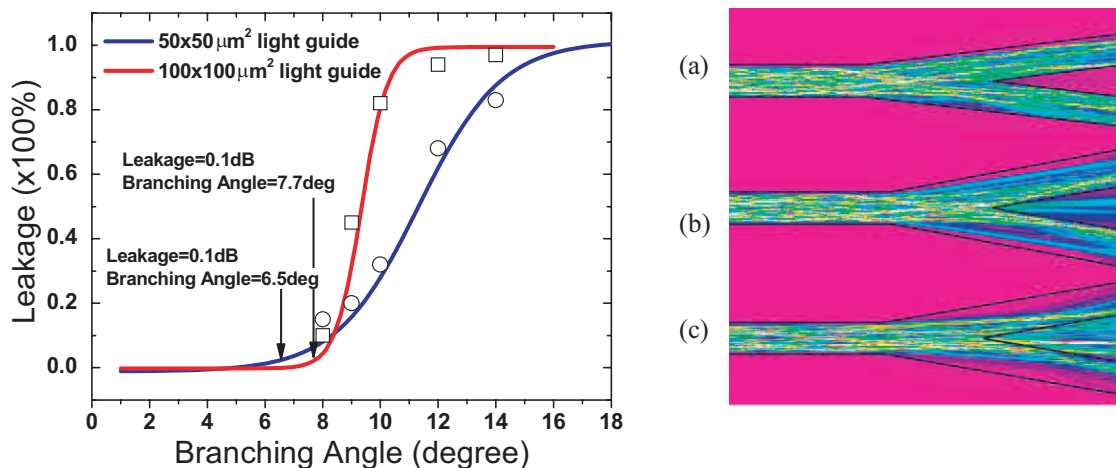
Figure 5: (left) Leakage v.s. branching angle of polymer waveguides with cross-sections of $50 \times 50 \, \mu m^2$ and $100 \times 100 \, \mu m^2$. Open symbols stand for experimental measurement. (right) BPM simulation of Y-branching with a cross-section of $50 \times 50 \, \mu m^2$ at a branching angle of (a) $3°$, (b) $8°$, and (c) $11°$.

3.3. Mode Scrambling Dependence

Moreover, multimode waveguides exhibit many special transmission properties over short-range (mm to cm) applications [6, 7], such as mode scrambling and shifting of the center of beam intensity. Thus the single-mode beam from VCSEL should be transformed into multimode within a short distance before its further propagation in the optical layer. In our practical layout designs, it was found that a non-uniform mode distribution would influence follow-up circuit output by a great deal and makes it nearly impossible to predict output power from final terminals. A mode scrambler is thus a necessity in order to make beam output as planned. In case of board-level interconnection, however, instead of a particularly designed structure or a coil of fibers, a short segment of polymer rectangular waveguide with a suitable length can serve as the scrambler.

To obtain a good estimation of the shortest propagation length needed to scramble a VCSEL beam, we conducted an analysis of beam mode scrambling in terms of intensity profile as well as distribution of ray directionality. We define VCSEL beam filling factor as the fraction of FWHM (full width at half maximum) of intensity profile over waveguide width. For two typical cross-sections of waveguides used in interconnects, $50 \times 50 \, \mu m^2$ and $100 \times 100 \, \mu m^2$, the beam filling factor was calculated as a function of propagation distance using Monte Carlo simulation and result is shown in Figure 6(left), it follows that minimum lengths of 1.3 mm and 2.5 mm are needed, respectively, for the beam to uniformly fill up the 50×50 or $100 \times 100 \, \mu m^2$ waveguides in terms of intensity regardless of incident direction. Four profiles of VCSEL beam filling factor of 10%,
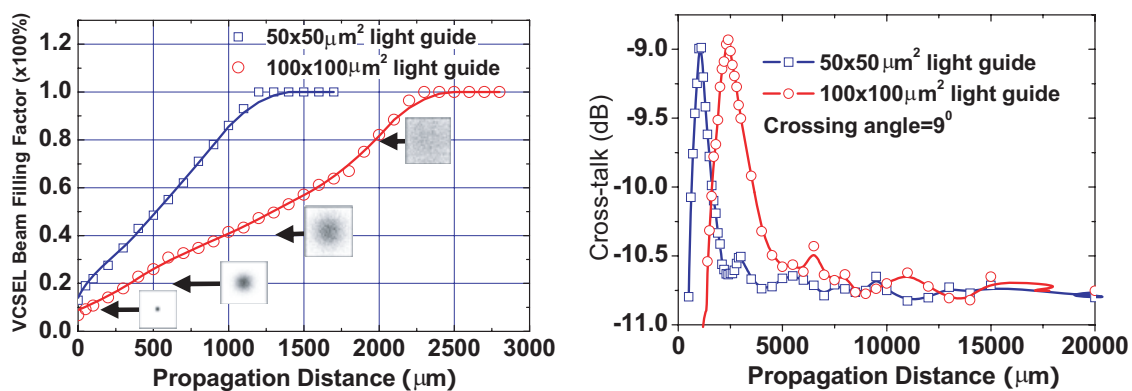

Figure 6: (Left) Calculated beam filling of light guide as a function of propagation distance when a typical SM VCSEL ($10°1/e^2$) is coupled to a straight rectangular light guide. Squares denote a waveguide of a cross-section of $50 \times 50 \, \mu m^2$, circles for $100 \times 100 \, \mu m^2$, with inset figures showing intensity profiles; (Right) Crosstalk as a function of propagation distance from the VCSEL to crossing point at a cross angle of $9°$.

25%, 40%, and 80% are shown in the insets of Figure 6(left). However, it actually takes a much longer propagation distance for ray direction to become completely scrambled. To demonstrate the effect, we calculated crosstalk at a 9° cross angle as a function of straight distance between a single mode VCSEL and crossing point and the result is displayed in Figure 6(right), in which a peak in crosstalk from below $-11\,\mathrm{dB}$ to $-9\,\mathrm{dB}$ was shown at 1.05 mm and 2.38 mm for 50×50 and $100 \times 100\,\mu\mathrm{m}^2$ light guide respectively, followed with three satellite peaks spanning to 15 mm of propagation distance.

4. CONCLUSIONS

Our analysis from the fabricated prototypes, demonstrate that this proposed integrated architecture of a polymeric optical interconnection for conventional PCB implementation is advantageous in the aspects of misalignment tolerance. The 2×2 cross-over circuit, as detailed in this paper, can achieve acceptable crosstalk at a cross angle of greater than 12° in weak confinement. We have also obtained suitable branching angle of a 1×2 branching node for the purpose of reducing leakage in the Y-junction. We have finally found proper lengths of propagation in a branching arm for the consideration of distribution uniformity of light rays. The results have a high degree of applicability to future optics-integrated PCBs featuring soft-lithography-fabricated interconnect structures.

ACKNOWLEDGMENT

REFERENCES

1. Savage, N., "Linking with light," *IEEE Spectrum*, Vol. 39, 32–36, 2002.
2. Berger, C., M. A. Kossel, C. Menolfi, T. Morf, T. Toifl, and M. L. Schmatz, "High-density optical interconnects within large-scale systems," *VCSELs and Optical Interconnects*, H. Thienpont, J. Danckaert, eds, *Proc. SPIE*, Vol. 4942, 222–235, 2003.
3. Cho, H., P. Kapur, and K. C. Saraswat, "Power comparison between high-speed electrical and optical interconnects for interchip communication," *J. Lightwave Technol.*, Vol. 22, 2021–2033, 2004.
4. Choi, C., L. Lin, Y. Liu, and R. T. Chen, "Polymer-waveguide-based fully embedded board-level optoelectronic interconnects," *Photonic Devices and Algorithms for Computing IV*, K. M. Iftekharuddin, A. A. S. Awwal, eds, *Proc. SPIE*, Vol. 4788, 68–72, 2002.
5. Steenberge, G. V., P. Geerinck, S. V. Put, J. V. Koetsem, H. Ottevaere, D. Morlion, H. Thienpont, and P. V. Daele, "MT-Compatible laser-ablated interconnections for optical printed circuit boards," *J. Lightwave Technol.*, Vol. 22, 2083–2090, 2004.
6. Kokubun, Y., T. Fuse, and K. Iga, "Optimum length of multimode optical branching waveguide for reducing its mode dependence," *Appl. Opt.*, Vol. 24, 4408–4413, 1985.
7. Eskiyerly, M., A. Garcia-Valenzuela, and M. Tabib-Azar, "Mode conversion and large angle transmission in symmetric multimode Y-junction couplers," *Integrated Optics and Microstructures*, T. A. Massood, D. L. Polla, eds, *Proc. SPIE*, Vol. 1793, 70–82, 1993.

FDTD Simulation for Statistical Properties of Microwave Scattering and Attenuation Due to Randomly Distributed Rainfalls

Yasumitsu Miyazaki[1], **Koichi Takahashi**[1], **and Nobuo Goto**[2]

[1]Department of Media Informatics, Aichi University of Technology
50-2 Manori, Nishihasama-cho, Gamagori 443-0047, Japan
[2]Institute of Technology and Science, The University of Tokushima
2-1 Minamijosanjima-cho, Tokushima 770-8506, Japan

Abstract— Rain measurement system using propagation characteristics of microwave and millimeter wave is very important for disaster prevention system to foresee and prevent the occurrence of disasters caused by strong rainfalls and for the sensor of ITS to support safe and automatic driving and cruising. In this study, propagation region is modeled as random media with randomly distributed rainfalls. Propagation of electromagnetic wave in random media is analyzed by three-dimensional FDTD method. FDTD analysis demonstrates the dynamic characteristics of wave scattering and absorption phenomena in rain region where raindrops are distributed randomly. Scattered field by single raindrop is compared with Rayleigh scattering theory to test the accuracy of numerical simulation. The effects of multiple scattering and absorption due to randomly distributed raindrops are evaluated numerically using small area analysis.

1. INTRODUCTION

Rain measurement system using propagation characteristics of microwave and millimeter wave is very effective for disaster prevention system to disasters caused by strong rainfalls and also it is very useful for the sensor of ITS (Intelligent Transport Systems) to support safe and automatic driving and cruising. Measurement technique of electromagnetic scattering and attenuation characteristics by rain is one of useful evaluation methods of rainfall rate. Rain measurement system using microwave and millimeter wave has the advantage of rapid measurement of rainfall rate in wide area compared with the direct measurement by rain gauges. To support safe driving and cruising in automatic driving system, sensor of meteorological condition on the road, particularly the amount of rain, fog and snow are indispensable [1]. The sensor using propagation characteristics of microwave and millimeter wave are very useful to detect these road conditions. Also, the efficient design for radio communication channel and radar sensing system requires accurate propagation characteristics of microwaves, it is very important to investigate the characteristics of electromagnetic wave propagation and multiple scattering phenomena in random media with many particles such as raindrops [2]. The application systems of rain measurement using microwave propagation characteristics are shown in Fig. 1.

In this study, propagation region is modeled as random media with randomly distributed raindrops. Propagation and scattering of electromagnetic wave in random media is analyzed by three-dimensional FDTD method. FDTD method is very effective for analysis of scattering and absorption phenomena which depend on polarization of incident wave when the shape of raindrop is non-spherical.

2. RAIN MEASUREMENT SYSTEM USING MICROWAVE

Rain measurement system using microwave is shown in Fig. 2. A parabolic antenna with diameter $2a = 1\,\text{m}$ is used to transmit and receive microwave of 10–20 GHz. Gaussian beam with beam spot $r_0 = 0.15\,\text{m} \cong 5\lambda \sim 10\lambda$ is transmitted. Reflected wave from a conducting plate is received by the parabolic antenna and specific attenuation $A\,(\text{dB/Km})$ is evaluated by

$$A = \alpha \cdot 10^3/(2\ell), \quad \alpha = -10\log_{10}\frac{P(2\ell)}{P_0(2\ell)} \tag{1}$$

where, for example, $\ell = 20\,(\text{m})$ is the distance from the antenna and the conducting plate. Here, P_0 and P are the received powers in free space and rain region. 2ℓ is propagation distance. The conducting plate needs sufficiently area to reflect transmitted wave with spreading beam width. Radiated field $E_y(x, y, z)$ is microwave beam given by Eq. (2) when the distribution of electric
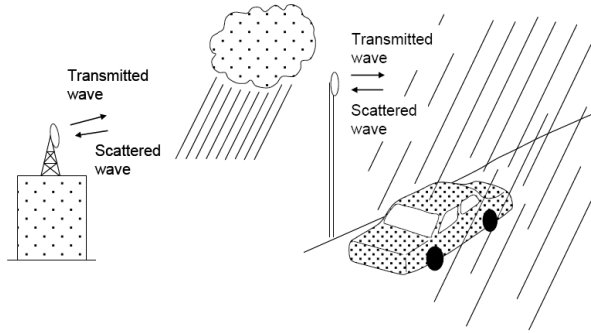
Figure 1: Rain measurement in weather measurement system and ITS.
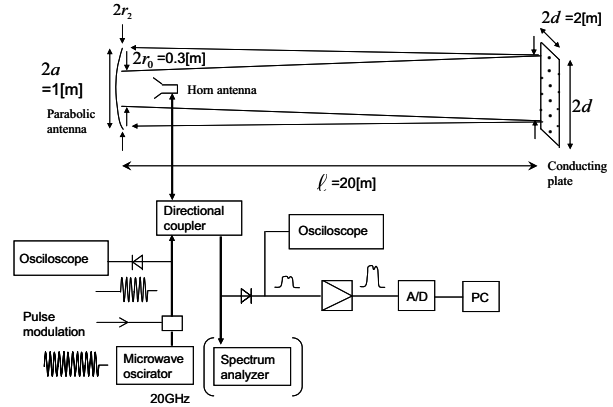

Figure 2: Measurement system of rain attenuation using microwave.

field on the incident plane S_0 is Gaussian with beam spot size r_{0x} and r_{0y} in x and y coordinate, respectively.

$$
E_y(x, y, z) = \frac{jk}{4} \frac{e^{-jk\sqrt{x^2+y^2+z^2}}}{\sqrt{x^2 + y^2 + z^2}} E_0 \left(1 + \frac{z}{\sqrt{x^2 + y^2 + z^2}}\right) \cdot \frac{r_{0x}}{\sqrt{1 + j\frac{kr_{0x}^2}{2z}}} \frac{r_{0y}}{\sqrt{1 + j\frac{kr_{0y}^2}{2z}}}
$$
$$
\cdot e^{-\frac{(kr_{0x})^2}{(2z)^2+\left(kr_{0x}^2\right)}x^2+j\frac{(kr_{0x})^3 r_{0x}}{(2z)^3+2z\left(kr_{0x}^2\right)^2}x^2} \cdot e^{-\frac{(kr_{0y})^2}{(2z)^2+\left(kr_{0y}^2\right)}y^2+j\frac{(kr_{0y})^3 r_{0y}}{(2z)^3+2z\left(kr_{0y}^2\right)^2}y^2} \tag{2}
$$

Amplitudes of incident field at propagation distances $z = 5, 10, 20$ and $40\,(\mathrm{m})$ are shown in Fig. 4. The beam width at $z = 20\,(\mathrm{m})$ can be covered by conducting plate with a side length $2d = 2\,(\mathrm{m})$.
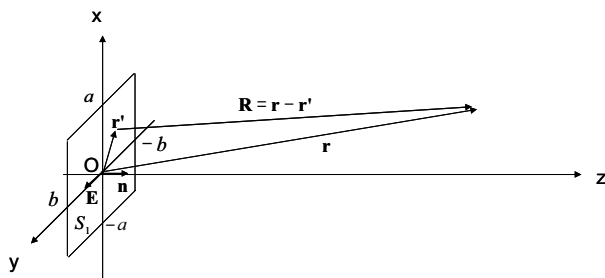
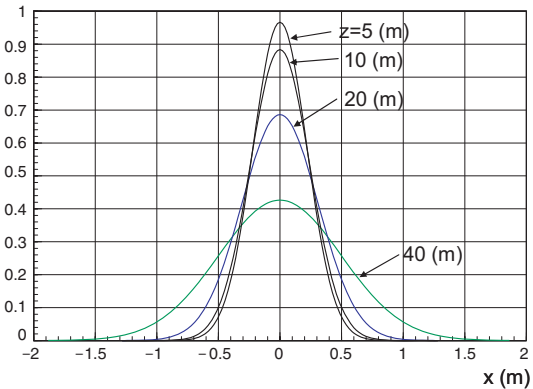Figure 3: Radiation field by rectangular aperture $\max|E_{inc}|\,(\mathrm{V/m})$.

Figure 4: Incident electric field ($f = 20\,\mathrm{GHz}$, $\lambda = 0.015\,\mathrm{m}$, $r_{0x} = r_{0y} = r_0$).

3. FDTD ANALYSIS OF ELECTROMAGNETIC SCATTERING AND ATTENUATION BY RAIN

Electromagnetic scattering and attenuation characteristics due to randomly distributed raindrops are very important information for ITS and weather measurement systems. These characteristics can be numerically analyzed by FDTD computer simulation [3–5]. Three-dimensional analysis model of rain attenuation is shown in Fig. 5. Analysis region is defined as $\ell_x \ell_y \ell_z \cdot z = \ell_z$ is the observation plane of the transmitted electromagnetic field to obtain the propagation loss per the distance ℓ_z. Raindrops are distributed randomly in rain region. The incident wave is assumed the Gaussian beam of frequency $20\,\mathrm{GHz}$ radiated from the surface S_0 located at $z = 0$.

In FDTD method, the analysis region is discretized as $x = i\Delta s\,(0 \leqq i \leqq Nx)$, $y = j\Delta s\,(0 \leqq j \leqq Ny)$, $z = k\Delta s\,(0 \leqq j \leqq Nz)$ and $t = n\Delta t\,(0 \leqq n \leqq T)$. When radiation field in the antenna

is expressed by equivalent current $J_y(x, y, z, t)$, the electromagnetic fields at point (i, j, k) at time $n\Delta t$ are calculated by difference equation from Maxwell's equation,

$$
\begin{aligned}
E_y^n(i, j, k) &= c_1 E_y^{n-1}(i, j, k) - c_2 J_y^{n-1}(i, j, k) \\
&\quad + c_3 \left\{ H_x^{n-1}(i, j, k) - H_x^{n-1}(i, j, k-1) - H_z^{n-1}(i, j, k) + H_z^{n-1}(i, j, k) \right\} \quad (3) \\
H_y^n(i, j, k) &= H_y^{n-1}(i, j, k) - c_4 \left\{ E_x^n(i, j, k+1) - E_x^n(i, j, k) - E_z^n(i+1, j, k) + E_z^n(i, j, k) \right\} \quad (4)
\end{aligned}
$$

where

$$
c_1 = \frac{1 - \sigma \Delta t / (2\varepsilon)}{1 + \sigma \Delta t / (2\varepsilon)}, \quad c_2 = \frac{\Delta t / \varepsilon}{1 + \sigma \Delta t / (2\varepsilon)}, \quad c_3 = \frac{\Delta t / \varepsilon}{1 + \sigma \Delta t / (2\varepsilon)} \frac{1}{\Delta s}, \quad c_4 = \frac{\Delta t}{\mu \Delta s} \quad (5)
$$

Here, Δs is cell size and Δt is time step. ε, μ and σ are the dielectric constant, conductivity, and permeability of material. Mur's boundary condition is applied to obtain the electromagnetic fields at the boundary of the analysis space.

In raindrop region, complex dielectric constant ε^* and complex refractive index n_r^* are given by $\varepsilon^* = \varepsilon + \sigma / (j\omega) = n_r^{*2} \varepsilon_0$, $n_r^* = n_1 - j n_2$. Here, we consider the time harmonic electromagnetic field with angular frequency ω. By using $n_r^* = n_1 - j n_2$, dielectric constant and conductivity of raindrops are $\varepsilon = \varepsilon_0 \left(n_1^2 - n_2^2 \right)$, $\sigma = 2 \varepsilon_0 \omega n_1 n_2$. The dielectric constant of background medium is ε_0. In this analysis, the shape of a raindrop is assumed to be cubic and generated by giving the length of a side length of squares a_i, positions of one apex (x_i, y_i, z_i) and relative dielectric constant $\varepsilon_r(i, j, k) = \varepsilon(i, j, k) / \varepsilon_0$.

Equivalent current source of the incident Gaussian beam of y polarization is given by

$$
\begin{aligned}
J_y^n(i, j, 1) &= J_0 \varphi(i, j) p(n), \\
\varphi(i, j) &= \exp\left\{ - \left(\frac{i\Delta s - x_0}{r_0} \right)^2 \right\} \exp\left\{ - \left(\frac{j\Delta s - y_0}{r_0} \right)^2 \right\}
\end{aligned} \quad (6)
$$

where, $x_0 = \ell_x / 2$, $y_0 = \ell_y / 2$, r_0 is the beam spot of incident Gaussian beam.

3.1. Electromagnetic Scattering by Single Raindrop

Scattered field of incident beam wave with the relatively large beam waist by a sphere with diameter a' and dielectric constant $\varepsilon = n_r^{*2} \varepsilon_0$ is approximately obtained by Rayleigh scattering when diameter is smaller than $1/10$ of wavelength. The amplitudes of scattered fields by theoretical and numerical analysis are compared to test the accuracy of FDTD analysis. Scattered field at $\mathbf{r} = (x, y, z)$ from a raindrop at $\mathbf{r}_i = (x_i, y_i, z_i)$ is given by

$$
\mathbf{E}_{scatt}^{(1)} = \frac{e^{-j\beta r_i'}}{r_i'} C E_{inc}(\mathbf{r}_i) \left(\cos\theta_i' \sin\phi_i' \mathbf{i'}_{\theta i} + \cos\phi_i' \mathbf{i'}_{\phi i} \right) \quad (7)
$$

where $C = \beta^2 a'^3 \frac{\varepsilon - \varepsilon_0}{\varepsilon + 2\varepsilon_0}$, $x_i' = x - x_i$, $y_i' = y - y_i$, $z_i' = z - z_i$, $r_i' = |\mathbf{r} - \mathbf{r}_i| = \sqrt{x_i'^2 + y_i'^2 + z_i'^2}$, $\cos\theta_i' = \frac{z_i'}{r_i'}$, $\cos\phi_i' = \frac{x_i'}{\sqrt{x_i'^2 + y_i'^2}}$, $\sin\phi_i' = \frac{y_i'}{\sqrt{x_i'^2 + y_i'^2}}$.

Figure 7 shows the comparison of scattered field amplitudes of cubic and spherical raindrop model. In this figure, a is a side length of cubic model and a' is a diameter of spherical model. Theoretical amplitude is given by Eq. (7) and $|E_{sy}(x, y_0, z)| = |\mathbf{i}_y \cdot \mathbf{E}_{scatt}^{(1)}(x, y_0, z)|$. This figure shows that in this case, cubic model with $a = 1\,(\text{mm})$ in FDTD analysis corresponds to spherical model with $a' = 1.7\,(\text{mm})$ and cubic model with $a = 2\,(\text{mm})$ in FDTD analysis corresponds to spherical model with $a' = 3.6\,(\text{mm})$.

3.2. Electromagnetic Scattering by Randomly Distributed Raindrops

As an example of rain scattering and attenuation of microwave propagation, numerical results in small area are discussed.

For $p(n)$ in Eq. (6) is given by

$$
p(n) = H_0(n\Delta t) \sin(2\pi f n\Delta t), \quad H_0(n\Delta t) = \begin{cases} 1 & |n\Delta t - t_0| \leq T \\ 0 & |n\Delta t - t_0| > T \end{cases} \quad (8)
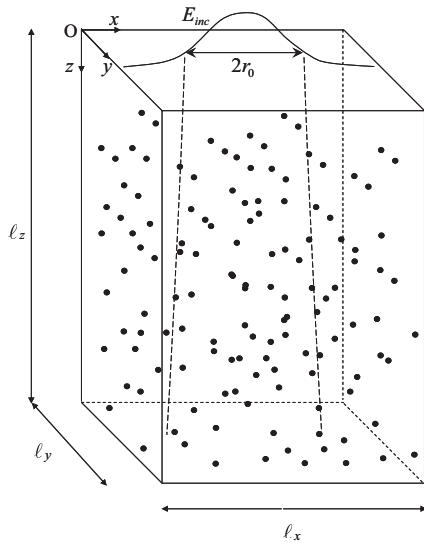$$

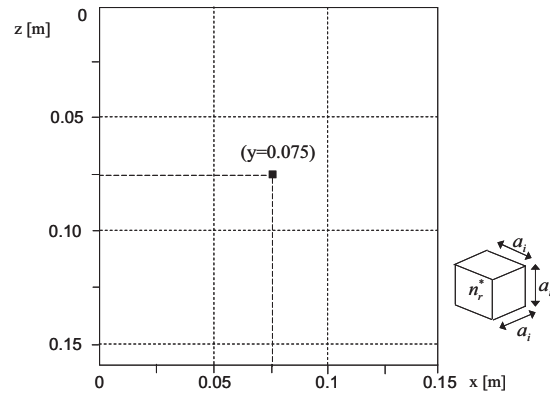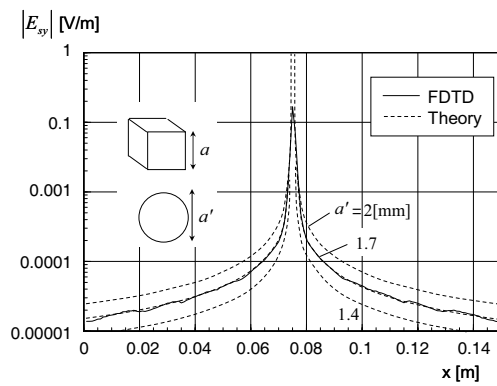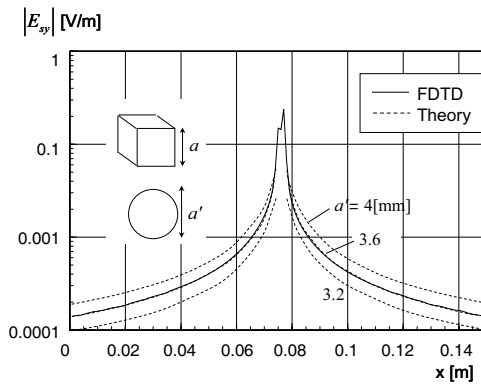Figure 5: Three dimensional analysis model for rain scattering and attenuation.

Figure 6: Analysis model of scattering by single raindrop ($N = 1$, $a_i = 1 \sim 2\,\text{mm}$, $n_r^* = 6.46 - j2.81$).

(a) Cubic model (a = 1 mm)

(b) Cubic model (a = 2 mm)

Figure 7: Comparison of scattered field amplitudes of cubic and spherical raindrop ($y = z = 0.075\,\text{m}$).

and $\ell_x = \ell_y = 5\lambda = 0.075\,(\text{m})$, $x_0 = y_0 = 0.0375\,(\text{m})$, $r_0 = \lambda = 0.015\,(\text{m})$ are used. The electric field of incident wave is shown in Fig. 8. Fig. 9 shows a random media model, where $N = 20$ is the number of raindrops, $a_i = 2\,(\text{mm})$ is a side length of raindrops and $n_r^* = 6.46 - j2.81$ is the refractive index of raindrops. Fig. 10 shows the numerical results of difference ΔE_y between total field and the incident field as scattered field. This figure shows that spherical scattering waves are
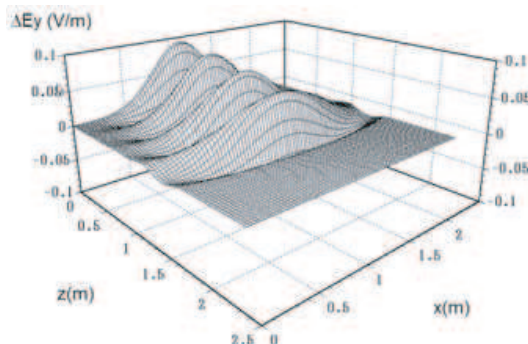

Figure 8: Electric field of incident wave at $y = 0.0375\,(\text{m})$, $t = 250\Delta t = 0.1875\,(\text{ns})$.

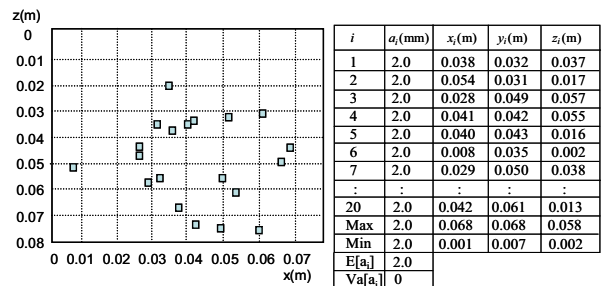| i	a_i(mm)	x_i(m)	y_i(m)	z_i(m)
1	2.0	0.038	0.032	0.037
2	2.0	0.054	0.031	0.017
3	2.0	0.028	0.049	0.057
4	2.0	0.041	0.042	0.055
5	2.0	0.040	0.043	0.016
6	2.0	0.008	0.035	0.002
7	2.0	0.029	0.050	0.038
:	:	:	:	:
20	2.0	0.042	0.061	0.013
Max	2.0	0.068	0.068	0.058
Min	2.0	0.001	0.007	0.002
E[a_i]	2.0			
Va[a_i]	0			

Figure 9: Cross sectional vies of raindrop distribution, $N = 20$, $a_i = 2\,(\text{mm})$, $n_r^* = 6.46 - j2.81$.

generated from the point of scatterers and the distribution of electric field is relatively complex due to the interference of multiple scattering. Dominant term of multiple scattering could be expressed by the first and second terms of multiple scattering effects

$$\mathbf{E}_{scatt} = \sum_{i=1}^{N} \mathbf{p}_i \left(\mathbf{E}_{inc} \right) \mathbf{G} + \sum_{j=1}^{N} \sum_{i=1}^{N} \mathbf{p}_j \left(\mathbf{p}_i \left(\mathbf{E}_{inc} \right) \mathbf{G} \right) \mathbf{G} \tag{9}$$

where $\mathbf{p}_i$ is the equivalent dipole of Rayleigh scattering and $\mathbf{G}$ is Green's function of free space.
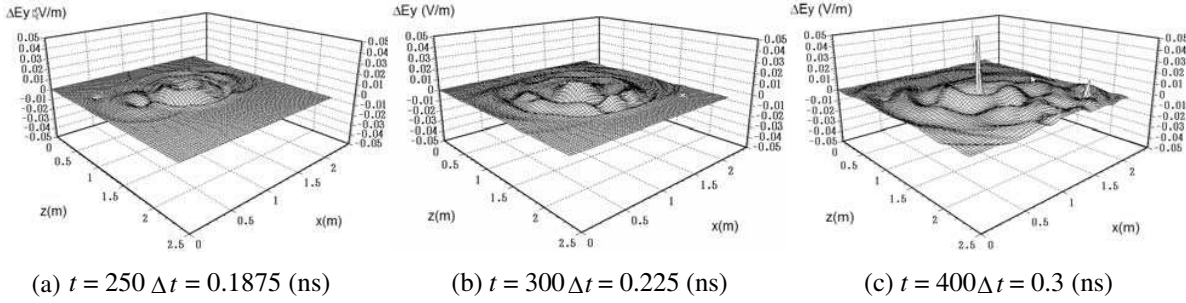

(a) $t = 250\,\Delta t = 0.1875$ (ns)   (b) $t = 300\,\Delta t = 0.225$ (ns)   (c) $t = 400\Delta t = 0.3$ (ns)

Figure 10: Difference of electric field ΔE_y at $y = 0.0375$ (m).

Table 1: Simulation parameters.

Parameter	Values
f: Frequency of incident wave	20 GHz
λ: Wavelength of incident wave	0.015 m (c/f)
ℓ_x: Length of analysis space (x)	0.075 m (5λ)
ℓ_y: Length of analysis space (y)	0.075 m (5λ)
ℓ: Propagation distance	0.065 m (4.3λ)
ℓ_z: Length of analysis space (z)	0.08 m (5.3λ)
V: Volume of rain region	$3.66 \times 10^{-4}\,\mathrm{m}^3$
x_0: Center point of beam (x)	0.0375 m
y_0: Center point of beam (y)	0.0375 m
r_0: Beam spot at $z = 0$	0.015 m (λ)
Δs: Length of a cell	$5 \times 10^{-4}\,\mathrm{m}$ $(\lambda/30)$
Δt: Time increment	0.75 ps $(T/66.7)$
N: Number of raindrops	20
a_i: Length of a side of a raindrop	2 mm
n_r^*: Complex refractive index of raindrops	$6.46 - j2.81$

4. CONCLUSIONS

In this paper, electromagnetic scattering by single raindrop and randomly distributed raindrops in small area are analyzed by FDTD method. Scattered field by single raindrop is compared with Rayleigh scattering theory to test the accuracy of numerical simulation. In random media model, multiple scattering is observed due to the presence of many raindrops. These results show the fundamental characteristics because the scattered wave which does not reach the receiving antenna contributes to rain attenuation. To obtain the rain attenuation in large area which contains hundreds of raindrops, parallel and successive computation of FDTD is indispensable. By using these algorithms, FDTD computation will be demonstrated with less memory requirements and electromagnetic wave scattering and attenuation characteristics in rain can be evaluated numerically.

REFERENCES

1. Oguchi, T., "Electromagnetic wave propagation and scattering in rain and other hydrometers," *IEEE Proc.*, Vol. 71, 1029–1078, 1983.
2. Miyazaki, Y., "Light scattering of laser beams by random micro-inhomogeneities in glasses and polymers," *Jpn. J. Appl. Phys.*, Vol. 13, No. 8, 1238–1248, 1974.
3. Miyazaki, Y., K. Takahashi, and N. Goto, "FDTD parallel computing of microwave scattering and attenuation characteristics due to randomly distributed rainfalls," *PIERS Online*, Vol. 4, No. 2, 245–250, 2008.
4. Rodriguez, G., Y. Miyazaki, and N. Goto, "Matrix-based FDTD parallel algorithm for big areas and its applications to high-speed wireless communications," *IEEE Trans. Antennas and Propagation*, Vol. 54, No. 3, 785–796, 2006.
5. Takahashi, K., Y. Miyazaki, and N. Goto, "FDTD parallel computational analysis of grid-type scattering filter characteristics for medical X-ray image diagnosis," *IEE J. Trans. EIS*, Vol. 127, No. 12, 1973–1981, 2007.

Research and Application on Scattering Matrixes of the Radar Target under Different Polarization Bases

Jian-Xun Liu, **Qiang Xu**, **Hou-Jun Sun**, **and Xin Lv**

School of Information Science and Technology, Beijing Institute of Technology
Beijing 100081, China

Abstract— The deduction and application of transition matrix of radar target polarization scattering matrix under different linear polarization bases were presented in this paper. It was proved to be correct by exemplifying the relation of scatter matrixes of two polarization bases, h/v and $(\hat{45}^\circ, \hat{135}^\circ)$, and applied to virtual polarization synthesis successfully. Moreover, the curve of co-polar scattering parameter to the inclining angle is presented. This method can be used to identify the inclining angle of a plane when it is landing in the dark.

1. INTRODUCTION

Any electromagnetic wave with different polarization can be decomposed according to the orthogonal polarization bases, such as $(\widehat{h}, \widehat{v})$, $(\widehat{R}, \widehat{L})$ or $(\hat{45}^\circ, \hat{135}^\circ)$. The polarization scattering matrix which is used in target detection, recognition and anti-interference varies with different polarization bases. As far as it is concerned, the deduction of the transition matrix between different polarization scattering matrixes is of great importance in practice. The relation between two scattering matrixes of circular and linear polarization is already given in some papers [1]. However, some detection radars work with linear polarization and different inclining angle. Therefore, it is necessary to find the relation between different linear polarization scattering matrixes. The deduction of transition matrix of radar target polarization scattering matrixes under different linear polarization bases were presented in this paper, then, some research on the transition matrix is given in practice and in theory.

2. ABOUT SCATTERING EQUATION AND POLARIZATION

Scattering equation is a basic equation of radar polarimetry, the expression under $(\widehat{x}, \widehat{y})$ polarization base is $E_S(XY) = \overline{S}(XY)E_T(XY)$, which shows the polarization characteristics of a scattering wave.

For example, for $(\widehat{h}, \widehat{v})$ polarization base, the equation is as follows.

$$\begin{bmatrix} E_{Sh} \\ E_{Sv} \end{bmatrix} = \begin{bmatrix} S_{hh} & S_{hv} \\ S_{vh} & S_{vv} \end{bmatrix} \begin{bmatrix} E_{ih} \\ E_{iv} \end{bmatrix} \tag{1}$$

for $(\widehat{x}, \widehat{y})$ polarization base, the equation is as follows.

$$\begin{bmatrix} E_{Sx} \\ E_{Sy} \end{bmatrix} = \begin{bmatrix} S_{xx} & S_{xy} \\ S_{yx} & S_{yy} \end{bmatrix} \begin{bmatrix} E_{ix} \\ E_{iy} \end{bmatrix} \tag{2}$$

Every polarized status can be found on Poincaré sphere. All linear polarizations located on the equator as shown in Figure 1. Only linear polarization bases are studied in this paper.

3. TRANSITION MATRIX

The broken line represents rotated linear polarization base $\widehat{x}$, $\widehat{y}$ as shown in Figure 2, which can be combined by $\widehat{h}$, $\widehat{v}$ as (3) and (4)

$$\widehat{x} = \frac{1}{\sqrt{1 + \text{tg}^2\theta}} \left(\widehat{h} + \text{tg}\theta \cdot \widehat{v} \right) \tag{3}$$

$$\widehat{y} = \frac{1}{\sqrt{1 + \text{tg}^2\theta}} \left(-\text{tg}\theta \cdot \widehat{h} + \widehat{v} \right) \tag{4}$$
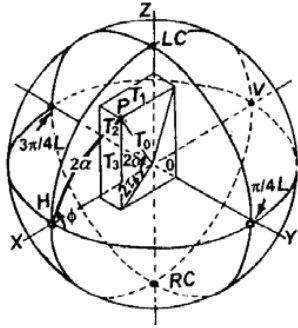
Figure 1: Poincaré sphere.
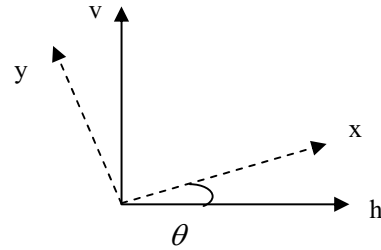
Figure 2: The relation between $(\widehat{x},\,\widehat{y})$ and $(\widehat{h},\,\widehat{v})$.

The coefficient $\frac{1}{\sqrt{1+\mathrm{tg}^2\theta}}$ guarantee $\widehat{x}$, $\widehat{y}$ unitary power density. (3) and (4) can be rewritten as

$$(\widehat{x},\widehat{y}) = \left(\widehat{h},\widehat{v}\right) \frac{1}{\sqrt{1+\mathrm{tg}^2\theta}} \begin{bmatrix} 1 & -\mathrm{tg}\theta \\ \mathrm{tg}\theta & 1 \end{bmatrix} \tag{5}$$

Thus, transition matrix is

$$U = \frac{1}{\sqrt{1+\mathrm{tg}^2\theta}} \begin{bmatrix} 1 & -\mathrm{tg}\theta \\ \mathrm{tg}\theta & 1 \end{bmatrix} \tag{6}$$

The relation between electric field $E(XY)$ and $E(HV)$ is

$$E(XY) = U^{-1}E(HV) \tag{7}$$

where U is unitary matrix, $U^H = U^{-1}$, H represents Hermite transpose.

The relation between two polarization scattering matrixes $\bar{S}(XY)$ and $\bar{S}(HV)$ can be written as

$$\bar{S}(XY) = U^H \bar{S}(HV) U \tag{8}$$

It can be deduced from (6) and (8),

$$\begin{bmatrix} S_{xx} \\ S_{xy} \\ S_{yx} \\ S_{yy} \end{bmatrix} = \frac{1}{1+\mathrm{tg}^2\theta} \begin{bmatrix} 1 & \mathrm{tg}\theta & \mathrm{tg}\theta & \mathrm{tg}^2\theta \\ -\mathrm{tg}\theta & 1 & -\mathrm{tg}^2\theta & \mathrm{tg}\theta \\ -\mathrm{tg}\theta & -\mathrm{tg}^2\theta & 1 & \mathrm{tg}\theta \\ \mathrm{tg}^2\theta & -\mathrm{tg}\theta & -\mathrm{tg}\theta & 1 \end{bmatrix} \begin{bmatrix} S_{hh} \\ S_{hv} \\ S_{vh} \\ S_{vv} \end{bmatrix} \tag{9}$$

by which a new scattering matrix can be obtained from $\bar{S}(HV)$.

4. EXPERIMENT RESULTS AND APPLICATIONS

Firstly, validation of Equation (9) in virtual polarization synthesis is brought forward. When it is required to get horizontal and vertical information where two antennas were placed inclined at 45° and 135°, as shown in Figure 3. The scattering matrix in 45° and 135° polarization bases can be obtained from scattering matrix in horizontal and vertical polarization bases and transition matrix as follows,
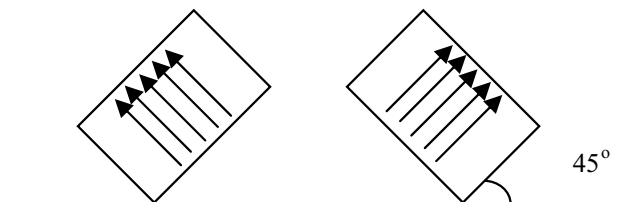

Figure 3: Two rotated antennas.

$$
\begin{bmatrix} S_{4545} \\ S_{45135} \\ S_{13545} \\ S_{135135} \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & 1 & -1 & 1 \\ -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} S_{hh} \\ S_{hv} \\ S_{vh} \\ S_{vv} \end{bmatrix} \tag{10}
$$

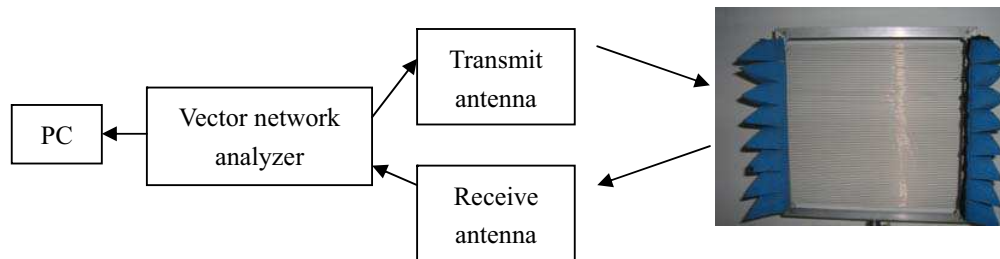The experiment to validate Equation (10) is shown in Figure 4.

Figure 4: Polarization scattering matrix measure.

The target in this experiment is a group of level copper threads. $\overline{S}(HV)$ can be measured at CW 15.8 GHz.

Deduced results and measuring results of S_{4545} and S_{135135} of the target are shown in Figures 5 and 6.
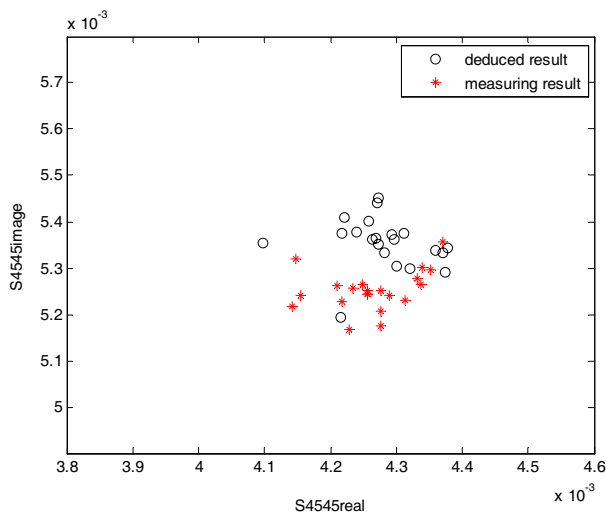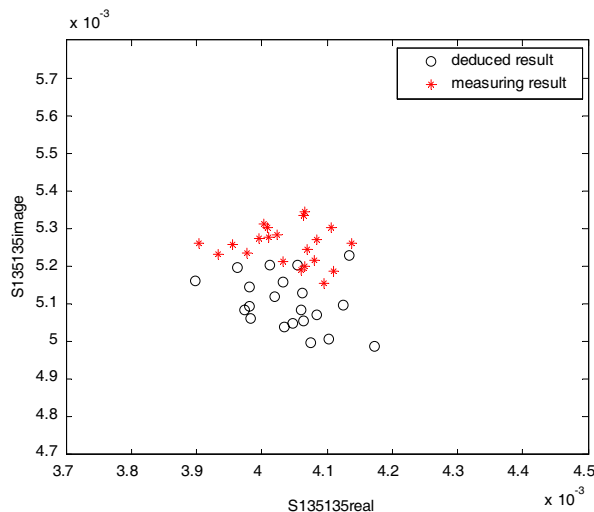

Figure 5: S_{4545} deduced and measuring results.

Figure 6: S_{135135} deduced and measuring results.

Where "∘" is deduced result, and "∗" is measuring result. Horizontal coordinate is real part and vertical coordinate is image part.

The results shown above indicate that Equation (10) and Equation (9) further are correct.

Secondly, application of inclining angle recognition of radar target is presented.

$\overline{S}(HV)$ can be measured when target is placed horizontally, and $\overline{S}(XY)$ can be given when it is placed inclined, thus, the inclining angle can be computed by placing $\overline{S}(HV)$ and $\overline{S}(XY)$ to Equation (9). Furthermore, the relation between co-polarization parameter $\overline{S}(XX)$ and inclining angle is shown as Figure 7.

The upper right-hand point in the figure is $\overline{S}(XX)$ which is measured when radar target' inclining angle $\theta = 0$, the lower left-hand point is $\theta = pi/2$, and the middle point is $\theta = pi/4$. Horizontal coordinate is the real part, vertical coordinate is the image part, where "∗" is theoretical result, and "△" is measuring result. Radar target' inclining angle can be found in the figure according to $\overline{S}(XX)$.

This method can be applied to judge whether a plane is landing or flying-off safely in the dark. The approach is that polarization scattering matrix $\overline{S}(HV)$ under the base $(\hat{h}, \hat{v})$ can be obtained by measuring the polarization matrix of the plane when it is level beforehand, just like $\overline{S}(XY)$

under the same base can be got when the plane is landing slantingly. Accordingly, the inclining angle θ can be calculated by Equation (9), thus, plane' Pitch Attitude is clear.
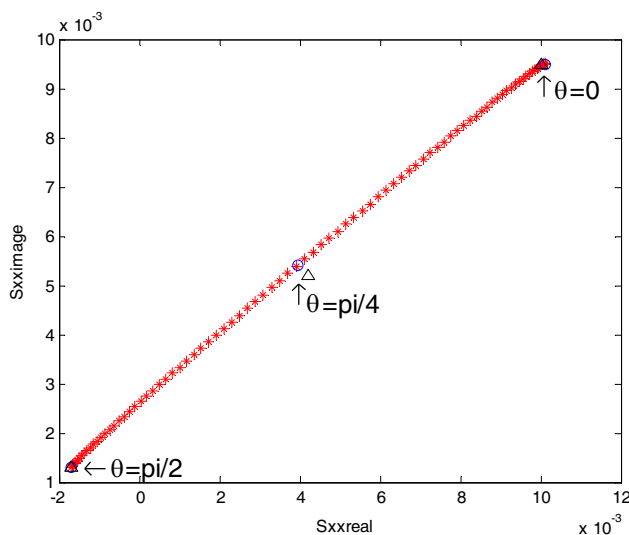

Figure 7: Co-polarization parameter S_{xx} to inclining angle.

5. CONCLUSION

The relation between two scattering matrixes under polarization base h/v and other linear polarization bases was discussed. The transition matrix between different linear polarization bases was deduced which can be applied to virtual polarization synthesis and identifying the inclining angle of radar target.

REFERENCES

1. Zhuang, Z.-W., *Radar Polarization Information Processing and Application*, Beijing, 2000.
2. Agrawal, A. P. and W.-M. Boerner, "On the concept of optimal polarizations in problem of radio wave propagation and probing," *1988 International Symposium on Radio Propagation*, 222–226, CIE, Beijing, China, 1988.
3. Kiseliov, V. K., "Investigation of polarization scattering characteristics metal plates in the near-millimeter waveband by the quasioptical waveguide modeling method," *MSMW'07 Symposiutm Proceedings*, June 2007.

Air Gap Tuning Effect on the Resonant Frequency and Half-power Bandwidth of Superconducting Microstrip Patch

T. Fortaki[1], **S. Benkouda**[1], **M. Amir**[1], **and A. Benghalia**[2]

[1]Electronics Department, University of Batna, Algeria
[2]Electronics Department, University of Constantine, Algeria

Abstract— An accurate analysis of an air gap tuned high T_c superconducting microstrip antenna is presented. To include the effect of the superconductivity of the microstrip patch in the Full_wave analysis of the tunable rectangular microstrip patch, a surface complex impedance is considered. This impedance is determined by using London's equation and the model of Gorter and Casimir. Numerical results obtained are found to be in excellent agreement with the theoretical and experimental data available in the literature. Finally, numerical results for the air gap tuning effect on the operating frequency and bandwidth of the high T_c superconducting microstrip antenna are also presented.

1. INTRODUCTION

Rectangular microstrip patches can find an application in microwave integrated circuits as planar resonators for oscillators and filters. Also, rectangular microstrip patches can be used as resonant antennas fed by means of either coaxial probes, or microstrip lines. Since the bandwidth of microstrip patch resonators and antennas around their operating resonant frequencies is known to be very narrow, it is important to develop accurate algorithms for the computation of those resonant frequencies. The resonant frequency value of the rectangular microstrip patch depends on the structural parameters, and it is evident that if the resonant frequency is to be changed, a new microstrip antenna is needed. In order to achieve tunable resonant frequency characteristics, an adjustable air gap layer can be inserted between the ground plane and the substrate, resulting in a two-layer structure [1, 2].

Apart from tunable microstrip patches, in the last few years, there has been a growing interest in the use of superconducting materials in microwave integrated circuits, which is due to their main characteristics, such as: very small losses, which means low-attenuation and low-noise, very small dispersion up to frequencies of several tens of GHz, smaller devices due to the lower losses, which leads to larger integration density and reduction in the time of propagation of the signals in the circuits [3].

In this paper, a rigorous Full-wave analysis of an air gap tuned high T_c superconducting microstrip antenna is presented. To the best of our knowledge, this structure has not been analyzed previously. Only results for the case of perfectly conducting patch have been reported in the open literature [1, 2]. In Section 2, the authors provide details of the application of the Galerkin's method in the Fourier transform domain to the analysis of high T_c superconducting microstrip antennas with air gaps. In Section 3, the validity of the solution is tested by comparing the computed results with theoretical and experimental data available in the literature. Numerical results for the air gap tuning effect on the operating frequency and half power bandwidth of the high T_c superconducting microstrip patch antenna are also presented. Finally, concluding remarks are summarized in Section 4.

2. OUTLINE OF THE NUMERICAL PROCEDURE

The high T_c superconducting microstrip patch antenna considered in this work is shown in Figure 1. It was obtained by depositing a superconducting patch of thickness e on a dielectric layer. The dielectric layer of thickness d_2 is characterized by the free-space permeability μ_0 and the permittivity ε_0, ε_r (ε_0 is the free-space permittivity and the relative permittivity ε_r can be complex to account for dielectric loss). An adjustable air gap layer of thickness d_1 is inserted between the substrate and the perfectly conducting ground plane. Following the conventional spectral domain approach, the Fourier transform domain relationship between the surface electric field at $z = d_1 + d_2$ and surface current on the patch is given by

$$\begin{cases} \tilde{E}_x(k_x, k_y) = G_{xx}\tilde{J}_x(k_x, k_y) + G_{xy}\tilde{J}_y(k_x, k_y) \\ \tilde{E}_y(k_x, k_y) = G_{yx}\tilde{J}_x(k_x, k_y) + G_{yy}\tilde{J}_y(k_x, k_y) \end{cases} \tag{1}$$
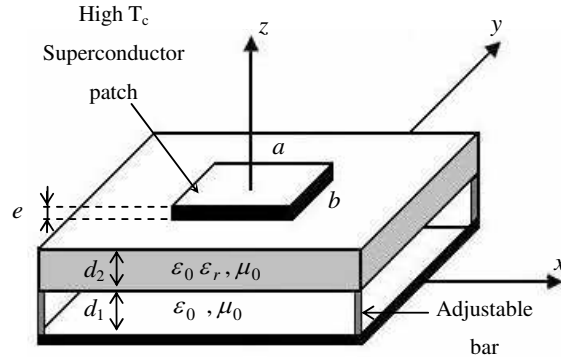
Figure 1: Geometry of the tunable high T_c superconducting microstrip antenna.

where G_{xx}, G_{xy}, G_{yx}, and G_{yy} are the components of the spectral dyadic Green's function. They can be deduced from [4, Equation (20)] together with a simple transformation of the (TM, TE) representation to the (x, y) coordinates. The surface electric fields at $z = d_1 + d_2$ can be written as a superposition of an electric field in the patch and another out of the patch, this yields

$$\begin{cases} \tilde{E}_x = \tilde{E}_x^i + \tilde{E}_x^o \\ \tilde{E}_y = \tilde{E}_y^i + \tilde{E}_y^o \end{cases} \tag{2}$$

The electric field in the superconducting patch is given by

$$\begin{cases} \tilde{E}_x^i = Z_s \tilde{J}_x \\ \tilde{E}_y^i = Z_s \tilde{J}_y \end{cases} \tag{3}$$

where Z_s is the surface impedance of the superconducting patch. When the thickness of the superconducting patch is less than three times the penetration depth λ at a temperature $T = 0\,\mathrm{K}$, the surface impedance can be approximated as follows

$$Z_s = \frac{1}{e\sigma} \tag{4}$$

where σ is the complex conductivity of the superconducting film. It is determined by using London's equation and the model of Gorter and Casimir as [5]

$$\sigma = \sigma_1 - i\sigma_2 \tag{5}$$

The resistive part of the complex conductivity (σ_1) may arise from normal electron conduction within non-superconducting grains and scattering from grain boundaries, flux vibration at pinning centers and normal electron conduction due to thermal agitation in the superconducting state. The temperature dependence of σ_1 is as follows:

$$\sigma_1 = \sigma_n \left(T/T_c \right)^4 \tag{6}$$

where σ_n is the normal conductivity. The reactive part of the conductivity $(-i\sigma_2)$ arises from the lossless motion of the superconducting carries which may be derived from the Lorentz-force equation as [5]

$$\sigma_2 = \frac{1}{\omega\mu_0[\lambda(T)]^2} \tag{7}$$

with

$$\lambda(T) = \frac{\lambda(0)}{\sqrt{1 - \left(\frac{T}{T_c}\right)^4}} \tag{8}$$

In the above two equations, ω is the angular frequency and T_c is the transition temperature of the superconductor. Substituting Equations (2) and (3) in Equation (1) yields

$$\begin{cases} \tilde{E}_x^o = (G_{xx} - Z_s)\tilde{J}_x(k_x, k_y) + G_{xy}\tilde{J}_y(k_x, k_y) \\ \tilde{E}_y^o = G_{yx}\tilde{J}_x(k_x, k_y) + (G_{yy} - Z_s)\tilde{J}_y(k_x, k_y) \end{cases} \tag{9}$$

Now that we have include the effect of the superconductivity of the patch in the Green's function formulation, the well-known Galerkin procedure of the moment method can be easily applied to Equation (9) to obtain the resonant frequencies and the bandwidths of the resonant modes of the air gap tuned high T_c superconducting microstrip antenna shown in Figure 1.

3. RESULTS AND DISCUSSION

3.1. Comparison of Numerical Results

In order to confirm the computation accuracy, our numerical results are compared with those obtained from the cavity model [6]. The patch is fabricated with a YBCO superconducting thin film with parameters $\sigma_n = 10^6$ S/m, $T_c = 89$ K, $\lambda(0) = 140$ nm and $e = 350$ nm. The patch of length $a = 1630\,\mu$m is printed on a lanthanum aluminate substrate with $\varepsilon_r = 23.81$ and $d_2 = 254\,\mu$m. The high permittivity of the lanthanum aluminate substrate allows the growth of high-quality (low surface resistance) YBCO films [6]. Table 1 summarizes our computed resonant frequencies and those obtained via the cavity model [6] for three different patch widths and differences between these two results of less than 2% are obtained. In Table 2, we have also compared our results with the experimental and theoretical data of Abboud et al. [7] in the case where the tunable microstrip antenna is fabricated by a perfectly conducting patch. The comparison is done for two different air separations. Not that the small differences between our results and the calculated data in [7] can be attributed to the fact that the cavity model used for computing the resonant frequencies in [7] do not account rigorously for the effect of fringing fields. The above comparisons show a good agreement between our results and those of the literature. This validates the theory presented in this paper.

Table 1: Comparison of our calculated resonant frequencies with those of Richard et al. [6]; $a = 1630\,\mu$m, $\sigma_n = 10^6$ S/m, $T_c = 89$ K, $\lambda(0) = 140$ nm, $e = 350$ nm, $T = 77$ K, $\varepsilon_r = 23.81$, $d_1 = 0$ and $d_2 = 254\,\mu$m.

$b\,(\mu\text{m})$	Resonant frequencies (GHz)		Error (%)
	Cavity model [6]	Our results	
935	28.95	28.76	0.66
1050	26.12	26.29	0.65
1100	25.05	25.33	1.12

Table 2: Comparison of our calculated resonant frequencies with those of Abboud et al. [7] in the case where the tunable microstrip antenna is fabricated by a perfectly conducting patch; $\varepsilon_r = 2.33$.

a (mm)	b (mm)	d_2 (mm)	Resonant frequencies (GHz)			
			$d_1 = 0$ mm		$d_1 = 0.5$ mm	
			Measured [7]	Our results	Calculated [7]	Our results
57	38	3.175	2.31	2.41	2.51	2.55
45.5	30.5	3.175	2.89	2.96	3.05	3.11
29.5	19.5	3.175	4.24	4.43	4.42	4.59
17	11	1.524	7.87	7.75	8.16	8.24

3.2. Air Gap Tuning Effect on the Superconducting Antenna Characteristics

The effect of the air separation on the resonant frequency and bandwidth of the high T_c superconducting microstrip antenna shown in Figure 1 is investigated. The superconducting patch of dimension 8×5 mm is fabricated with a YBCO thin film and the substrate thickness is $d_2 = 0.4$ mm. In Figure 2(a), the resonant frequency versus the air separation d_1 for various substrate materials is shown. These materials are Gallium arsenide (GaAs), oxide of Magnesium (MgO) and oxide of Berilium (BeO) with relative permittivities $\varepsilon_r = 6.6$, 9.6 and 12.5, respectively. It is observed that when the air separation grows, the resonant frequency increases rapidly until achieving a maximum operating frequency at a definite air separation $d_{1\text{fmax}}$. Note that the effect of the air gap is more

pronounced for small values of d_1. When the air separation exceeds $d_{1\text{fmax}}$, increasing the air gap width will decrease slowly the resonant frequency. Extreme care should be taken when designing an antenna with thin air gap; since small uncertainty in adjusting d_1 can result in an important detuning of the frequency.

Graphical representation of the half-power bandwidth is shown in Figure 2(b). Note that it increases monotonically with increasing air separation.
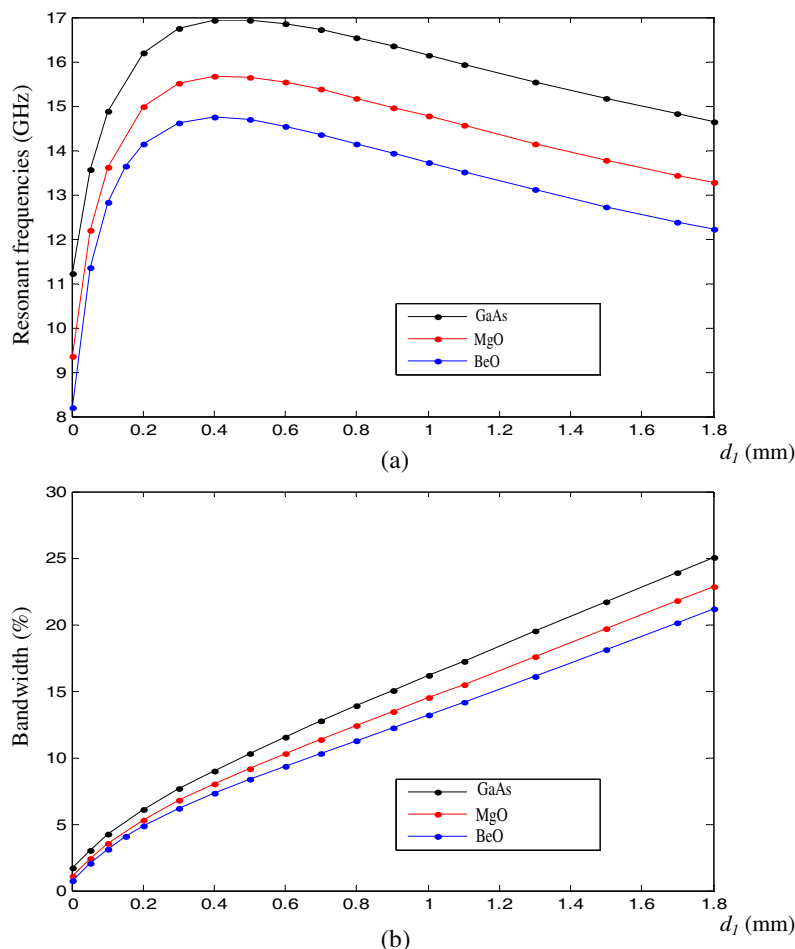

Figure 2: Resonant frequency and bandwidth of a superconducting microstrip antenna versus the air separation; $a = 8\,\text{mm}$, $b = 5\,\text{mm}$, $\sigma_n = 7.46 10^6\,\text{S/m}$, $T_c = 89\,\text{K}$, $\lambda(0) = 180\,\text{nm}$, $e = 150\,\text{nm}$, $T = 77\,\text{K}$ and $d_2 = 0.4\,\text{mm}$.

4. CONCLUSIONS

A spectral domain approach has been used for the numerical calculation of the characteristics of a high T_c superconducting microstrip antenna with an air gap. To include the effect of the superconductivity of the microstrip patch in the Full_wave analysis of the tunable rectangular microstrip patch, a surface complex impedance has been considered. This impedance has been determined by using London's equation and the model of Gorter and Casimir. The numerical results obtained via Galerkin's method in the Fourier transform domain have been compared with theoretical and experimental data, and good agreement has been found. Computations show that the air separation can be adjusted to have the maximum operating frequency of the superconducting microstrip antenna. The half-power bandwidth, on the other hand, increases monotonically with increasing the air gap width. Extreme care should be taken when designing a microstrip antenna with thin air gap; since small uncertainty in adjusting the air separation can result in an important detuning of the frequency.

REFERENCES

1. Gurel, C. S. and E. Yazgan, "New computation of the resonant frequency of a tunable equilateral triangular microstrip patch," *IEEE Trans. Microwave Theory Tech.*, Vol. 48, No. 3, 334–338, 2000.

2. Gurel, C. S. and E. Yazgan, "Resonant frequency of an air gap tuned circular disc microstrip antenna," *Int. J. Electron.*, Vol. 87, No. 8, 973–979, 2000.

3. El-Ghazaly, S. M., R. B. Hammond, and T. Itoh, "Analysis of superconducting microwave structures: Application to microstrip lines," *IEEE Trans. Microwave Theory Tech.*, Vol. 40, No. 3, 499–508, 1992.

4. Fortaki, T. and A. Benghalia, "Rigorous full-wave analysis of rectangular microstrip patches over ground planes with rectangular apertures in multilayered substrates that contain isotropic and uniaxial anisotropic materials," *Microwave Opt. Technol. Lett.*, Vol. 41, No. 6, 496–500, 2004.

5. How, H., R. G. Seed, C. Vittoria, D. B. Chrisey, J. S. Horwitz, C. Carosella and V. Folen, "Microwave characteristics of high T_c superconducting coplanar waveguide resonator," *IEEE Trans. Microwave Theory Tech.*, Vol. 40, No. 8, 1668–1673, 1992.

6. Richard, M. A., K. B. Bhasin, and P. C. Claspy, "Superconducting microstrip antennasd: An experimental comparison of two feeding methods," *IEEE Trans. Antennas Propagat.*, Vol. 41, No. 7, 967–974, 1993.

7. Abboud, F., J. P. Damiano, and A. Papiernik, "Accurate model for the input impedance of coax-fed rectangular microstrip antenna with and without air gaps," *Proceedings of Sixth International Conference on Antennas and Propagation*, Vol. 1, 102–106, April 1989.

Radiation Characteristics of a Wideband Triangular Antenna for Wireless Communications

Adel M. Abdin

Department of Communications, Faculty of Engineering, Shorouk Academy
Shorouk City, Cairo, Egypt

Abstract— A new design of a wideband printed triangular antenna with a slot and steps is fed by a microstrip transmission line is presented which is suitable for ultra-wide band application. Antenna impedance and bandwidth are optimized by testing the suitable slot dimension and position and using the suitable dimension of the ground plane. A prototype antenna was fabricated on a dielectric FR4 (substrate thickness = 1.5 mm, $\varepsilon_r = 4.65$) and measured for VSWR characteristics. The antenna operates in the (1.88109–3.01712) GHz frequency range and has a bandwidth of 46% within 2 : 1 VSWR. Range of antenna gain is between 2.2 dBi to 3.08 dBi which indicate that the proposed antenna can be used for various applications. Details of the antenna design and a comparison of the simulated and measured results are presented and discussed.

1. INTRODUCTION

Microstrip antennas (MSAs) have several advantages over conventional microwave antennas (MAs). They have small volume, light weight, low profile, and planar configuration. These antennas are used in UHF to millimeter-wave frequency bands. The main disadvantage of these antennas is their narrow bandwidth (BW). Fortunately, the BW can be increased by using a low dielectric constant with a thick substrate, stacked electromagnetically coupled or aperture-coupled patches, planar gap-coupled and directly coupled multi-resonators, log-periodic configurations, ferrite substrates, and impedance-matching techniques [1–3]. Ultra-wide band (UWB) antennas have had many applications in communication systems with UWB and spread-spectrum features in radar systems. A dual wideband folded microstrip-patch antenna was presented in the 3.5–4 GHz frequency range for wireless local-area network (WLAN) applications [4]. In [5–7] different slot antennas with microstrip line feeding for wideband applications are presented. The recent systems need an antenna with multiband or UWB characteristics for future communication systems.

In this paper, a compact, wideband, printed, microstrip antenna suitable for ultra-wide band application is presented. The proposed antenna covering major wireless communication bands like GSM, AWS, WCDMA, UMTS, DSR, Wi. Bro, ISM application (Wi-Fi), Wi-max, Fixed microwave links and DMB, Onboard aircrafts internet based on the AMSS.

2. DESCRIPTION OF THE PROPOSED ANTENNA

The triangular patch with a rectangular slot is presented. More than one technique to have wideband is considered. The ground plane is chosen to be a partial ground and a slot with rectangular shape is considered. The bandwidth (BW) value depends on the dimension and position of the slot and on the dimension of the partial ground. These dimensions are optimized using Zeland IE3D software to have the required response.

The configuration of the optimized antenna is shown in Figure 1. The proposed antenna comprises a $50\,\Omega$ microstrip line feed of width $W_f = 3$ mm, the ground plane of length $L_g = 50$ mm and width $W_g = 82$ mm, and a rectangular patch with steps each one has length and width (1.5×1 mm). It is etched on a dielectric FR4 substrate of thickness $h = 1.5$ mm and a relative permittivity $\varepsilon_r = 4.65$.

The proposed antenna is optimized by using Zeland IE3D software which is based on the method of moment technique [8].

3. SIMULATED RESULTS

3.1. VSWR Characteristic

Numerical simulations and measurements were conducted on the compact wideband printed triangular antenna with a slot. Simulation and optimization are based on Zeland. It should be noted that the antenna was simulated over the full range of frequencies (0–14) GHz, which is the limit of the Zeland software. The VSWR characteristic of the antenna is presented in Figure 2. It shows that it has an impedance bandwidth for (VSWR $\leq$ 2) from (1.88109–3.01712) GHz.
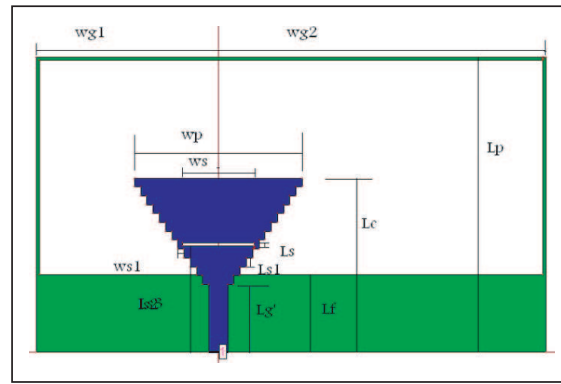
Figure 1: Geometry of printed rectangular slot antenna fed by $50\,\Omega$ microstrip line. Designed structure parameters are: $W_{g1} = 29.2992\,\text{mm}$, $W_{g2} = 52.678\,\text{mm}$, $L_g = 50\,\text{mm}$, $L_f = 13.5\,\text{mm}$, $L_{g'} = 11.5\,\text{mm}$, $W_p = 27\,\text{mm}$, $W_S = 11.5\,\text{mm}$, $L_p = 49.5\,\text{mm}$, $L_S = 0.5\,\text{mm}$, $L_{S1} = 1.5\,\text{mm}$, $W_{S1} = 1\,\text{mm}$, $L_{s2} = 18\,\text{mm}$, $h = 1.5\,\text{mm}$, $\varepsilon_r = 4.65$, $t = 0.035\,\text{mm}$.
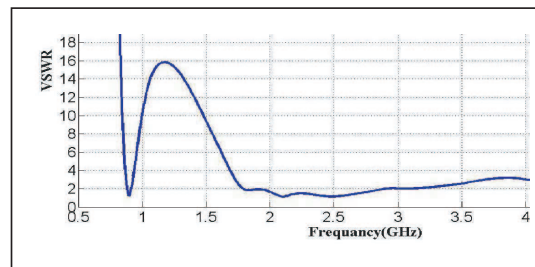
Figure 2: Simulated VSWR curve of the proposed antenna.

3.2. Input Impedance and Smith Chart

Figure 2 illustrates the simulated VSWR curve of the optimal design of the antenna. The corresponding Smith chart curve is plotted in Figure 3. It is noticed on the Smith Chart that the input impedance around the impedance matching point within the circle of (VSWR $\leq$ 2) as illustrated in Figure 3.
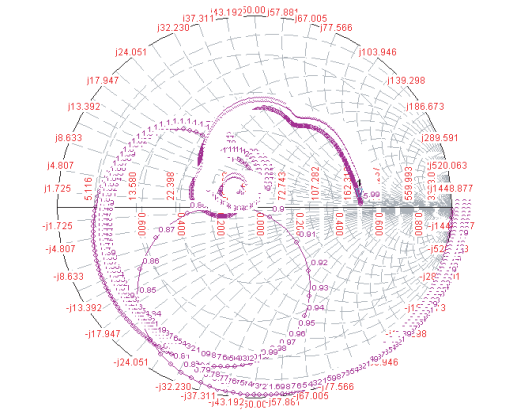
Figure 3: Simulated Smith chart of the presented antenna.

3.3. Current Distribution

The detailed EM behavior of the antenna can only be revealed by examining the current distributions or radiation patterns. The typical current distributions on the antenna close to the resonance frequencies are plotted in Figure 4.
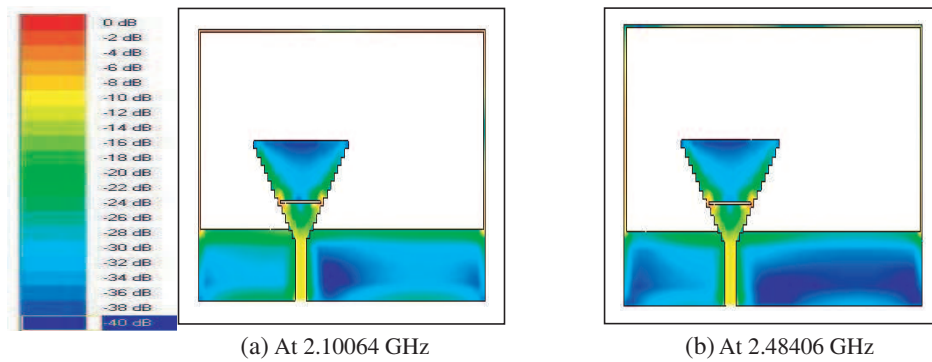
(a) At 2.10064 GHz   (b) At 2.48406 GHz

Figure 4: Simulated current distributions at different resonant frequencies of the proposed antenna.

3.4. Radiation Pattern

Simulated results of the radiation patterns of the proposed antenna are presented in Figure 5. The results include the E-plane and H-plane at different frequencies 2.10064 and 2.48406 GHz respectively.
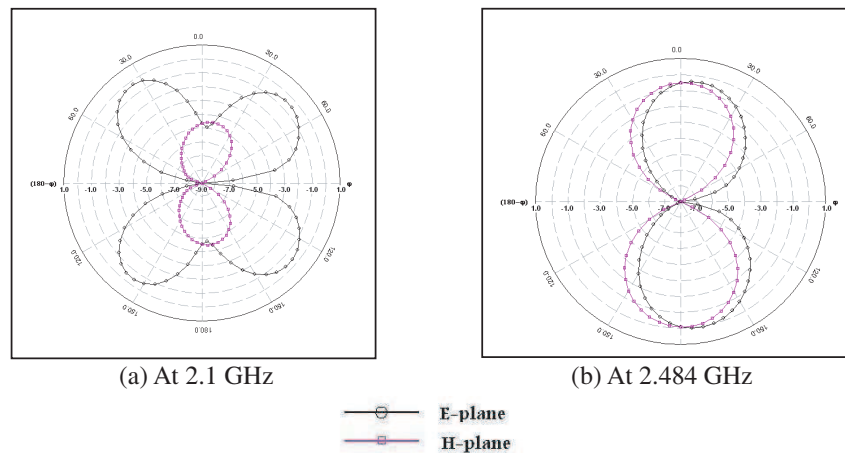

(a) At 2.1 GHz   (b) At 2.484 GHz

E-plane
H-plane

Figure 5: Simulated radiation patterns of the antenna. (a) $F = 2.10064$ GHz. (b) $F = 2.48406$ GHz.

3.5. Gain

Figure 6 shows the simulated antenna gain in a frequency range (1.79713–2.86897) GHz. The gain variations are between 2.2 dBi and 3.08 dBi. The antenna has a slightly higher gain through out the entire band.
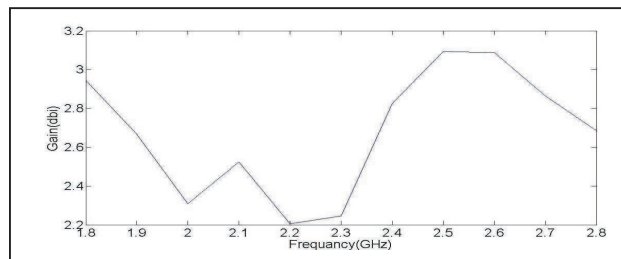

Figure 6: Simulated antenna gain in a frequency range (1.79713–2.86897) GHz.

4. MEASURED RESULTS

New compact wideband antenna with partial ground was fabricated and tested (Figure 7). The simulated and measured return loss are shown in Figure 8. The fundamental resonant mode is

at (2.1 and 2.48) GHz for the simulated results and (2.27660 and 2.725074) GHz for the measured results. The -10 dB impedance bandwidth is (1.79713–2.86897) GHz for the simulated result and (1.88109–3.01712) GHz for the measured antenna. There are discrepancies between the measured and simulated results. This is due to the simulation do not correctly model substrate losses as the real antenna. Furthermore, the 50-SMA feeding port is not taken into account in the simulation so as to ease the computational requirements and also due to the fabrication inaccuracy and some technological challenges. Figure 9 shows the proposed antenna connected to S-parameter network analyzer (Ailment-8719ES) and the measured return loss.
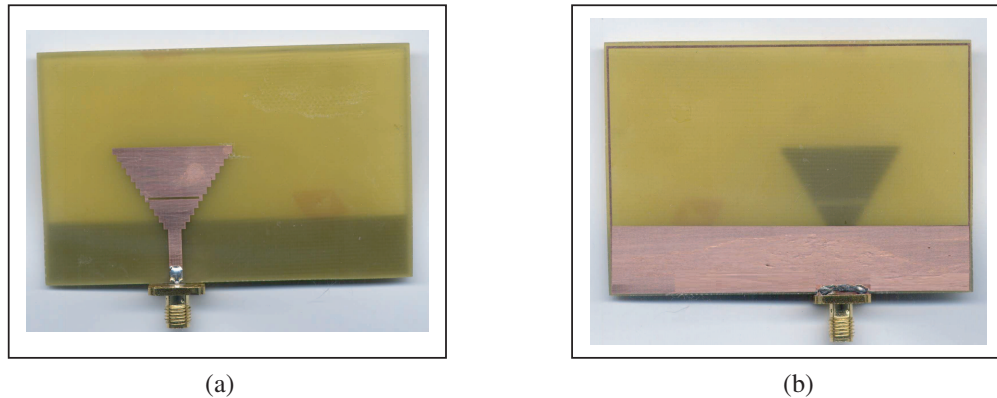

(a)   (b)

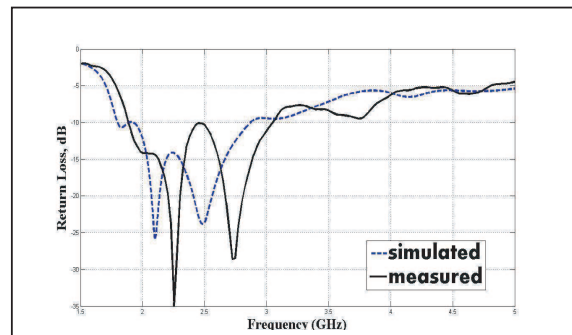Figure 7: Photo of fabricated antenna with SMA connector (a) from front view, (b) from back view.
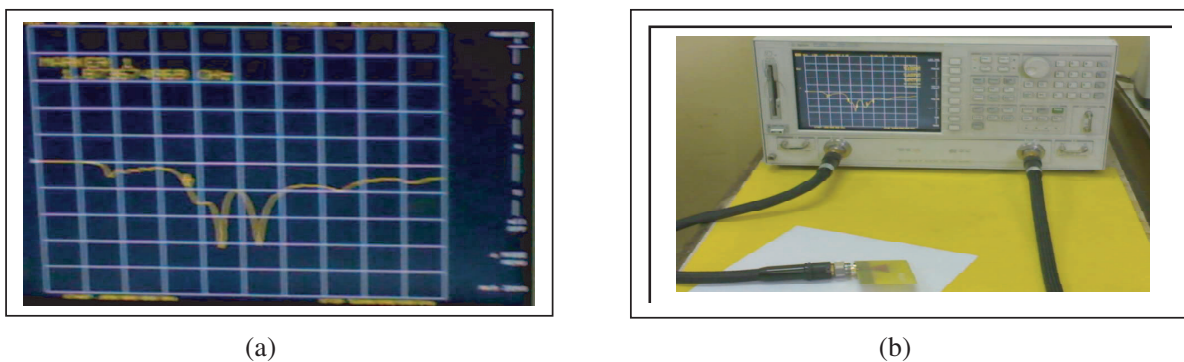

Figure 8: Simulated and measured return loss.

(a)   (b)

Figure 9: (a) The measured return loss. (b) Shows the proposed antenna connected to S-parameter network analyzer (Ailment-8719ES).

The simulated and measured impedance curves and the measured Smith chart of the antenna are shown in Figure 10.
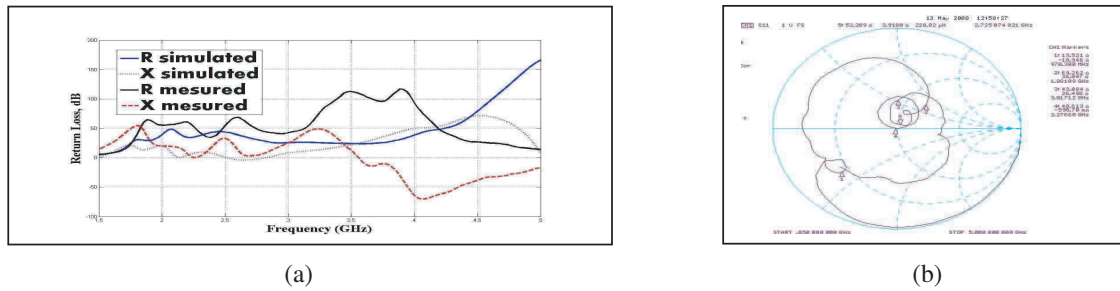
(a)   (b)

Figure 10: (a) Simulated and measured impedance curves of the compact wideband antenna. (b) Measured Smith chart of the proposed antenna.

5. CONCLUSIONS

A new compact wideband antenna was presented. A band of (1.88109–3.01712) GHz was achieved. The gain is good enough over the band of interest. Experimental results reveal that the antenna shows very good performance in a wideband. The radiation properties of the antenna were tested and discussed. The simulated results were conducted by Zeland IE3D software. They have good agreement with the measured ones.

REFERENCES

1. Kumar, G. and K. P. Ray, *Broadband Microstrip Antennas*, Artech House, 2003.
2. Ray, K. P., "Broadband, dual-frequency and compact microstrip antennas," Ph.D. Thesis, Indian Institute of Technology, Bombay, India, 1999.
3. Wong, K. L. and W. H. Hsu, "Broadband triangular microstrip antenna with U-shaped slot," *Electronics Letters*, Vol. 33, No. 25, 2085–2087, 1997.
4. Sittironnarit, T. and M. Ali, "Analysis and design of a dual band folded microstrip patch antenna for handheld device application," *IEEE Southeast Conference Proceedings*, 255–258, 2002.
5. Lee, H. L., H. J. Lee, J. G. Yook, and H. K. Park, "Broadband planar antenna having round corner rectangular wide slot," *Proc. IEEE Antennas and Propagation Society*, Vol. 2, 460–463, Jan. 16, 2002.
6. Sze, J. Y. and K. L. Wong, "Bandwidth enhancement of a microstripline-fed printed wide-slot antenn," *IEEE Trans. Antennas Propag.*, Vol. 49, No. 7, 1020–1024, July 2001.
7. Kim, M. K., K. Kim, Y. H. Suh, and I. Park, "A T-shaped microstripline-fed wide slot antenna," *Proc. IEEE Int. Symp. APS*, 1500–1503, July 2000.
8. IE3D Zeland software version 11.54.

Characteristics of a Multi-bandwidth Gear Microstrip Antenna Using a Taper for Feeding

Adel M. Abdin

Dept. of Communications, Faculty of Engineering, Shorouk Academy
Shorouk City, Cairo, Egypt

Abstract— In this paper, a gear microstrip antenna (GMA) using a taper as a feeder is presented. Several techniques are demonstrated to overcome its inherent disadvantage of narrow bandwidth of microstrip antenna. For broadening the bandwidth, different configurations for the presented antenna using various materials are suggested. Partial ground with different dimensions and the GMA with a slot and without a slot are tested. Different substrates such as RT5880 substrate of thickness $h = 0.78$ mm and a relative permittivity $\varepsilon_r = 2.2$, FR4 substrate of $h = 1.5$ mm and $\varepsilon_r = 4.65$ and RT/Duroid 6010 substrate of $h = 1.25$ mm and $\varepsilon_r = 10.5$ are also taken into account in the analysis. All these cases are analyzed using Zeland IE3D software and High Frequency Structure Simulator (HFSS). As much as 163.6% and 26% SWR < 2 bandwidths are reported for the best case. Application of the GMA in modern wireless systems covering AWS, DCS, DECT, PCS, PHS, GSM, WCDMA, UMTS, DSR, Wi.Bro, ISM, DMB, C band and Ku band satellite, IMT2000, Wi-max, Wi-Fi, Fixed microwave links, DAB, and Bluetooth bands is widely demonstrated. The antenna comprises at two bands, from $(1.5–15)$ GHz and $(20–26)$ GHz having $(2.9–4.5)$ dBi gain. This design has achieved a high average gain and low cross-polarization across its wide operating frequencies. The measured and simulated results are presented and discussed.

1. INTRODUCTION

In recent years ultra-wideband (UWB) systems have received a lot of attention. This is due to the promising applications in sensors, radars, home systems, and modern wireless systems. Therefore, the antenna of wide-bandwidth is required for a UWB system. Recently, microstrip slots antennas have been studied [1–3]. The study shows that by using several slot shapes and a tapered feed line, a significant widening of the bandwidth could be achieved [4].

In this paper, a gear microstrip antenna (GMA) using a taper as a feeder is presented. Several techniques are demonstrated to overcome the inherent disadvantage of microstrip antenna which is the narrow bandwidth. For broadening the bandwidth, different configurations using different techniques are presented. Various materials, partial ground with different dimensions, the GMA with a slot and without a slot, and a tapered feed line are suggested and tested. Different substrates such as RT5880 substrate of thickness $h = 0.78$ mm and a relative permittivity $\varepsilon_r = 2.2$, FR4 substrate of $h = 1.5$ mm and $\varepsilon_r = 4.65$ and RT/Duroid 6010 substrate of $h = 1.25$ mm and $\varepsilon_r = 10.5$ are taken into account in the analysis. All these cases are analyzed using Zeland IE3D software and HFSS. The antenna comprises at two bands, from $(1.5–15)$ GHz and $(20–26)$ GHz, having $(2.9–4.5)$ dBi gain. This enables the GMA antenna to work with modern wireless systems such as mobile and satellite applications.

2. ANTENNA STRUCTURE

The antenna structure is shown in Fig. 1. The RT5880 substrate has a dielectric constant of $\varepsilon_r = 2.2$. The thickness of the substrate is 0.78 mm and the dimension is 45.5 mm × 44 mm. A rectangular slot with half circle at each edge is etched at the centre of the patch. A taper microstrip line is used as a feeder line with width equal to 1.25 mm at the antenna input and 2.74 at the other side.

3. RESULTS AND DISCUSSION

The two numerical techniques used to model the structure are Zeland IE3D software and HFSS. The simulated return loss plot using HFSS for the different substrates is shown in the Fig. 2. The S_{11} parameter is also tested for the different substrates using the two numerical techniques and their results are very similar. Figs. 3, 4, and 5 represent the return loss for the substrates RT5880, FR4, and RT/Duroid 6010 respectively. The return loss is also tested using Zeland IE3D software for different lengths of the ground plane (Hg). Fig. 6 shows that the length $Hg = 20$ mm gives the widest bandwidth which is chosen in the design. The bandwidth is also tested for GMA with

a slot and without the slot. The widest bandwidth is due to the multiple resonances introduced due to the square slot combined with the half circles ones. One can deduce from Fig. 7 that, the bandwidth is wider by more than 5 GHz for the existence of the slot. It was found from the previous comparisons that the return loss plot is better for RT5880 substrate with $Hg = 20$ mm, and with a slot than the others. An impedance bandwidth of $(1.5–15)$ GHz and $(20–26)$ GHz that corresponds to 163.6 and 26 per cent of the centre frequency is obtained.

The simulated Smith chart curve is illustrated in Fig. 8. Most of the resonant loops in the Smith chart are inside the circle of VSWR < 2.
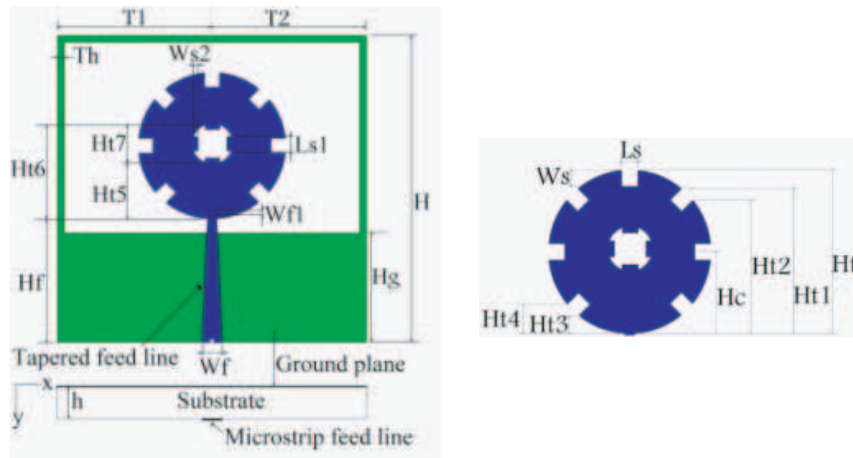

Figure 1: Geometry for a typical gear microstrip antenna. $T1 = 21$, $T2 = 24.5$, $Hg = 20$, $H = 44$, $Th = 1$, $Wf = 2.74$, $Wf1 = 1.250978$, $Hf = 21.9397$, $Ht = 19.9295$, $Hc = 9.9695$, $Ht1 = 17.6911$, $Ht2 = 16.2691$, $Ht3 = 2.1721$, $Ht4 = 3.5709$, $Ls = 2$, $Ws = 1.9491$, $Ht5 = 7.5117$, $Ht6 = 11.2541$, $Ht7 = 3.7424$, $Ls1 = 2.05$, $Ws1 = 0.65839$ (all dimension in mm).
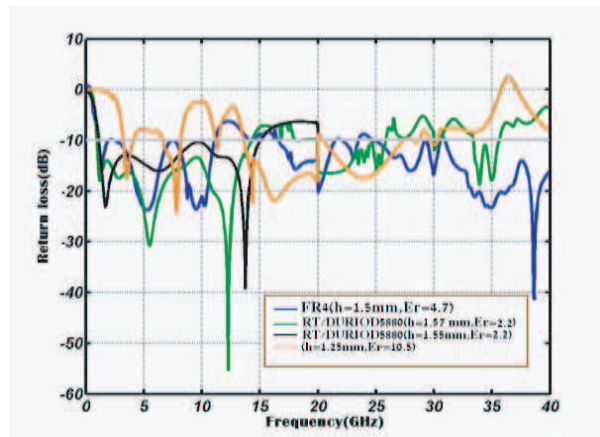

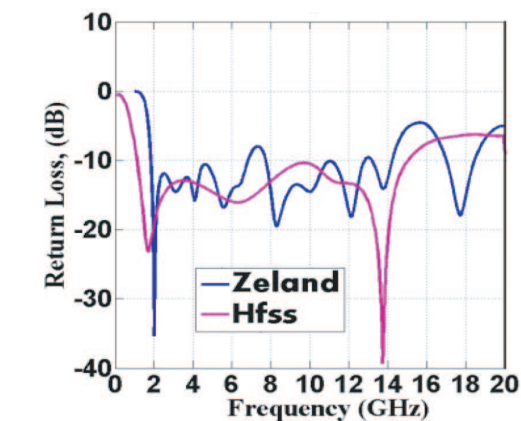Figure 2: The simulated return loss plot for the different substrates.

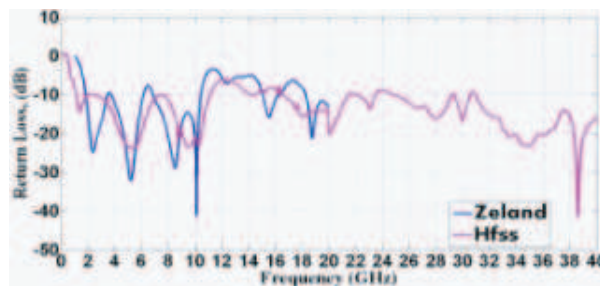Figure 3: Simulated S_{11} for RT5880 substrate.

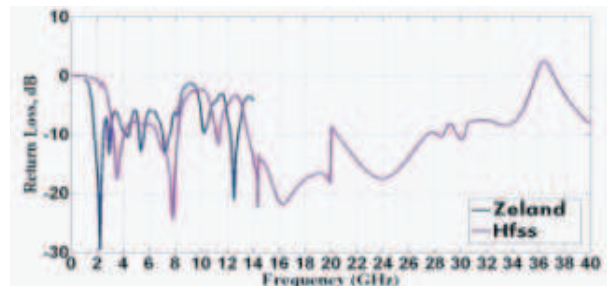Figure 4: Simulated S_{11} for FR4 substrate.

Figure 5: Simulated S_{11} for RT/Duroid 6010 substrate.

Figure 9 illustrates the variation in the input impedance against frequency for the proposed antenna. The impedance behavior was mostly controlled by the type of the substrates, dimensions of the ground plane, and the existence of the slot. Fig. 9 shows that the real the impedance and the imaginary parts of the GMA is changed around $50\,\Omega$ and $0\,\Omega$ respectively. The $50\,\Omega$ SMA feeding port is not taken into account in all of the simulations so as to ease the computational requirement.
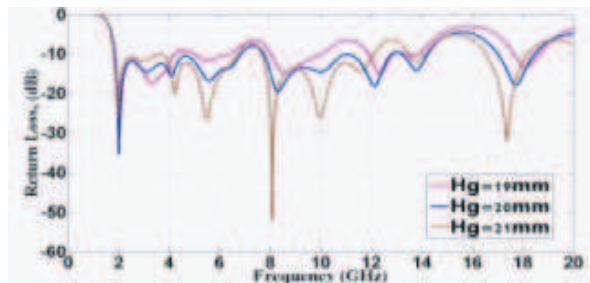

Figure 6: Simulated return loss curves of GMA for different ground lengths.
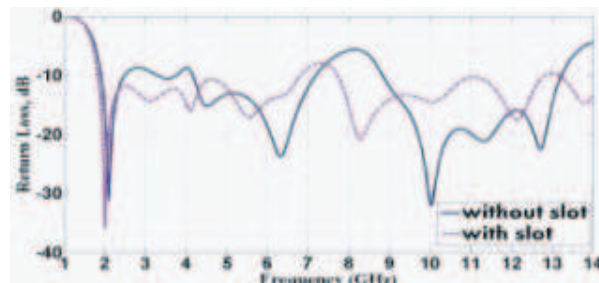

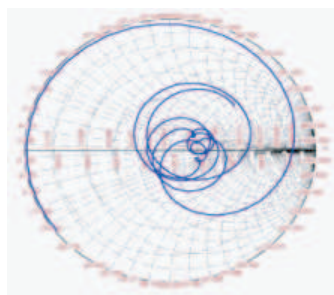Figure 7: The effective of the slot on the return loss.

Figure 8: Simulated Smith chart of the compact wide-band GMA.
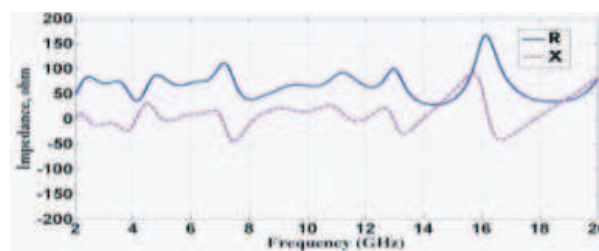

Figure 9: Simulated impedance curves of the compact wide-band antenna.

(a) At 2 GHz　　(b) At 3 GHz　　(c) At 4 GHz　　(d) At 5.6 GHz

(e) At 8.3 GHz　　(f) At 12 GHz　　(g) At 13.8 GHz
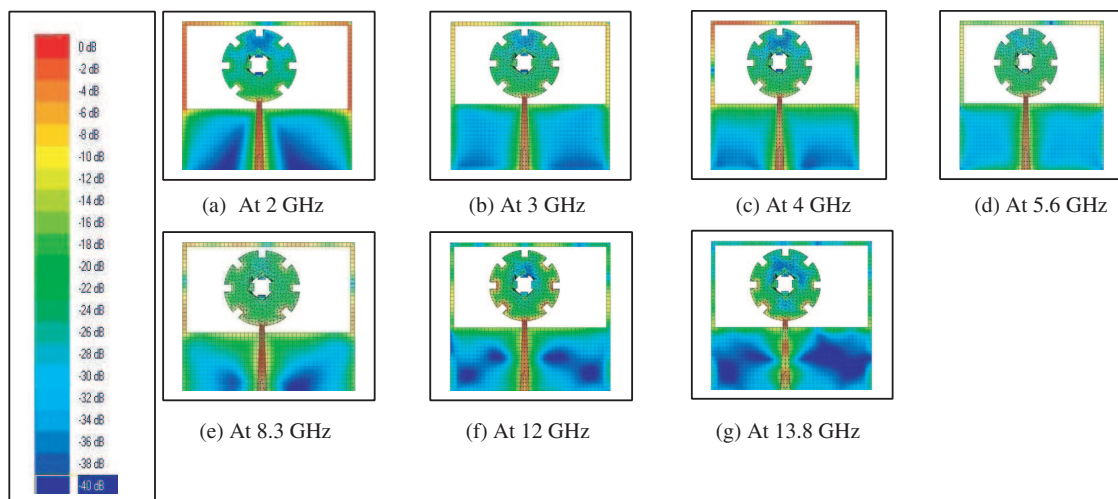
Figure 10: Simulated current distributions of the gear antenna.

The current distribution analyses are performed in order to provide additional information for the performance of the antenna. The placements of the taper have been decided considering the current distributions so that good radiation can occur at the new resonant frequencies. The current distribution shows a different pattern for each of the seven frequencies. i.e., the first harmonic around 2 GHz in Fig. 10(a), the second harmonic around 3 GHz in Fig. 10(b), the third harmonic around 4 GHz in Fig. 10(c), the fourth harmonic around 5.6 GHz in Fig. 10(d), the fifth harmonic

around 8.3 GHz in Fig. 10(e), the sixth harmonic around 12 GHz in Fig. 10(f), and the seventh harmonic around 13.8 GHz in Fig. 10(g). This confirms that the gear slot is capable for supporting wide-band resonant modes.

The radiation patterns are shown for the proposed antenna in polar form with a dB strength scale. Simulated results of the radiation patterns of the proposed antenna are presented in Fig. 11. The results include the E-plane and H-plane at different frequencies 2, 3, 4, 5.6, 8.3, 12, and 13.8 respectively.

Gain versus frequency plot is shown in Fig. 12, in the frequency range (2–17.7) GHz. One should not expect an excellent gain bandwidth for this design as the gain maximization in the center frequency has been the goal function for this optimized antenna. The gain variations are between 2.9 dBi and 4.5 dBi. The antenna has a slightly higher gain throughout the entire band, but it is lower at the edges of the band.

The photo of fabricated antenna with SMA connector is shown in Fig. 13.

For the return loss measurement, an edge-mount SMA connector was soldered to the edge of the board, and the return loss was measured on a network analyzer (Agilent 8719ES). The simulated

(a) At 2 GHz (b) At 3 GHz (c) At 4 GHz (d) At 5.6 GHz

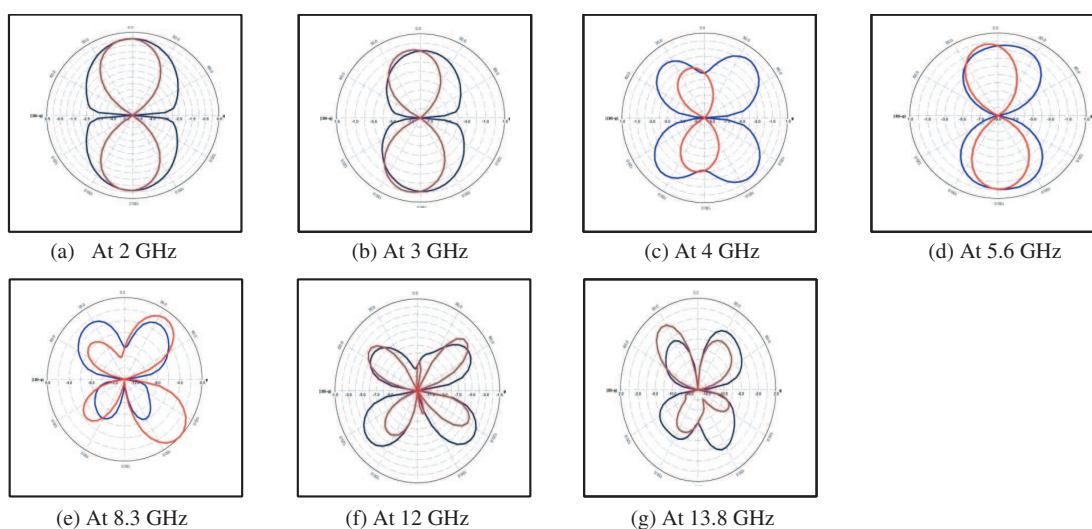(e) At 8.3 GHz (f) At 12 GHz (g) At 13.8 GHz

Figure 11: Simulated radiation patterns at different resonant frequencies for GMA.
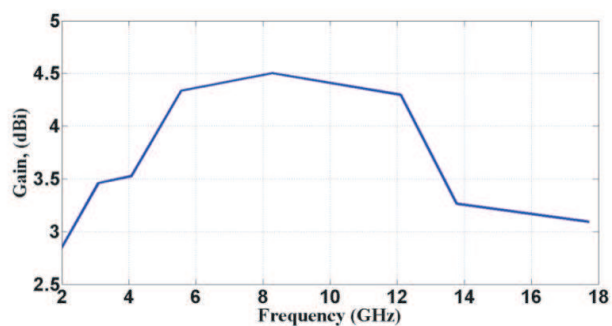

Figure 12: Simulated antenna gain in a frequency range (2–17.7) GHz.
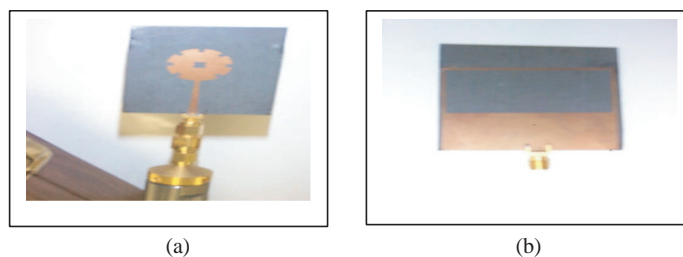

(a) (b)

Figure 13: Photo of fabricated antenna with SMA connector (a) from front view, (b) from back view.

and measured return loss versus frequency is presented in Fig. 14. There is a shift in the level of return loss between the measured and simulated about 10 dB. This is due to the SMA connector is not taken into consideration in the simulation and on the other side the measurement is done without using anechoic chamber but with using a man made chamber with some plates of rubber to reduce reflections. The difference between the two results can possibly be attributed to a small tolerance in the substrate RT5880 and over etching of the metal. The bandwidths of the measured prototype are 17.6%, 51.6%, 30.4%, 9.9%, 37.66%, 9.5%, 9.7%, and 9.6% while for the simulated design are 163.6% and 26%. The broadened bandwidth of the measured design is suitable for wideband applications.
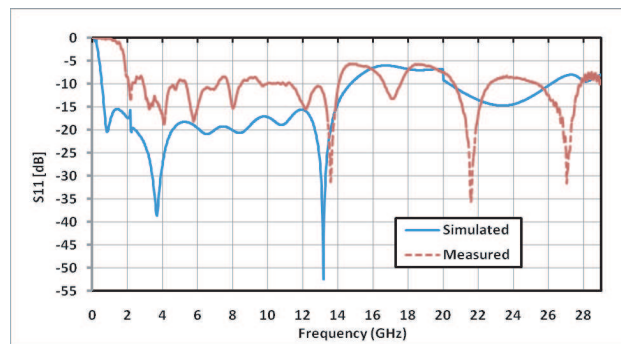

Figure 14: Simulated and measured return loss.

4. CONCLUSIONS

In this paper, a multi-bandwidth gear microstrip antenna fed by a taper is presented. The design is tested by Zeland IE3D and HFSS software and then measured. This design demonstrates a multi-wide impedance bandwidth. These are greatly due to using the appropriate substrate (RT5880), a slot with suitable dimensions, and a ground plane with an adjusted length. The proposed antenna is printed on a rectangle substrate which is processed easily. The multiple bands of the antenna and its suitable gain enable it for a wideband of applications for the wireless communication systems.

REFERENCES

1. Kharakhili, F. G., M. Fardis, G. Dadashzadeh, A. Ahmadi, and N. Hojjat, "Circular slot with a novel circular microstrip open ended microstrip feed for UWB applications," *Progress In Electromagnetics Research*, PIER 68, 161–167, 2007.
2. Jang, Y. W., "Characteristics of a large bandwidth rectangular microstrip fed inserted tri-angular patch in a circular slot antenna," *Microwave Journal*, Vol. 45, No. 5, 288–298, May 2002.
3. Habib, M. A., T. A. Denidni, and G. Y. Delisle, "Design of a new wide-band CPW-fed a circular slot antenna," *Conference IEEE*, 565–568, 2005.
4. Abuhalima, S. S. S., E. A. F. Abdallah, and D. A. E. Mohamed, "Ultrawideband elliptical microstrip antenna using different taper lines for feeding," *Proceedings of the 11th WSEAS International Conference on Communications*, 144–149, Agios Nikolaos, Crete Island, Greece, July 26–28, 2007.

Design and Analysis of Coplanar-waveguide-fed Dual-band Antenna by FDTD

Hou Zhang and Jian Wang

Airforce Engineering University, Xi'an 710021, China

Abstract— In this paper, an improved FDTD adopting mode match method is employed to design and analyze a CPW-fed dual-band antenna. Compared to the traditional FDTD, the improved FDTD can decrease both memory requirements and computational time. Finally, a good agreement between calculated and measured results is presented by comparison.

1. INTRODUCTION

With the rapid development of MMIC and the presence of low-loss dielectric material, coplanar-Waveguide has found many applications and attracted many researcher's attention on its advantages such as low dispersion and coplanar characteristics [1, 2]. However, a large number of literatures focused on the new structure of the device [2–7], there were fewer reports concerned with its numeric calculations [8]. In this paper, a dual-band antenna fed with CPW structure is analyzed and designed by using FDTD. The traditional FDTD is improved by introducing a mode-matching set of excitation source to decrease both memory requirements and computational time. The calculated results are in good agreement with the measured ones.

2. STRUCTURE OF DUAL-BAND ANTENNA

The antenna considered in the paper is shown in Figure 1. It is printed on a FR4 board with thickness of 1.6 mm and dielectric constant of 4.4. The area of the ground plane is $88 \times 75 \, \text{mm}^2$. A printed rectangular slot of $L \times W$ is loaded with an open metal ring with a strip width of d.
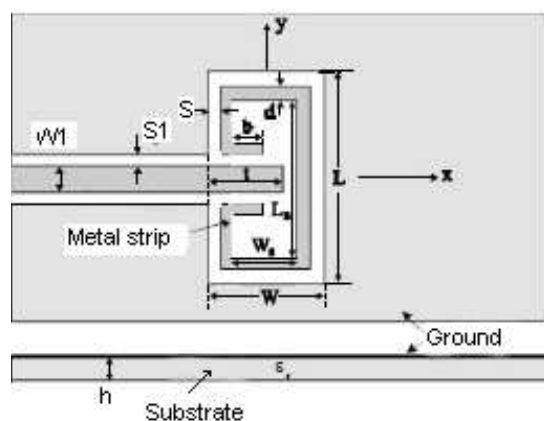
Figure 1: Geometry of the dual-band slot antenna fed by CPW.

The antenna can be divided into two slot antennas coupled with each other. The first antenna consists of a rectangular slot with dimension of $L_s \times W_s$, and the second one consists of a rectangular loop slot whose width is S. The two slot antennas are fed in capacitive way by CPW, the width of the CPW signal strip is denoted with $W1$, and $S1$ represents the gap between the ground and the CPW signal strip. A segment length of t of the CPW signal strip and a segment length of b of the metal strip can be adjusted to change the resonant mode and the impedance of the antenna. Thus two equivalent magnetic current loop circuits are established to work at two different operating frequency band.

By choosing the appropriate lengths of b and t the antenna resonance of two modes can be inspired and a good impedance matching can be achieved. The first resonant frequency is decided mainly by the rectangular perimeter of the loop antenna $(2L + 2W)$, and the second resonant frequency is largely dependent on L_s, the length of the rectangular slot antenna.

3. FDTD ALGORITHM

FDTD (Finite-Difference Time-domain Method, referred to as FDTD) is a numerical algorithm to solve electromagnetic field problems. The electromagnetic fields components are expressed with a finite difference format, space grids with the same electromagnetic properties of the antenna is used to simulate the objects, appropriate initial value and absorbing boundary conditions in the calculating space are selected to make simulation of electromagnetic wave transmission in space and interaction between the waves and the object, getting four-dimensional numerical solutions to Maxwell's equation [9].

3.1. Excitation Settings

According to the geometric symmetry of CPW structure, magnetic symmetry boundary conditions is applied to simplify the procedures for calculation, the simulation space can be reduced by half, which can accelerate the speed of calculation. Figure 2 shows the calculation model. The traditional FDTD has a demand of large memory and time-consuming with its usual excitation, this paper uses a new type of excitation approach producing an electric field distribution similar to that produced by CPW structure. This can be done by using electrostatic charge theory as illustrated in Figure 3. The method is known as mode match method.
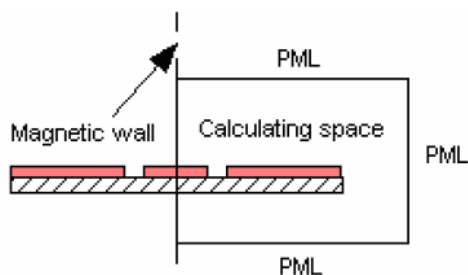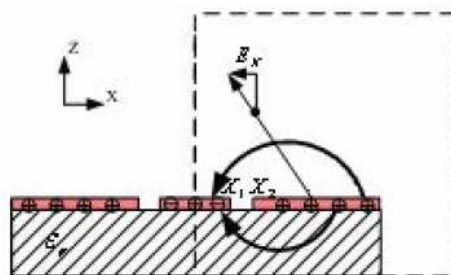

Figure 2: Calculation model.

Figure 3: Field distribution by electrostatic charge theory.

In Figure 3, the conductor is distributed with electric charge, which will produce electric field. Therefore the electric field at every point on X-Z plane can be calculated. This electric field distribution is basically same to the CPW electric field distribution, which can be used as the source of excitation. The calculation formula is as follows:

$$E_x = -\int_{-\infty}^{-X_2} \frac{x_s - x_0}{\left[(x_0 - x_s)^2 + (z_0 - z_s)^2\right]^{3/2}} + \int_{-X_1}^{X_1} \frac{x_s - x_0}{\left[(x_0 - x_s)^2 + (z_0 - z_s)^2\right]^{3/2}}$$
$$- \int_{X_2}^{\infty} \frac{x_s - x_0}{\left[(x_0 - x_s)^2 + (z_0 - z_s)^2\right]^{3/2}} \tag{1}$$

where x_0, z_0 is the coordinates of the field point, and x_s, z_s represent the coordinates of the source point.

3.2. Absorbing Boundary

FDTD algorithm needs Yee's grid to calculate electromagnetic fields in all regions. However the calculation space couldn't be unlimited, so there is a cut-off of the border in order to carry out the antenna radiation simulation. On the other hand, it is necessary to consider the true free-space simulation. A good set of non-reflective boundary conditions, that is, absorbing boundary conditions (Absorbing Boundary Conditions, ABC) was introduced. The commonly used absorbing boundary conditions include Mur absorbing boundary, Dispersion Boundary Condition (DBC) absorbing boundary, Berenger's Perfect matched layer (PML) [6], PML is t selected in the paper.

3.3. Fields Transformation from Near-fields to Far-fields

The fields calculated by FDTD are limited in near-fields region. The far-fields can be achieved by transformation from the calculated near-fields. Scalar Kirchhoff integral equation shows that the

relationships between the fields inside a closed volume and the fields on the surface of the volume are as follows [10]:

$$\phi\left(\vec{r}, t + \frac{R}{c}\right) = \frac{1}{4\pi} \oiint\limits_{s'} \hat{n} \cdot \left[\frac{\nabla'\phi}{R} - \frac{\vec{R}}{R^3}\phi - \frac{\vec{R}}{cR^2}\frac{\partial\phi}{\partial t}\right] ds' \tag{2}$$

where R denotes the distance between the field point and the source point, c is the light velocity, s' represents the surface of the source, ϕ is potential function, and time is expressed with t.

In the grid space, the integral is substituted with the sum, the differential with the central difference, The integral is calculated over a rectangular surface area. Thus the sum equation can be deduced. In this way, the near-field calculated by every step on the surface can be added to its far-field at all times. The transformation can be carried out with time-step iteration of FDTD at the same time.

3.4. FDTD Simulation Model

The antenna is designed by using the improved FDTD stated above, the upper and lower spaces in metal patches are divided into eight units, the step lengths of the grid in three coordinates direction are 0.6 mm, 0.6 mm and 0.5 mm, respectively. Frequency is limited to 1 GHz~3 GHz. The final geometry parameters of the antenna are illustrated in Table 1. Figure 4 and Figure 5 show the measured return loss and calculated return loss respectively. Figure 6 is electric field distribution of the antenna working at 1680 MHz and 2340 MHz. Figure 7 gives the calculated radiation pattern of the antenna at 1680 MHz and 2340 MHz.

Table 1: Geometry parameters of the antenna.

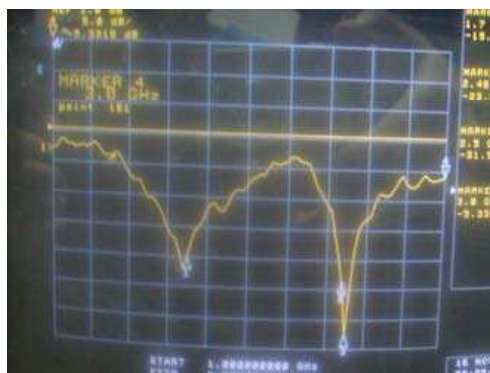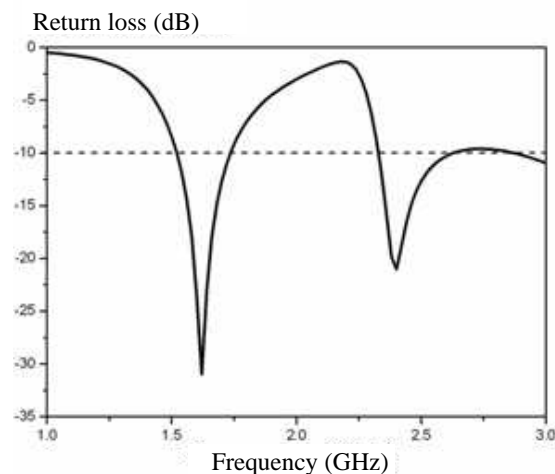| $L \times W$ (mm × mm)	$L_s \times W_s$ (mm × mm)	S (mm)	d (mm)	t (mm)	b (mm)
44.9 × 19.8	34.3 × 9.2	1.9	3.4	12.9	3.5

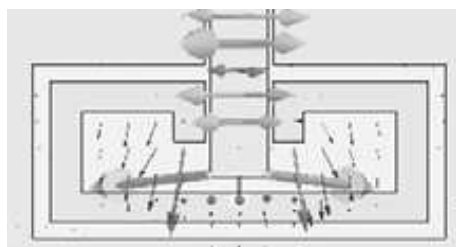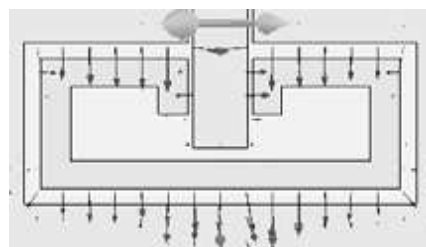Figure 4: Measured return loss.

Figure 5: Calculated return loss.

(a)

(b)

Figure 6: Electric field distribution of the antenna at (a) 1680 MHz and (b) 2340 MHz.
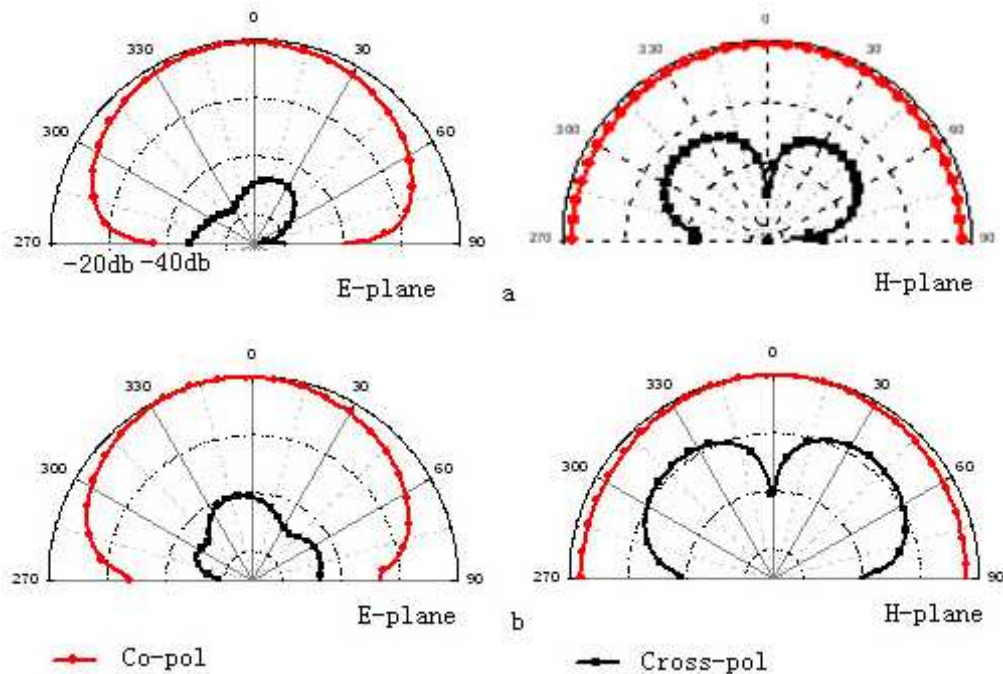
Figure 7: Radiation pattern of the antenna at (a) 1680 MHz and (b) 2340 MHz.

4. CONCLUSION

An improved FDTD adopting mode match method is employed to design and analyze a CPW-fed dual-band antenna. A good agreement between calculated and measured return loss is presented by comparison. Compared to the traditional FDTD, the improved FDTD can decrease both memory requirements and computational time.

REFERENCES

1. Ding, X. and A. F. Jacob, "CPW-fed slot antenna with side radiating apertures [J]," *IEE Proc. — Microw. Antennas Propag.*, Vol. 145, No. 2, 104–108, 1998.
2. Liang, X. L., S. S. Zhong, and W. Wang, "Tapered CPW-fed printed monopole antenna [J]," *Microwave and Optical Technology Letters*, Vol. 48, No. 7, 1242–1244, 2006.
3. Kim, Y. and D. H. Kwon, "CPW-fed planar ultra wideband antenna having a frequency band notch function [J]," *Electronics Letters*, Vol. 41, No. 7, 403–405, 2005.
4. Chen, M. E. and J. H. Wang, "CPW-fed crescent patch antenna for UWB applications [J]," *Electronics Letters*, Vol. 44, No. 10, 505–506, 2008.
5. Sze, J. Y., K. L. Wong, and C. C. Huang, "Coplanar waveguide-fed square slot antenna for broadband circularly polarized radiation [J]," *IEEE Trans. on Antennas and Propagation*, Vol. 51, No. 8, 2141–2144, 2003.
6. Lu, S. W., T. F. Huang, and P. Hsu, "CPW-fed slot loop coupled patch antenna on narrow substrate [J]," *Electronics Letters*, Vol. 35, 2067–2068, 1999.
7. Han, S., "Compact multiband slot-ring antennas [D]," Dissertation of Doctor of Philosophy Degree, Department of Electrical and Computer Engineering in the Graduate School, Southern Illinois University Carbondale, Southern Illinois, Nov. 2004.
8. Chen, W. S., "The search on dual-band antenna and slot antenna [D]," Dissertation of Doctor of Philosophy Degree, Natural Central University, Jungli, Taiwan, 2001.
9. Ge, D. B. and Y. B. Yan, *Electromagnetic Wave Finite-different Time-domain Method [M]*, (in Chinese), Xidian University Press, Xi'an, 2002.
10. Liu, P. G., "The study of UWB signal radiation and scattering [D]," (in Chinese), Dissertation of Doctor of Philosophy Degree, National University of Defense Technology, Changsha, 2001.

Design of an Optimum Antenna System for Maximum Power Transfer Using Statistical Design of Experiment Approach

Arnab Roy, **Sushanta Paul**, **A. S. M. Shamsuzzaman**, **and Md Ishfaqur Raza**

ECE Department, East West University, Mohakhali, Dhaka-1212, Bangladesh

Abstract— For optimum design of the feed of an antenna, the location of the feed with respect to the antenna is critical. The location and design of the feed is a function of the antenna target frequency, material characteristics of the substrate, and the physical measures of the antenna. The objective is to match the impedance of the feed to the impedance at the feed point. In this paper the design of an antenna system is implemented using the design of experiment statistical approach. A micro strip patch antenna is used as a test case to illustrate the design methodology. All the parameters that define the antenna and the feed are taken into consideration in the experiment to ensure maximum power transfer from the feed to the antenna and minimize return loss. A screening experiment is run to characterize the parameters that affect the maximum power transfer. Later a custom design is presented using the screened transmission system input variables to enable maximize power transfer. The result of this process creates a simple empirical function to design an efficient antenna.

1. INTRODUCTION

The design of a microwave system is instrumented for maximum power transfer from source to load using optimum impedance matching techniques [1]. Given that different components of the system are characterized by a range of physical and material parameters, a detailed analysis of the impact of the different parameters on the system performance is essential. The oversight of the impact of any parameter, particularly not accounting for sensitivity of system performance to variations in design definitions can render a system useless or at the least make the design inefficient. Mathematical and simulation tools do exist that allow for the modeling and simulation of point designs. However, this paper intends to highlight approaches which can be used to look not at point solutions or the performance variations over a range of values, but analyze impact of different parameters to system performance criteria.

In this paper, we have selected to study a micro strip patch antenna, which is widely used in communications and microwave applications. We will look at the impact of several critical design parameters, such as length and width of the patch, the dielectric thickness and its permittivity constant, the antenna feed impedance, width, and its locations. The performance metric is limited to the resonant frequency, which is the optimum frequency of operation for the antenna and return loss S11 (dB), which characterizes the matching of the feed impedance to the antenna. The impact of a low efficient product also has a chain effect on the overall performance of the system, end to end. Our approach here uses a statistical approach, commonly known as the Design of Experiment (DOE). Instead of writing our own tool to perform the analysis we have selected to use the popular statistical tool 'JMP' [2].

2. STATISTICAL DOE DESIGN

Design of Experiment (DOE) is a structured, organized method that is used to determine the relationship between the different factors affecting a process and the output of that process [3]. Design of Experiment involves designing a set of experiments, in which all relevant factors are varied systematically. When the results of these experiments are analyzed, they help to identify optimal conditions. To find out the parameters which affect the power transfer of the antenna, we use two designs. They are screening and custom design. In screening design, we examine many factors to see which have the greatest effect on the results of a process. Take for instance the patch antenna — the roughness of the copper used to manufacture the patch on the dielectric. The roughness of the copper of the patch is very likely to be different between antennas. There is a large list of parameters which are likely to be different between separate batches of antennas. However, a complete analysis considering all the parameters is not practical. The screening experiments are designed to screen the parameters and determine which parameter will be included in the comprehensive custom design phase. This phase of the design also identifies the appropriate levels or values that will be included in the custom analysis. The dependence of the system performance

metric (such as S11 and resonant frequency) on design parameters may be linear, quadratic, or even higher order. Screening designs are also used to estimate the order of dependence of these parameters.

Using all this information, a custom analysis is done to determine a prediction formula which will be a close fit to the actual system. The data to create this prediction formula is generated using data from numerous runs of experiments. With the wide range of physical parameters, the range of values these parameters can have, and the nature of dependence of the performance parameters on the design factors, the number of experiments needed to be run becomes unrealistically large. In each of the experiments, the values of the design parameters are generally different. Be it by simulation or by manufacturing, the design or construct of a different antenna for every single experiment is not feasible. The DOE statistical approach reduces the number of experiments to an optimal number which is then executed. The result of the custom design is then used to obtain a understanding of the impact of different design parameters on the system performance metrics.

3. ANTENNA SYSTEM-MICROSTRIP PATCH

A patch antenna is a popular antenna type, which gains its name from the fact that it basically consists of a metal patch suspended over a ground plane. A even height dielectric is sandwiched between the patch and the ground plane. Patch antennas are easy to design and manufacture. The physics and engineering behind the patch is very similar to the microstrip transmission line and resonator. The methods used to build the printed circuit board are widely used in building patch antenas. It should therefore, be noted that the final print-out of the manufacture house may have problems, such as registration (the patch is not exactly where it is supposed to be), the width and location of the feed is not perfect, and the dielectric thickness is not what was desired. In some cases the dielectric constant can also be different due to improper dielectric build. All these issues raises the need to have a good understanding of the impact of different parameters on the antenna performance.
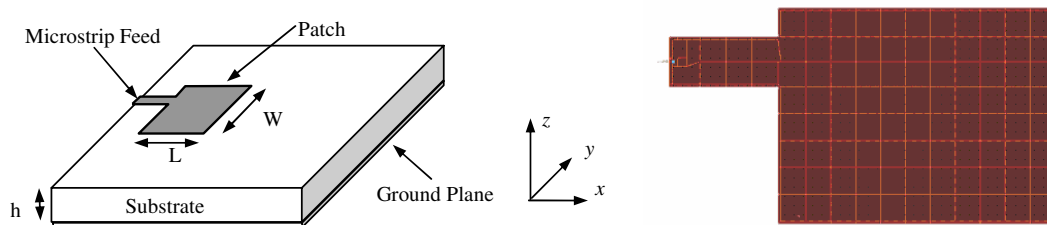
Figure 1: Microstrip Patch Antenna, schematic-left and ADS model-right.

The most popular microstrip antenna is certainly the rectangular patch. Small impedance bandwidth is a common property of patch antennas, which is about 2 to 3%, which is the result of the high quality factor of patches. The resonant frequency of the patch antenna is given by Equation (1), where L and W are the width of the patch (see Figure 1) and eff is the effective dielectric constant of the patch substrate [4].

$$f_o = \frac{c}{2\sqrt{\varepsilon_{eff}}} \sqrt{\left[\left(\frac{m}{L}\right)^2 + \left(\frac{m}{W}\right)^2\right]} \tag{1}$$

The patch antenna is widely modeled as a transission line with the feed providing the input and the radiating edge (on the length side) acting as the load. The input admittance of the patch antenna using this model is

$$Y_{in} = Y_{in1} + Y_{in2} = Y_o \left[\frac{Y_e + jY_o \tan(\beta_p f)}{Y_o + jY_e \tan(\beta_p f)} + \frac{Y_e + jY_o \tan(\beta_p(W - f))}{Y_o + jY_e \tan(\beta_p(W - f))}\right], \text{ where } \beta_p = k_o\sqrt{\varepsilon_{eff}} \tag{2}$$

Here Y_e, Y_o, and f are the antenna radiating edge admittance, antenna patch admittance in the direction of the width, and the feed location on the patch (Figure 2).

Figure 2: Patch Antenna feed on non-radiating edge.

4. EXPERIMENTAL SETUP

For this paper we have used Agilent's Advanced Design System's (ADS) Momentum tool to model and evaluate patch antennas. Momentum is a 3-D planar electromagnetic (EM) simulator [5]. We modeled the parameters of this patch antenna so that its resonance frequency becomes around 7 GHz. The other metric of interest in this design is the return loss (S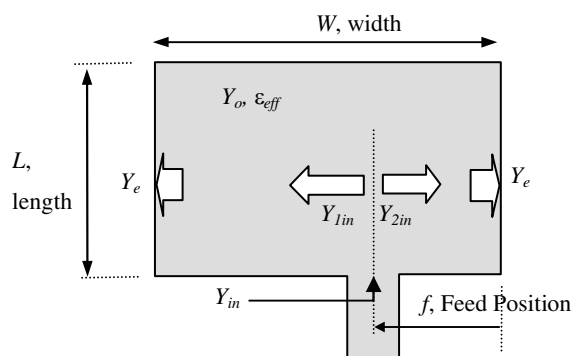11). Several screening experiments were done to identify the critical parameters and their range of interest. These parameters are Length of the antenna, Width of the antenna, Feed line position, Feed line width, Dielectric constant and Dielectric thickness. No particular dielectric material was selected for this analysis. Our DOE target is to match desired resonant frequency and maximize |S11|.

The table below shows the range of value selected for our simulation experiments.

Table 1: Range of DOE parameters.

	Length (mm)	Width (mm)	Feed Position (mm)	Feed Width (mm)	Dielectric Constant	Dielectric Thickness (mm)
Start Value	16	12.5	1.5	2	1.75	0.75
End Value	17.5	13.5	2.5	3	2.5	0.85

5. DATA ANALYSIS

ADS momentum is used to perform the simulations. From each experiment the resonant frequency was noted and the return loss recorded. For this range of design parameters, a total of 32 experiments were designed. The first 5 experiments are listed in Figure 3, which lists the values that are selected by the tool for simulation.

Table 2: Data sheet for the first 5 experiments.

	Length	Width	Feed Position	Feed Width	Dielectric Constant	Dielectric Thickness	Resonance Frequency	S11 dB
1	17.5	12.5	2.25	2.5	1.75	0.85	8.45	22
2	17.5	13.5	1.75	2	1.75	0.85	7.855	42
3	16	12.75	1.5	2.25	2.5	0.75	7.18	8
4	16	13.5	2.5	2	2.5	0.8	6.67	12.25
5	16.75	13.5	1.5	2.75	2.125	0.75	8.15	2.1

Depending on the sensitivity of the design to the design parameter, either the resonant frequency or the return loss would shift. Looking at Equation (1), the resonant frequency would be ore sensitive to dielectric constant and width of the patch. However, from Equation (2), the return loss is sensitive to most of the design parameters selected in these design analysis. This was also observed from the screening experiments. This is demonstrated in the following set of experiments. For a change in the dielectric constant, the change in resonance frequency is noted in Figure 4(a),

shifting between ∼ 7.8 GHz to 8.4 GHz. On the other hand a change in the feed width, which changes the impedance of the feed line, resulting in an impedance mismatch, and significant shift in S11 return loss, Figure 4(b).
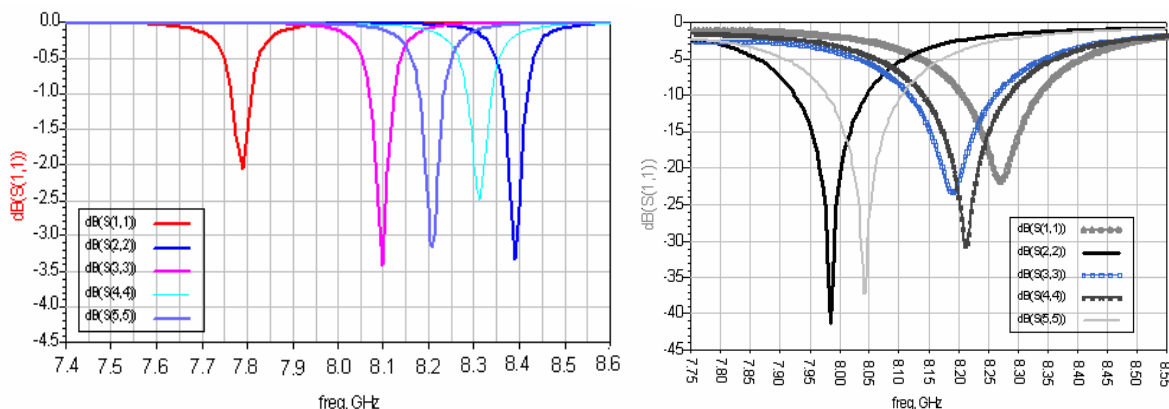
Figure 3: ADS Simulation of frequency shifting and S11 dB (a) left, (b) right.
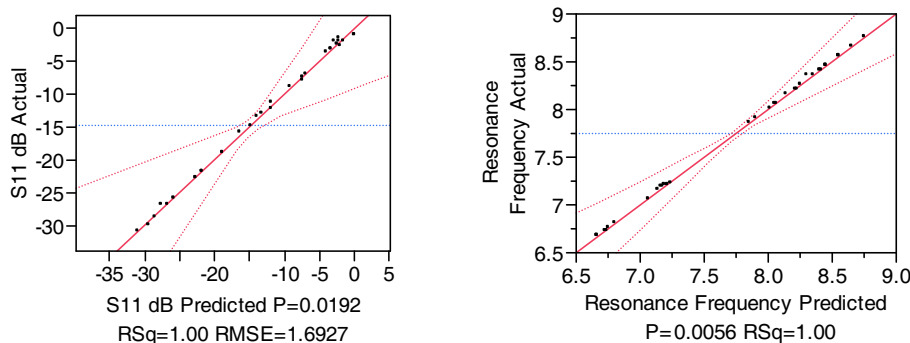
Figure 4: Actual vs. predicted plot for resonant frequency and S11 dB.

The result of the design analysis generates a formula designed for a least square error fit to the data generated by the experiments. This equation is used to predict the resonant frequency and return loss of the patch antenna. The plot of the predicted vs. actual S11 and Resonant frequency, shown in Figure 4, shows the RMSE (root mean square error) to be relatively small, indicating a good match between the prediction formulae for these output parameters. RSq (root square error) is found to be 0.999 for resonance frequency and 0.998 for S11, which means that the predicted data based on a prediction formula is very close to real. All this leads to a reliable model being setup, which can be used to design other antennas with the target to achieve desired performances.
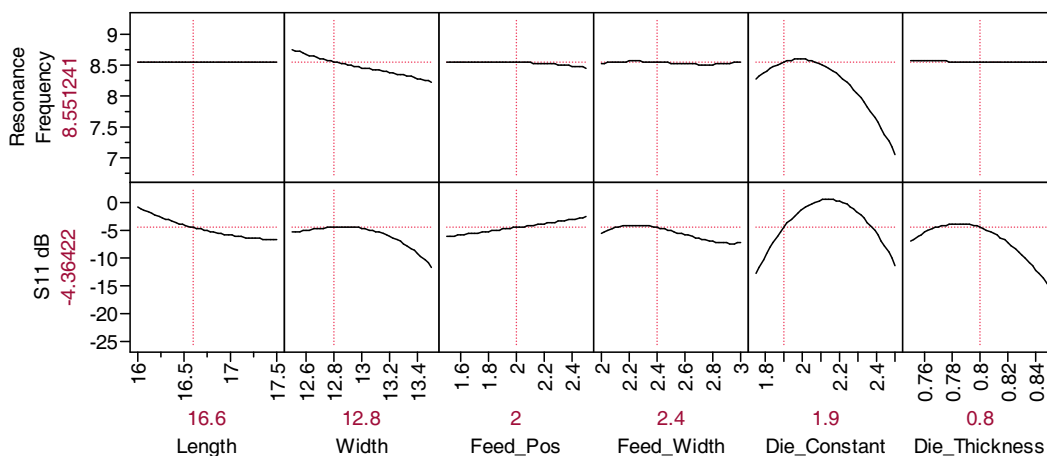
Figure 5: Prediction profiler.

Based on the prediction formulas, we are able to characterize the impact of each of the select design parameters on the performance of the antenna. The plot in Figure 5 is very intuitive as it indicates findings that reflect the theoretical basis of a patch antenna design. In the case of fundamental mode, TM_{01}, (see Equation (1)) where the feed is in the non-radiating edge (the length), $m = 0$, and $n = 1$, the resonant frequency is particularly sensitive to Dielectric constant and patch width. Return loss (S11) on the other hand is very much dependent on most of the parameters, as demonstrated in Figure 5.

6. CONCLUSIONS

This paper demonstrates the use of DOE statistical analysis method to do a comprehensive analysis of a micro strip patch antenna. It confirmed the theoretical basis of the antenna using numerical simulation data. The derived prediction formula can now be used to design antennas that can take into consideration variations in the design parameters and predict the performance of the high speed design system. The scope of such analysis can be broadened to include parameters that contribute to the performance of the system such as surface roughness.

REFERENCES

1. Pozar, D. M., *Microwave Engineering*, 3rd Ed., Wiley, New York, 2005.
2. http://www.jp.com
3. Staiculescu, D., N. Bushyager, and M. Tentzeris, "Microwave/milimeterwave metamaterial development using the design of experiments technique," *ACES*, 2005.
4. Balanis, C. A., *Antenna Theory, Analysis and Design*, 2nd ed., 740, Wiley, New York, 1996.
5. http://eesof.tm.agilent.com/products/momentum_main.html

Renormalization Group Application to Multi-port Model for Studying Fractal-shaped Structures' Diffraction

T. BenSalah, **C. L. Aguili**, **and T. Aguili**

L. Syscom, ENIT, B. P. 37, Le Belvedere 1002, Tunis, Tunisia

Abstract— This paper reports on the application of the renormalization group theory to the study of a fractal shaped structure's diffraction using the multi-port model proposed in the MS-MGEC, a MoM multi scale extension. Actually, the Surface Impedance Operator exposed in the latter method enables the formulation of expressions that lead to rigorous calculation of the stationary state of the studied system when the fractal iteration tends toward the infinity in an elegant manner. This corresponds to a multi-dimensional fixed point characterizing the invariance by renormalization of the structure regardless to fractal iteration.

1. INTRODUCTION

One of the most annoying characteristics of fractal [1] structures is their infinite number of scales. Existing Multi-scale methods ([2–4] among others) lack rigorous formulation for infinite scales. At their very best they can rely on convergence mechanisms when the fractal iteration tends toward infinity.

In a previous work we introduced an extension of the surface impedance model (MS-MCEG method [5, 6]) that we used for studying diffraction on some multi-scale obstacles pre-fractal shaped. We have shown that this model presents interesting results in terms of precision and also time of execution. According to its definition, it is well adapted to fractals. Indeed, a recursive application of the method describes well the iterative building of such structures. However, as we have stated, to study fractals at infinite iterations we need tend toward infinity and use convergence mechanisms.

To get rigorous modeling of the problem we propose the application of the renormalization group theory [7, 8] hard to circumvent in asymptotic phenomena studies. The method was firstly introduced in theoretical physics to help studying the behavior of some ferromagnetic materials around and at the temperature of Curie, where magnetic susceptibility presents a narrow peak and has an infinite value. We have already applied the renormalization group theory to the classical surface impedance model [9]. Here we are about to apply it to the Surface impedance operator.

The structure we are interested in (Figure 1) is the well known Cantor Set used as a thin surface obstacle in a rectangular waveguide. The waveguide is supposed with infinite length and with periodic boundaries.
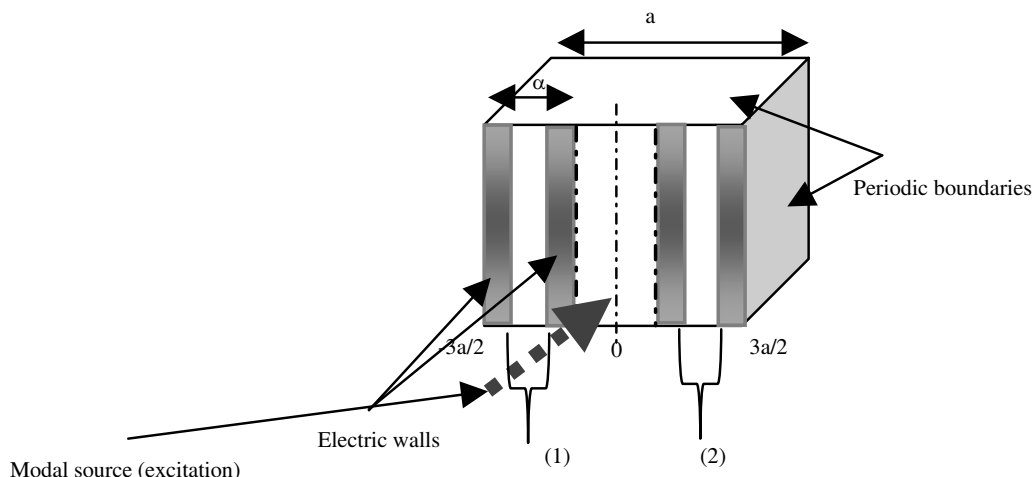

Figure 1: Cantor set obstacle in rectangular waveguide ($k = 2$).

2. MULTI SCALE METHOD OF GENERALIZED EQUIVALENT CIRCUIT

The MS-MCEG method is based on the divide and conquer paradigm. It makes possible the study of parts (which are supposed less complex) of a given structure as standalone structures. This is done regardless of coupling problems. It provides also formalization to help the characterization of the whole (complex) structure that becomes simpler as we come back in every iteration to the fractal at its very first iteration ($k = 1$).
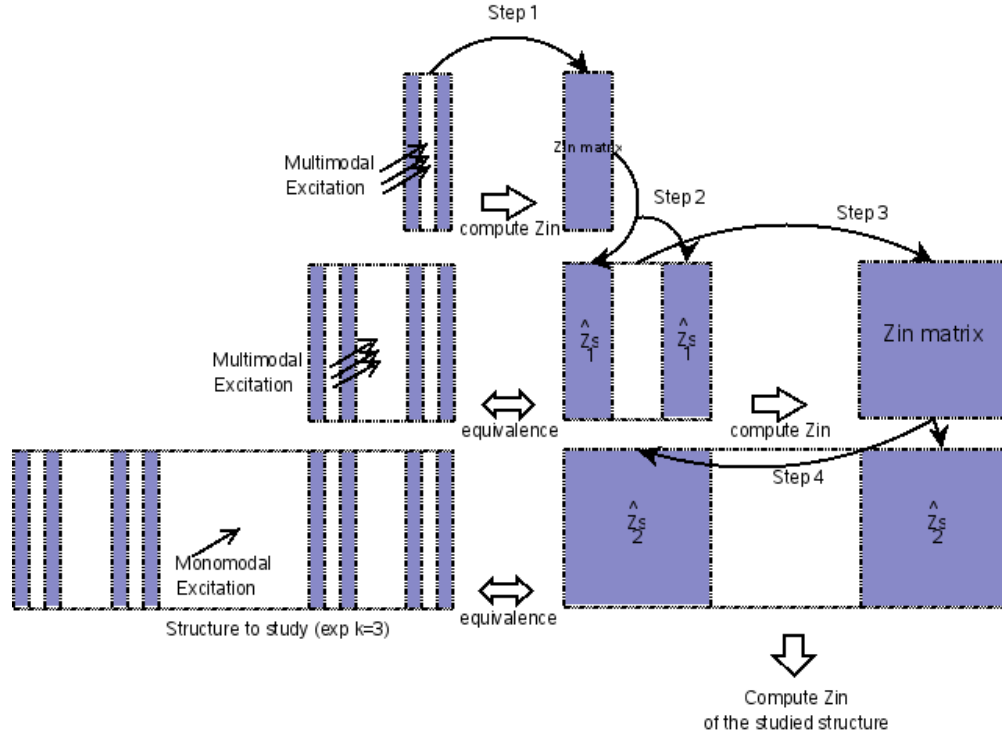
Figure 2: Application of the Z_s operator model for $k = 3$.

Indeed, as described in the Figure 2, when studying the Cantor Set as an obstacle (at iteration $k = 3$ for instance) with a mono-mode (fundamental mode) excitation to retrieve its input impedance Z_{in}, we start (step 1) studying the fractal at the first iteration in a multimodal excited (periodic) guide. The corresponding equivalent circuit is given in Figure 3. This type of excitation permits taking into consideration coupling between close metallic surfaces and thus improving precision. The resulting Z_{in} matrix is used as a Surface impedance operator in the next iteration (step 2) and leads to the equivalent circuit of Figure 4. The application of this latter equivalent circuit (step 3) is performed on the first iteration (scale $k = 1$) of the fractal (step 3) but with introduction of the Z_s operator model. The process is repeated until the maximum iteration is reached (here step 3 and 4 to get an equivalent structure to the one studied, i.e., Cantor Set at $k = 3$).

The $[Z_{in,k}]$ matrix expression for each iteration (for any k) using Galerkin method can be expressed as following:

$$[Z_{in,k}] = \left[\widehat{Z}_{s_k}\right] = \left([B_k]^t \left(\left[\hat{Z}_k\right] + \left[\hat{Z}_{s_{k-1}}\right]\right)^{-1} [B_k]\right)^{-1} \tag{1}$$

where

$$[B] = [\langle g_p, F_n \rangle]_{0 \leq p \leq NE \wedge 0 \leq n < m} \quad \text{and} \quad \hat{Z} = \sum |F_n\rangle z_n \langle F_n|$$

$\{g_p\}$ are the (current) test functions, $\widehat{Z}$ is the Impedance operator, $\{z_n\}$ are mode impedances and $\{F_n\}$ the (big) guide mode functions.

Now that we established an expression linking successive iteration impedances, we are about to apply renormalization group theory to find out the input impedance of the structure for an infinite value of k.
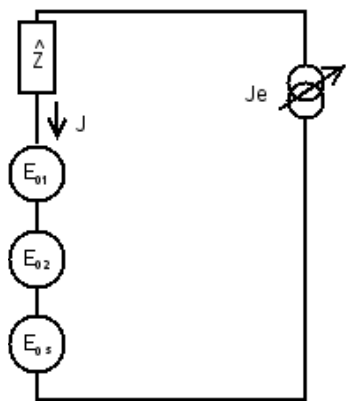
Figure 3: Equivalent circuit at step 1 ($k = 1$).
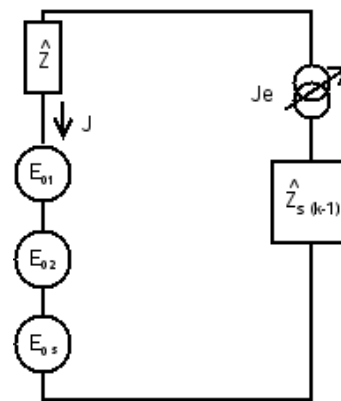
Figure 4: Equivalent circuit at step 3.

3. RENORMALIZATION GROUP

Renormalization Group relies on four steps: decimation, scaling, transformation and verification of invariance.

Decimation leads elementary components to be summed to produce groups of bigger and simpler components. This is also known as the constructive step or forward zoom step. For fractal structures, this often refers to the auto similarity characteristics of such structures. Indeed, a complex fractal is built upon a set of the same structure. In our case, a Cantor Set at iteration k is built of two structures at iteration $k - 1$ (Figure 5).
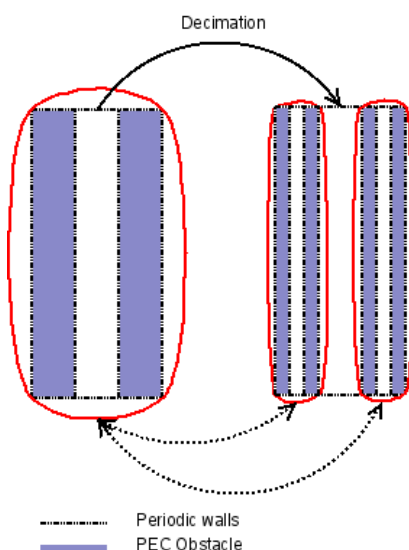
Figure 5: Decimation.
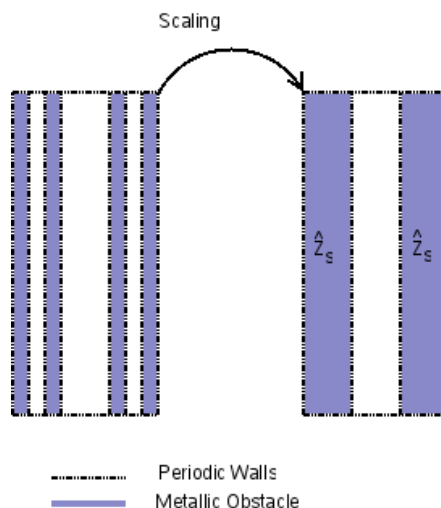
Figure 6: Scaling.

Scaling step consists in the opposite operation: backward zoom. This is a destructive step leading to the simplification of a complex structure: a length scale transformation according to initial form. This is made possible thanks to the surface impedance operator (multiport model) proposed in the MS-MCEG. The Cantor Set structure at iteration $k = 2$ is equivalent to the same structure at iteration $k = 1$ when changing parts (PEC and Dielectric) with a uniform Z_s area so that we fall back into the simplest structure ($k = 1$).

It is important to mention that decimation and scaling, when applied together, lead to a very interesting property: as far as we go deep in iterations, the studied structure remains as complex as the very first iteration one : the generator structure ($k = 1$).

The most difficult step in renormalization is the transformation. It consists in replacement of real parameters with effective ones. Effective parameters are very domain dependent.

First we rewrite the Equation (1) to inline backward recursion needed in renormalization;

$$[Z_{s_{k'}}] = \left([B_{k'}]^t \left(\left[\hat{Z}_{k'} \right] + \left[\hat{Z}_{s_{k'+1}} \right] \right)^{-1} [B_{k'}] \right)^{-1} \tag{2}$$

This is correct if we consider $k' = k_{\max} - k$. $k_{\max}$ is the final (infinite) iteration with a particular dimension (width) that we need to characterize (calculate Z_{in}). $Z_{s_{k'}}$ corresponds to a dimension of $a\alpha^k$. From now on we will use k instead of k' just for simplifying writings.

When applying the Equation (2), we choose mode functions of TE type (due to symmetry of the structure) and cosine typed current test functions. We can demonstrate that, when choosing normalized test functions (mode functions are always normalized) all scalar products (needed by Galerkin method) are scale independent (k independent). This is important because B matrix becomes independent of k.

Furthermore, in TE mode, impedance mode is given by:

$$z_n^k = \frac{j\omega\mu}{2\sqrt{\frac{4n^2\pi^2}{\alpha^{2k}a^k} - k_0^2}}$$

a is the width of the guide, α is the fractal quotient (Figure 1), n the mode index and k the fractal scale.

When choosing values of k_0^a that verify $\frac{\alpha^k k_0^a}{2n\pi} \ll 1$ (small dimensions relatively to wavelength) we can write $z_n^k = \alpha^k \frac{j\omega\mu a}{4n\pi}$.

Using this approximation, we find out very interesting results. From one hand, the impedance operator $\left[\hat{Z}_k \right]$ may be rewritten as follows:

$$\left[\hat{Z}_k \right] = \left[\sum_{s<n\leq N} \langle g_p, F_n \rangle z_n \langle F_n, g_q \rangle \right]_{p,q} = \alpha^k \left[z'_{p,q} \right]_{p,q}$$

where $z'_{p,q}$ is scale independent. N is the maximum number of modes and s is the number of excitation modes.

From the other hand, $\left[\hat{Z}_{s_k} \right]$ may also be rewritten in the next form :

$$\left[\hat{Z}_{s_k} \right] = \left[\sum_{m\leq s} \sum_{n\leq s} \langle g_p, f_m \rangle z_{s_{mn}} \langle f_n, g_q \rangle \right]_{p,q} = \alpha^k \left[\sum_{m\leq s} \sum_{n\leq s} \langle g_p, f_m \rangle \frac{z_{s_{mn}}}{\alpha^k} \langle f_n, g_q \rangle \right]_{p,q} = \alpha^k \left[z'_{sp,q} \right]_{p,q}$$

where $z'_{sp,q}$ is scale independent. $\{f_{mn}\}$ are the mode functions of the small guide.

This rewriting leads to an expression of the form $\left[\frac{\hat{Z}_{s_k}}{\alpha^k} \right] = \left([B]^t \left(\left[\hat{Z} \right] + \left[\frac{\hat{Z}_{s_{k+1}}}{\alpha^{k+1}} \right] \right)^{-1} [B] \right)^{-1}$.

We insist on the fact that B and $\hat{Z}$ are k independent.

At this point we can introduce the formal parameter $\hat{Z}'_{s_k} = \frac{\hat{Z}_{s_k}}{\alpha^k}$ of the transformation. It leads to an expression independent of k of the form

$$\hat{Z}'_{s_k} = \Psi \left(\hat{Z}_{s_{k+1}} \right) \tag{3}$$

This is the relation of renormalization that defines the renormalized problem.

The last step consists in finding the fixed points that describe the property of invariance of the system.

Numerical calculation ($s = 2$, $f = 4\,\text{GHz}$, $\alpha = \frac{1}{3}$, $a \cdot k_0 = 0.1$, $N = 5000$, $P = 30$) leads us to confirm that for any initial value $\hat{Z}'_{s_0}$, the numerical sequence Equation (3) converges to a single multi-dimensional (matrix) fixed point that defines the invariance of the system for infinite values of k.
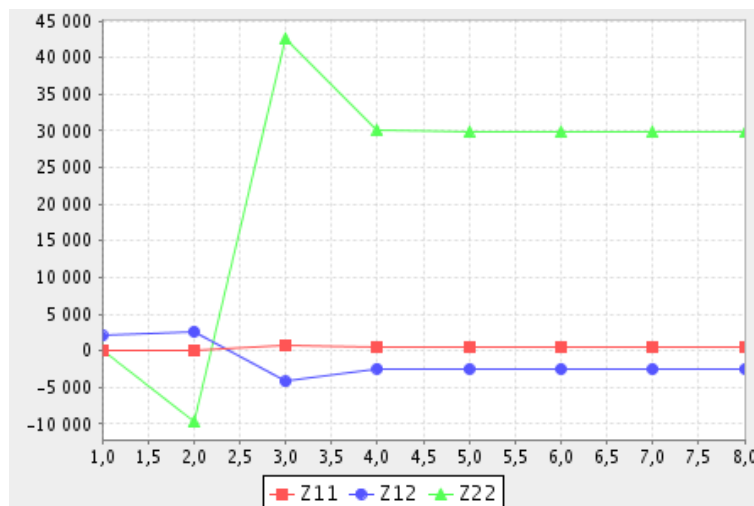
Figure 7: Multi dimensional fixed point for a couple of excitation modes.

4. CONCLUSION

In this paper we proposed the application of the renormalization group theory to the MS-MCEG method to help characterizing diffraction on a fractal shaped structure used as a plane obstacle in a rectangular waveguide. We have shown that renormalization is well adapted to fractal structures but its application is domain and structure specific particularly for finding the formal parameter. This transformation is important to define the relation of renormalization that leads to the calculation of a set of fixed points that characterize the invariance of structure toward scale. An application to the Cantor Set structure shows the existence of a single multidimensional (2 excitation modes) fixed point.

REFERENCES

1. Madelbrot, B., "Fractal objects," *Flammarion*, 1975.
2. Canning, F. X., "The impedance matrix localization method (IML) permits solution of large scatterers," *IEEE Magn.*, Vol. 27, 4275–4277, 1991.
3. Michielssen, E. and A. Boag, "A multilevel matrix decomposition algorithm for analyzing scattering from large structures," *IEEE Trans. Antennas Propagat.*, Vol. 44, 1086–1093, 1996.
4. Su, T. K. C., "A multiscale moment method for solving fredholm integral equation of the first kind," *Journal of Electromagnetic Waves and Applications*, Vol. 12, 173–197, 1998.
5. Ben Salah, T. and T. Aguili, "Software implementation of a new multi-scale method for fractal-shaped structures' diffraction analysis," *PIERS Proceedings*, 836–840, Cambridge, USA, July 2–6, 2008
6. Ebn Salah, T. and T. Aguili, "Study of sierpinski carpet fractal structure diffraction using MS-MGEC method," *Microwave Technology and Techniques Workshop*, Noordwijk, The Netherlands, 2008.
7. Wilson, K. G., "The renormalization group and critical phenomena," Nobel Lecture, Oct. 1982
8. Wilson, K. G., "Renormalisation group and critical phenomenon: Renormalization group and Kadanoff Scaling picture," *Phys. Rev. B*, 1971.
9. Aguili, C., T. Ben Salah, T. Aguili, A. Bouallegue, and H. Baudrand, "Study of electromagnetic waves diffraction by bi-dimensional fractal structures using the renormalization method," Elsevier, AUE, 2008.

Input Impedance of Gap-coupled Circular Microstrip Antennas Loaded with Shorting Post

P. Kumar, **Vivek K. Dwivedi**, **G. Singh**, and **S. Bhooshan**

Department of Electronics and Communication Engineering
Jaypee University of Information Technology, Solan 173215, India

Abstract— In this paper, we have presented the simulated results of the input impedance of two gap-coupled circular microstrip patch antennas loaded with shorting post. The input impedance of this proposed antenna is simulated for various gap distances between the adjacent edges of feed patch and parasitic patch. The effect on input impedance with variation of the radius of shorting post is also studied. This simulation is performed by using the commercially available IE3D simulator based on method-of-moments.

1. INTRODUCTION

There has been evergrowing demand for microstrip patch antennas in communication system applications that poses the highly desirable attributes such as multi-band and broadband, compact size, low profile and conformable, printed microstrip antennas, have been widely investigated and are very attractive for their configuration advantages, while bandwidth and efficiency capabilities still not be adequate for some future demands. The bandwidth of the microstrip antenna can be increased using various techniques such as by loading a patch, by using a thicker substrate, by reducing the dielectric constant and by using gap-coupled multi-resonator. The problem with a thicker substrate is generation of spurious radiation and there are some practical problems in decreasing the dielectric constant [1–5]. At lower frequencies, the size of the microstrip antennas is large. One approach to reduce the lateral dimensions of microstrip patch antenna is consists of loading the antenna with one or several shorting posts, that is metallic vias connecting the patch to the ground plane. This technique to reduce the resonance frequency has been proposed first time by Waterhouse [6] and has been demonstrated on a variety of different patch shapes [7]. Microstrip antennas were miniaturized by using shorting post in [8, 9]. As in [6], the maximum reduction in physical size can be achieved if a single shorting post is used. Here the radius of circular patch was reduced by a factor of three, making the antenna size suited for compact communication systems. In [10], circular microstrip antenna with dual frequency operation is designed by shorting the patch. The results are compared with the conventional circular microstrip antenna (without a shorting pin). In this article [10], it is observed that the size of the circular microstrip antenna can be reduced for the same frequency application, if patch is shorted at its edge. It is also observed that the resonant frequency of the pin shorted circular microstrip antenna can be varied by varying the location of shorting pin. In [11], a rectangular microstrip antenna with dual frequency operation
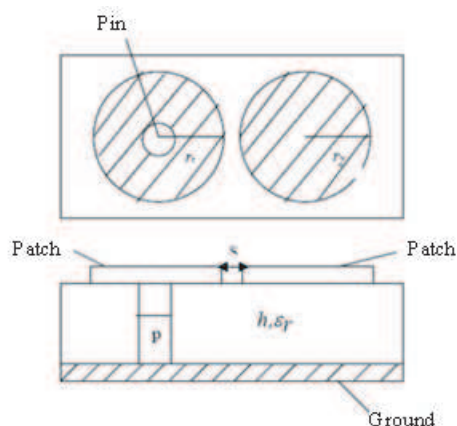
Figure 1: Geometry of gap-coupled circular microstrip patch antennas loaded with shorting post.

is designed by shorting the patch. It is observed that the size of the antenna can be reduced by shorting the patch at its edge. The bandwidth and input impedance are additional properties that are strongly affected by substrate thickness, often in desirable way. The other factors such as dielectric loss and feeding techniques can also be significantly play an important role in the electrical performance of the antennas.

In this paper, the input impedance of the two gap-coupled circular microstrip patch antennas loaded with shorting post is presented. The study is carried out using the method-of-moment based software IE3D. The organization of the paper is as follows. The Section 2 is concerned with the two gap-coupled circular microstrip patch antenna configuration. The Section 3 discusses about the simulated results regarding role of gap distance between the feed patch and parasitic patch as well as radius and position of the shorting post the proposed antenna. Finally, Section 4 concludes the work.

2. ANTENNA CONFIGURATION

The geometry of two gap-coupled circular microstrip antennas is shown in Fig. 1. The patch of radius r_1 is feed patch and patch of radius r_2 is the parasitic patch. The parasitic patch is excited by the gap-coupling whereas feed patch is excited by probe feeding technique. The fed patch is shorted by a pin of diameter p as shown in Fig. 1. The height and permittivity of the substrate is h and ε_r, respectively. The gap distance between the adjacent edges of the feed patch and parasitic patch is s.

Figure 2: Input impedance of the gap-coupled circular microstrip antennas loadedc with shorting post, (a) gap between adjacent edges = 0.5 mm, pin radius = 0.5 mm, TM_{01} mode (b) gap between adjacent edges = 0.5 mm, pin radius = 0.5 mm, TM_{21} mode, (c) gap between adjacent edges = 1 mm, pin radius = 1 mm, TM_{01} mode, (d) gap between adjacent edges = 1 mm, pin radius = 1 mm, TM_{21} mode.

3. RESULTS AND DISCUSSION

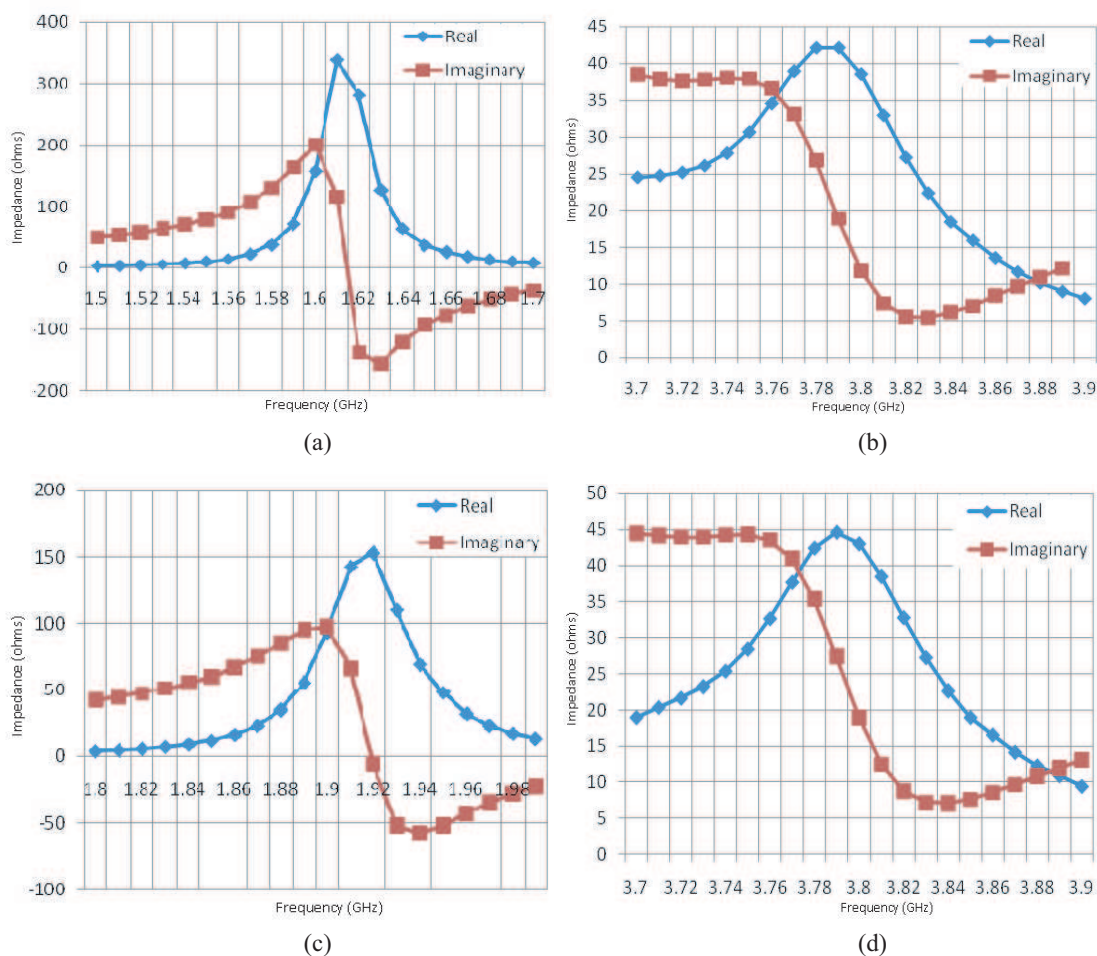The simulation work has been performed using the method-of-moment based software (IE3D from M/S Zeland software). For the simulation two circular patches are placed close to each other as shown in Fig. 1. The dimensions of the antenna used for simulation are given in Table 1. The input impedance of the antenna for different gaps between the adjacent edges of the patches and for different pin radius for TM_{01} and TM_{21} modes is shown in Fig. 2. The variation of the real and imaginary part with gap between adjacent edges at different frequencies is shown in Fig. 3(a) and Fig. 3(b) respectively for TM_{01} mode. From Fig. 3, it can be concluded that as gap between adjacent edges of the patches increases, resonant frequency increases and the value of real part increases and the value of imaginary part of the impedance is almost constant for TM_{01} mode. The variation of the real and imaginary part with gap between adjacent edges at different frequencies is shown in Fig. 4(a) and Fig. 4(b) respectively for TM_{21} mode. The resonant frequency and the value of real part of impedance decreases on increasing the gap between adjacent edges and the imaginary part increases at lower frequencies and decreases at higher frequencies on increasing the gap between adjacent edges for TM_{21} mode. The variation of the real and imaginary part with radius of shorting post is shown in Fig. 5(a) and Fig. 5(b), respectively for TM_{01} mode. The resonant frequency depends critically on the position and dimension of the shorting post. The resonant frequency increases on increasing the radius of shorting post and the value of real and imaginary part of impedance decreases on increasing the radius of shorting post for M_{01} mode as shown in Fig. 5. The variation of the real and imaginary part with radius of shorting post is shown in Fig. 5(a) and Fig. 5(b) respectively for TM_{21} mode. From Fig. 5, it is clear that the resonant frequency increases on increasing the radius of shorting post for TM_{21} mode.
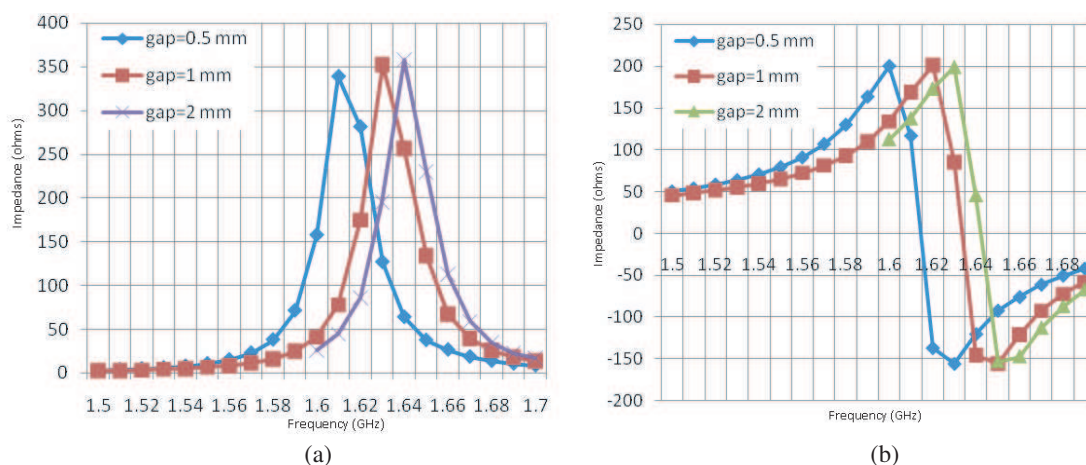
Figure 3: Variation of (a) real part, (b) imaginary part of the input impedance with gap between adjacent edges for TM_{01} mode (pin radius = 0.5 mm).
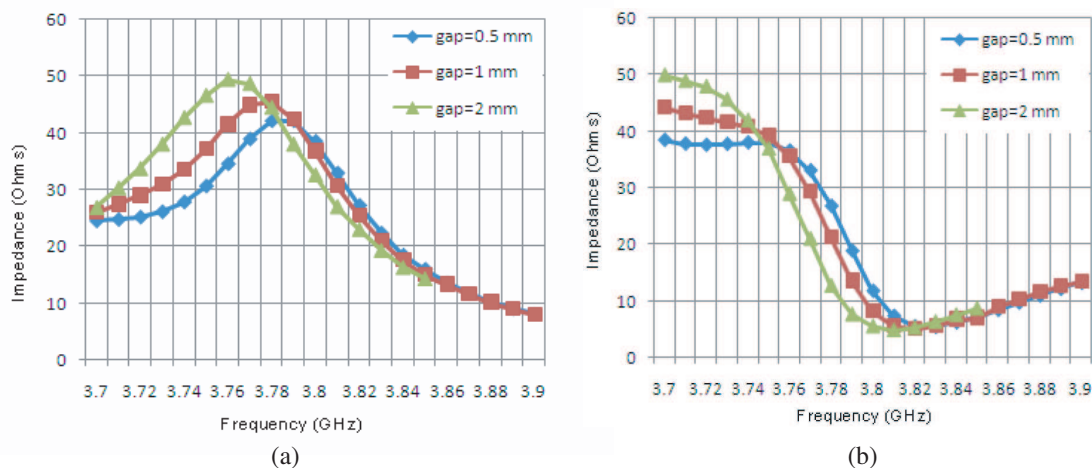
Figure 4: Variation of (a) real part, (b) imaginary part of the input impedance with gap between adjacent edges for TM_{21} mode (pin radius = 0.5 mm).

(a)   (b)
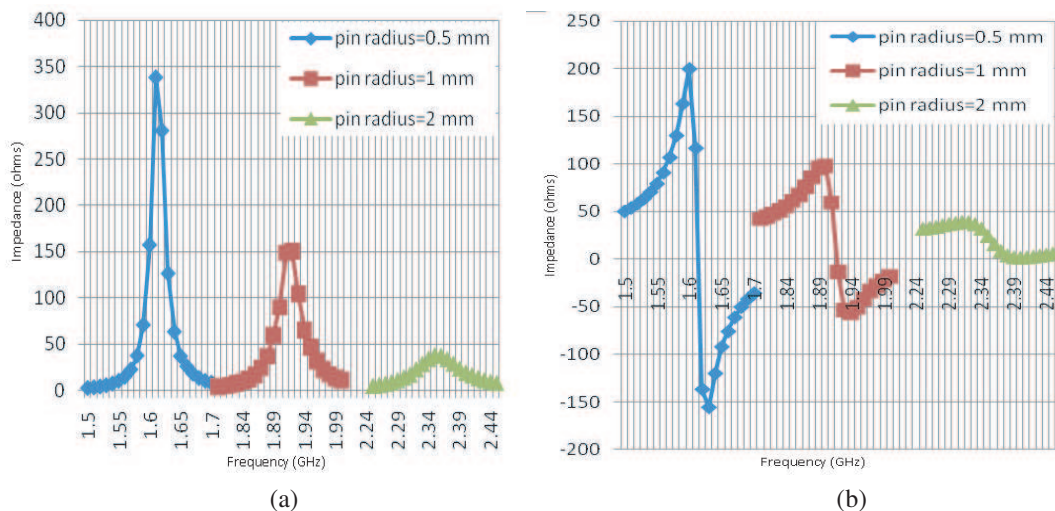
Figure 5: Variation of (a) real part, (b) imaginary part of the input impedance with pin radius for TM_{01} mode (gap between adjacent edges $= 0.5\,\text{mm}$) Frequency (GHz), Impedance (Ohms).

(a)   (b)
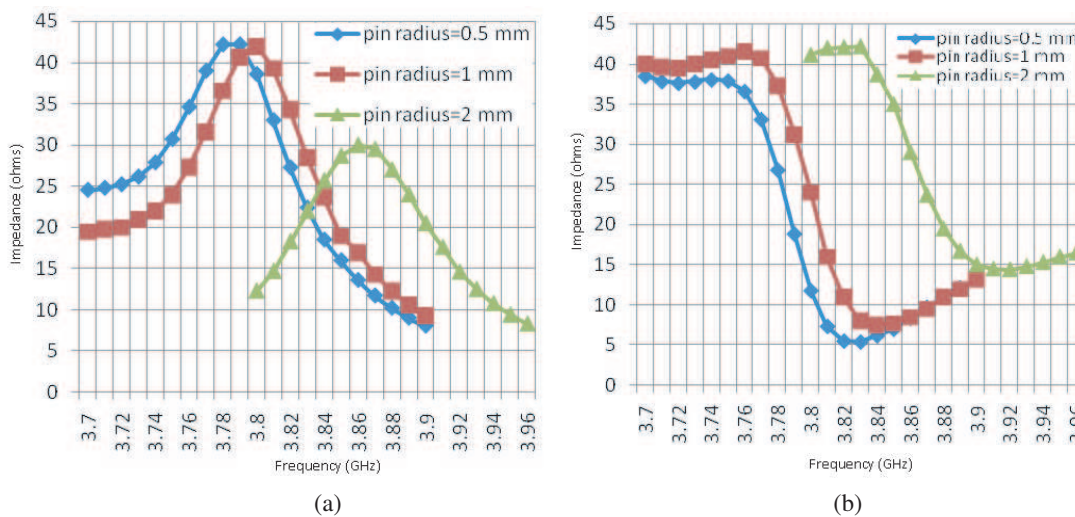
Figure 6: Variation of (a) real part, (b) imaginary part of the input impedance with pin radius for TM_{21} mode (gap between adjacent edges $= 0.5\,\text{mm}$).

Table 1: Dimension of the antenna.

Radius of shorted patch (r_1) (mm)	Radius of parasitic patch (r_2) (mm)	Thickness of substrate (h) (mm)	Dielectric constant (ε_r)
15	15	1.59	2.2

4. CONCLUSION

The size reduction of the microstrip antenna can be done by shorting the patch. Simulation studies on input impedance for gap coupled circular microstrip antennas loaded with shorting post in terms of $Z(1,1)$ parameter is performed. The variation of real and imaginary part of input impedance, for different modes, with gap between adjacent edges of patches as well as with radius of shorting post is discussed.

REFERENCES

1. Garg, R., P. Bhartia, I. Bahl, and A. Ittipiboon, *Microstrip Antenna Design Handbook*, Artrech House Publishers, 2001.

2. Kumar, P., A. K. Singh, G. Singh, T. Chakravarty, and S. Bhooshan, "Terahertz technology-a new direction," *Proc. of IEEE Int. Symposium on Microwave*, 195–201, Bangalore, India, 2006.

3. Kumar, P., G. Singh, and S. Bhooshan, "Gap-coupled microstrip antennas," *Proc. of Int. Conf. on Computational Intelligence and Multimedia Applications*, Vol. 4, 434–437, Sivakasi, India, 2007.

4. Pozar, D. M., "Microstrip antennas," *Proc. IEEE*, Vol. 80, 79–91, Jan. 1992.

5. Ray, K. P., S. Ghosh, and K. Nirmala, "Compact broadband gap coupled microstrip antennas," *IEEE Antennas and Propagation Society Int. Symposium*, 3719–3722, Jul. 2006.

6. Waterhouse, R., "Small microstrip patch antenna," *Electronics Letters*, Vol. 31, 604–605, 1995.

7. Dey, S. and R. Mitra, "Compact microstrip patch antennas," *Microwave Opt. Tech. Lett.*, Vol. 13, 12–14, Sept. 1996.

8. Porath, R., "Theory of miniaturized shorting post microstrip antennas," *IEEE Trans. Antennas Propagation*, Vol. 80, 41–47, Jan. 2000.

9. Kan, H. K. and R. Waterhouse, "Size reduction technique for shorted patches," *Electronics Letters*, Vol. 35, 948–949, 1999.

10. Tang, C. L., H. T. Chen, and K. L. Wong, "Small circular microstrip antenna with dual frequency operation," *Electronics Letters*, Vol. 33, 1112–1113, 1997.

11. Wong, K. L. and W. S. Chen, "Compact microstrip antenna with dual-frequency operation," *Electronics Letters*, Vol. 33, 646–647, Apr. 1997.

UWB Rectangular Ring Microstrip Antenna with Simple Capacitive Feed for Breast Cancer Detection

Sangam Kumar Singh and Arun Kumar Singh

Punjab Engineering College, Deemed University, Chandigarh 160012, U.T., India

Abstract— Breast cancer is a second leading cause of cancer in women today after lung cancer and is the most common cancer among women according to WHO. At present X-ray Mammography is approved technique for early breast cancer detection. But this technique have up to 20% false detection rate.

There will be approximately 250,000 new cases of breast cancer in India by 2015. At present, India reports around 100,000 new incidence of breast cancer. We can eliminate this false detection using Microwave imaging. Microwave imaging is defined as "seeing" the internal structure of an object by means of electromagnetic fields at microwave frequencies (300 MHz–10 GHz).

Antenna is a key component in microwave imaging system. Microstrip antenna can become right candidate for Microwave imaging system. Because it is compact and planar but major drawback is its narrow bandwidth. But using different technique we can increase bandwidth up to 70%.

Single Layer Rectangular Microstrip Antennas with simple capacitive feed, offering an impedance bandwidth up to 50% is designed in IISc., Bangalore, India 2007 [1]. The proposed antenna is modification of above antenna. In this antenna, we replace rectangular patch with rectangular ring patch which increases impedance bandwidth more than 66.44% having resonance frequency 6.05 GHz.

1. INTRODUCTION

X-Ray mammography is only detection method which is approved by U.S. Food and Drug Administration (FDA). In X-Ray Mammography, low dose X-ray are scattered on breast surface. But by this technique, we can not get 100% result, and 100% result is desirable in early stage breast cancer detection so that we can cure cancer at primary stage. X-Ray mammography is not recommended for young girl by FDA also. But we can eliminate above limitation if we use Microwave Imaging technique. At microwave frequency, there is significant difference of dielectric constant between normal and malignant breast tissue so that we can get clearer picture of malignant tissue [4, 5].

There are many UWB antenna are available for wireless application such as Vivaldi antenna, Horn antenna, Unipolar antenna, Bowtie antenna. But they have some merit and some demerit such as bulky, non planar, costly and bidirectional radiation pattern. Antenna array becomes complex and non planar surface if antenna is bulky and non planar. Bidirectional radiation pattern also reduces antenna gain which is not desirable [6–15]. Microstrip antenna can become right candidate for Microwave imaging system because it is in small size, light weight, low cast and also planar [16]. Major drawback of microstrip antenna is narrow bandwidth, typically 1–5% impedance bandwidth [16]. But using different techniques, we can get larger impedance bandwidth. One of such technique is aperture couple feeding method in stack configuration. Using this technique, first time in 2007 Aperture Coupled Stacked Patch Microstrip Antenna is designed with 77% impedance bandwidth at University of Bristol, U.K. for breast cancer detection [17]. But these configuration cause alignment issues while assembling and hence may increase the production cast. And Primary advantage of MSA lies in there eases of fabrication by standard lithographic techniques. Single Layer Rectangular Microstrip Antennas with simple capacitive feed, offering an impedance bandwidth up to 50% is designed in IISc., Bangalore, India 2007 [1]. The proposed antenna is modification of above antenna. In this antenna, we replace rectangular patch with rectangular ring patch which increases impedance bandwidth more than 66.4% having resonance frequency 6.05 GHz due to inductive impedance match with some approximation and also it has no misalignment problem, so antenna will be cheaper.

2. ANTENNA DESIGN AND DISCUSSION

We have design antenna using IE3D v.12 antenna simulator. Following are the antenna parameters.

Width of the radiator patch $(W) = 25.5$ mm.
Length of the radiator patch $(L) = 15.5$ mm.

Length of the feed strip $(s) = 3.7$ mm.
Width of the feed strip $(l) = 1.2$ mm.
Separation of the feed strip from the patch $(d) = 0.5$ mm.
Air gap between substrates $(g) = 6.0$ mm.
Relative dielectric constant $(\varepsilon_r) = 3.0$.
Thickness of substrate $(h) = 1.56$ mm.
Blank inner square area $= 5 \times 5$ mm^2.

Here, two type of rectangular patches are used, one patch are rectangular ring and other are small rectangular patch. Rectangular ring patch is used for radiation and small rectangular patch is used for coaxial probe feed. Hole on radiating patch will increase the bandwidth of antenna [16 18 19]. But it doesn't mean that increasing more and more area of hole, increase bandwidth of antenna. Because at certain limiting area of hole, antenna will go higher mode of propagation [19].

Length (L) and width (W) of rectangular ring patch is 15.5 mm and 25.5 mm. Length and width the feed strip is 3.7 mm and 1.2 mm. Above infinite ground plane, there is air gap of thickness 6.0 mm which have dielectric constant 1.07 and loss tangent 0.0009. A substrate having dielectric constant 3.0 cover air gap and thickness and loss tangent of substrate is 1.56 mm and 0.0013. Over substrate rectangular ring patch and feed stip patch are sited. Coaxial probe feed is used for excitation of antenna. Separation of rectangular ring from feed is 0.5 mm.

Figure 2 shows return loss of the proposed antenna. Proposed antenna have return loss less than -9.5 dB loss from frequency range 4.4 GHz to 8.42 GHz and bandwidth of antenna is 4.02 GHz. This wide bandwidth is useful for taking image with high resolution and it also help in fast signal processing. This antenna has 66.44% impedance bandwidth so it is UWB antenna because we all know that UWB devices have at least 20% impedance bandwidth.

(a) Cross-sectional View   (b) Top View
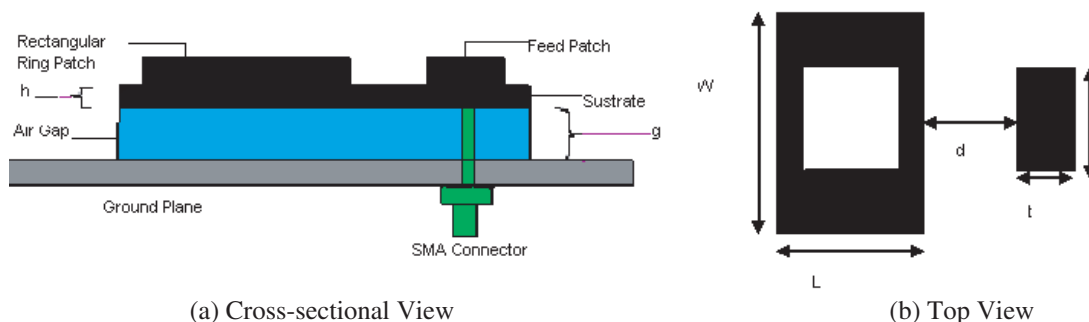
Figure 1: Geometry of rectangular ring microstrip antenna with small capacitive feed.
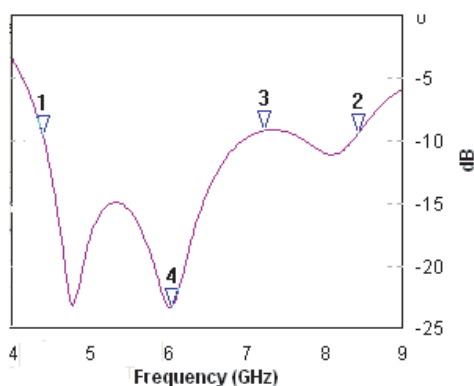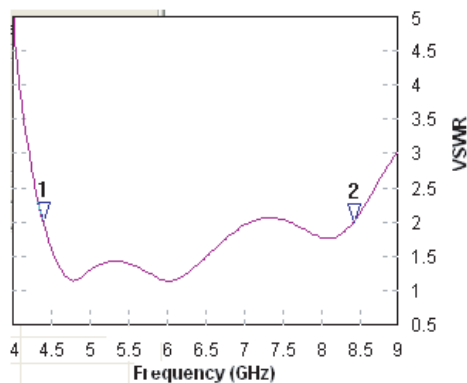

Figure 2: Return loss with frequency.   Figure 3: VSWR with frequency.

Figure 3 shows variation of VSWR with frequency which also show satisfactory result i.e., at all frequency VSWR within value 2.0 except 7.22 GHz where VSWR value is 2.06. Figure 4 shows two dimensional radiation pattern of antenna and HPBW is 163.64 degree. This large HPBW is

desirable in breast cancer detection because larger HPBW cover larger surface of breast. Figure 5 shows variation of antenna efficiency with frequency. Here maximum antenna efficiency is 87% and minimum antenna efficiency is 77.17% which are quit enough for image processing application.

Antenna gain is shown in Figure 6. All antenna gain is up to 6.88 dBi. And Figure 7 shows three dimensional view of radiation pattern.

All above results such as 66.44% impedance bandwidth, 163.64 degree HPBW, antenna efficiency more than 77.17%, antenna gain up to 6.88 dBi are very satisfactory results for microstrip antenna in microwave imaging system for early breast cancer detection. But radiation pattern are not flat in whole bandwidth and retune loss of antenna have one peaks between frequency range 4.4 to 8.42 GHz. Such things are undesirable but it will not create any significant error in image processing.
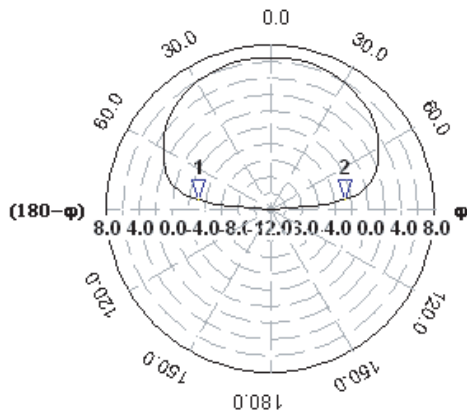
Figure 4: 2D radiation pattern of antenna.
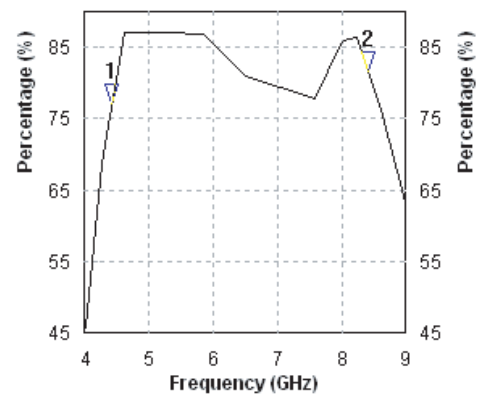
Figure 5: Efficiency of antenna.
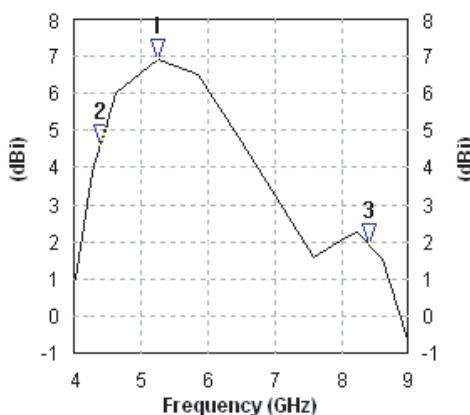
Figure 6: Gain with frequency for antenna.
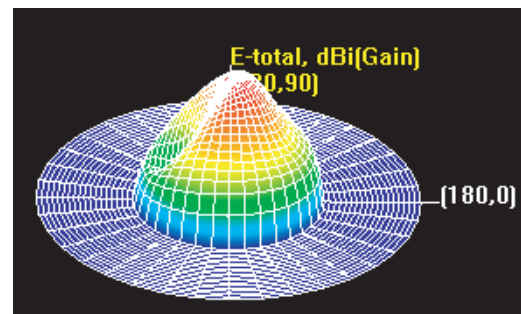
Figure 7: 3D radiation pattern of antenna.

REFERENCES

1. Kasabegoudar, V. G., D. S. Upadhyay, and K. J. Vinoy, "Design studies of ultra-wideband microstrip antenna with a small capacitive feed," *International Journal of Antenna and Propagation*, Vol. 2007, Article ID 67503, 8, 2007, doi:1155/2007/67503.
2. Mayhew-Ridger, G., J. W. Odendaal, and J. Joubert, "Single layer capacitive feed for wideband probe-feed microstrip antenna elements," *IEEE Transactions on Antennas and Propagation*, Vol. 51, No. 6, June 2003.
3. Vandenbosch, G. A. E. and A. R. Van de Capelle, "Study of the capacitive feed microstrip antenna elements," *IEEE Transactions on Antennas and Propagation*, 1994.
4. Fear, E. C., et al., "Confocal microwave imaging for breast cancer detection: Localization of tumors in three dimension," *IEEE Transaction on Biomedical Engineering*, Vol. 49, No. 8, August 2002.

5. Fear, E. C. and M. A. Stuchy, "Microwave breast tumor detection: Antenna design and characterization," *IEEE*, 2000.

6. Yun, X., E. C. Fear, and R. H. Johnston, "Compact antenna for radar-based breast cancer detection," *IEEE Transaction of Antennas and Propagation*, Vol. 53, No. 8, August 2005.

7. Shannon, C. J., E. C. Fear, and M. Okoniewski, "Dielectric fielded slotline bowtie antenna for breast cancer detection," *Electronics Letters*, March 31, 2005.

8. Chiappe, M. and G. L. Gragnani, "Vivaldi antennas as detectors for microwave imaging: Further steps in its radiation feature analysis," *IEEE International Workshop on Imaging Systems and Techniques*, Niagara Falls, Sheration Hotel on the Falls, May 13–14, 2005.

9. Yun, X., E. C. Fear, and R. Johnson, "Broadband cross-polarized bowtie antenna for breast cancer detection," *IEEE Antennas and Propagation Society International Symposium*, 2003.

10. Targonski, S. D. and R. B. Waterhouse, "Design of wide-band aperture stacked patch microstrip antennas," *IEEE Transaction on Antennas and Propagation*, Vol. 46, No. 9, September 1998.

11. Shannon, C. J., M. Okoniewski, and E. C. Fear, "A dielectric filled ultra-wideband antenna for breast cancer detection," *IEEE Antennas and Propagation Society International Symposium*, 2003.

12. Singhal, P. K., B. Dhaniram, and S. Banerjee, "A stacked square patch slotted broadband microstrip antenna," *Journal of Microwave and Optoelectronics*, Vol. 3, No. 2, August 2003.

13. Kanj, H. and M. Popovic, "Miniaturized microstrip-fed dark eyes antenna for near field microwave sensing," *IEEE Antennas and Wireless Propagation Letters*, Vol. 4, 2005.

14. Liang, J., C. C. Chiau, X. Chen, and C. G. Parini, "Printed circular disc monopole antenna for ultra-wideband applications," *Electronics Letters*, Vol. 40, No. 20, September 30, 2004.

15. Choi, S. H., J. K. Park, S. K. Kim, J. Y. Park, "A new ultra-wideband antenna for UWB applications," *Microwave and Optical Technology Letters*, Vol. 40, No. 5, March 5, 2004.

16. Kumar, G. and K. P. Ray, *Broadband Microstrip Antennas*, Artech House, Norwood, Mass, USA, 2003.

17. Nilavalan, R., I. J. Craddock, A. Preece, J. Leendertz, and R. Benjamin, "Wideband microstrip patch antenna design for breast cancer tumour detection," *IET Microwave Antenna Propagation*, Vol. 1, No. 2, 277–281, 2007.

18. Palanisamy, V. and R. Garg, "Rectangular ring and H-shaped microstrip antennas-altenatives to rectangular patch antenna," *Electronics Letter*, Vol. 21, No. 19, September 12, 1985.

19. Deshmukh, A. A. and G. Kumar, "Formulation of resonance frequency for compact microstrip antennas," *IEEE Antennas and Propagation Society International Symposium*, 2006.

Multiband Rectangular Ring Microstrip Antenna for UWB Wireless Applications

Sangam Kumar Singh and Arun Kumar Singh

Punjab Engineering College, Deemed University, Chandigarh 160012, U.T., India

Abstract— In this paper, we have designed rectangular ring micostrip antenna for UWB (3.1 GHz–10.6 GHz) wireless applications. The beauty of this antenna is the use of single patch which make it easy to fabricate consequently cost of antenna becomes cheaper. We have designed tri-band antenna which have resonance frequency 4.49 GHz, 6.17 GHz and 9.64 GHz and this antenna have 13.5%, 24.42%, and 6% impedance bandwidth respectively. Return loss at all three centre frequency have less than -24 dB. And antenna efficiency at all center frequencies are more than 95%. All simulations have been done using IE3D v.12 which is based on method of moment.

1. INTRODUCTION

Microstrip patch antennas are widely used because of there many merits such as the low profile, light weight, low cast and planar also. However patch antenna have a disadvantage of narrow bandwidth typically 1–5% impedance bandwidth. Researchers have made many efforts to overcome this problem and many configurations have been presented to increase the bandwidth [1–4]. Many configurations while increase the bandwidth but it also increase the production cast of antenna due to misalignment of stacked patches. To overcome misalignment problem, we have to use lesser no of patches. In this paper we have design antenna using only single radiation patch which make it easier to fabricate consequently antenna becomes cheaper. We have use coaxial probe feed technique which create inductive reactance to compensate inductive impedance, we use capacitive patch. And coaxial probe are connected with capacitive patch [2–4].

2. ANTENNA DESIGN AND DISCUSSION

Antenna which can work in more than one frequency region either for transmitting or receiving electromagnetic (EM) waves are termed as Multiband Antenna. Such antenna are usually used for dual-band, tri-band, quad-band and penta-band application [5–9]. Using only single antenna in the multi signal band is better than using the each antenna element in the separate way. Here we have design proposed antenna using IE3D v.12 simulator which are based on method of moment and moment method is very accurate than its counter part FDTD but its take more memory. Parameters used for design of proposed antenna are given below.

Width of the radiator patch (W) = 25.5 mm.
Length of the radiator patch (L) = 15.5 mm.
Length of the feed strip (s) = 3.7 mm.
Width of the feed strip (l) = 1.2 mm.
Separation of the feed strip from the patch (d) = 0.5 mm.
Air gap between substrates (g) = 6.0 mm.
Relative dielectric constant (ε_r) = 3.0.
Thickness of substrate (h) = 1.56 mm.
Blank inner square area = 10×10 mm^2.

Here, two type of rectangular patches are used, one patch are rectangular ring and other are small rectangular patch. Rectangular ring patch is used for radiation and small rectangular patch is used for coaxial probe feed. Length (L) and width (W) of rectangular ring patch is 15.5 mm and 25.5 mm. Length and width the feed strip is 3.7 mm and 1.2 mm. Above infinite ground plane, there is air gap of thickness 6.0 mm which have dielectric constant 1.07 and loss tangent 0.0009. A substrate having dielectric constant 3.0 cover air gap and thickness and loss tangent of substrate is 1.56 mm and 0.0013. Over substrate rectangular ring patch and feed stip patch are sited. Blank inner square area is 10×10 mm^2. Coaxial probe feed is used for excitation of antenna. Separation of rectangular ring from feed is 0.5 mm. Figure 1 shows cross-sectional and top view of Rectangular Ring Microstrip Antenna.

Figure 2 shows return loss of antenna with different frequencies. From Figure 2 we can see that return loss at all center frequencies are less than -10 dB. At center frequency 4.49 GHz return loss is
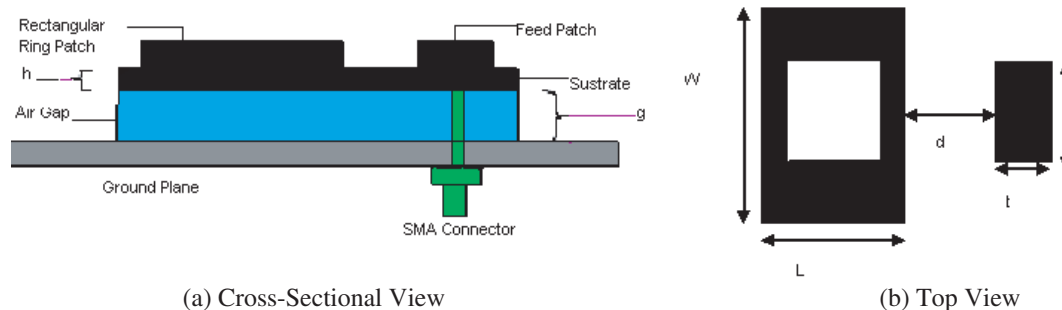
(a) Cross-Sectional View   (b) Top View

Figure 1: Prototype rectangular ring microstrip antenna.

-33.36 dB ($BW = 13.5\%$), at center frequency 6.17 GHz return loss is -24.25 dB ($BW = 24.42\%$) and at center frequency 9.68 GHz is -32.27 dB ($BW = 6\%$). Figure 3 shows VSWR variation with frequency for antenna. At all center frequencies, value of VSWR is less than 2.0. Figure 4 shows 2D Radiation pattern of antenna at different center frequencies which also give satisfactory result. Figure 5 shows efficiency of antenna at different frequencies. At all center frequencies antenna efficiency are more than 95% i.e., at center frequency 4.49 GHz antenna efficiency is 95%, at center frequency 6.17 GHz antenna efficiency is 96.29% and at center frequency 9.68 GHz antenna efficiency is 99.23. Figure 6 shows antenna gain with frequency. At all center frequency antenna gain is more than 5 dBi i.e., at center frequency 4.49 GHz antenna gain is 6.88 dBi, at center frequency 6.17 GHz antenna gain is 6.82 dBi and at center frequency 9.68 GHz antenna gain is 5.29. Figure 7 shows 3D radiation pattern of antenna at center frequencies 4.49 GHz, 6.17 GHz and 9.68 GHZ. All above result are quit satisfactory for UWB wireless application use.

3. RESULTS
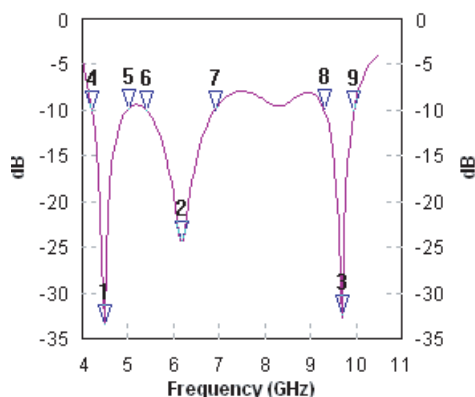
Figure 2: Return loss with frequency.
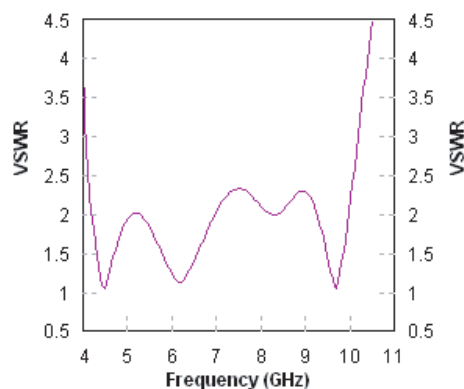
Figure 3: VSWR with frequency.

(a)

(b)

(c)
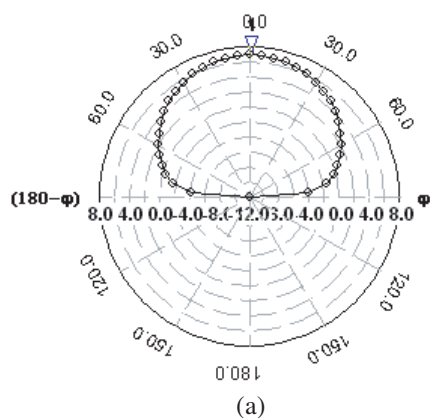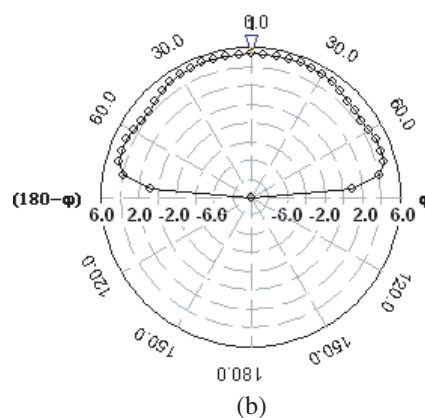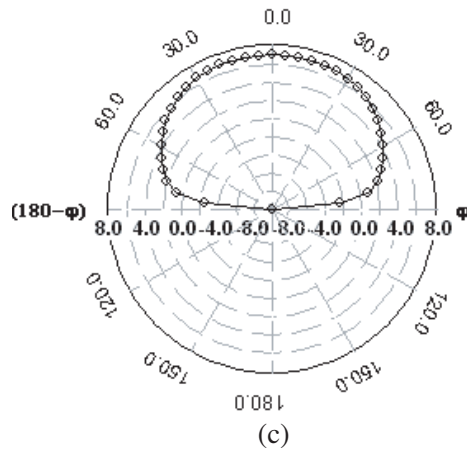
Figure 4: (a) 2D radiation pattern at 4.49 GHz. (b) 2D radiation pattern at 6.17 GHz. (c) 2D radiation pattern at 9.68 GHz.
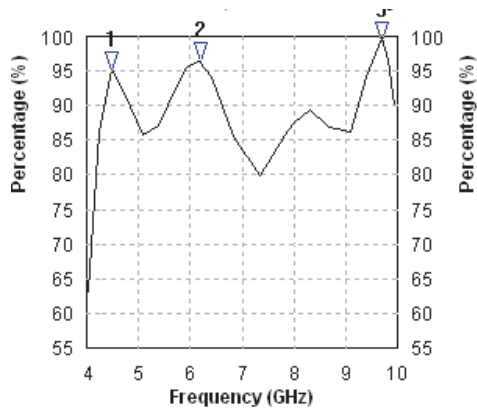
Figure 5: Antenna efficiency for antenna.
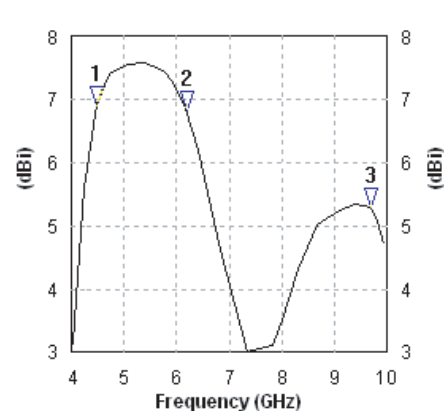
Figure 6: Antenna gain with frequency.
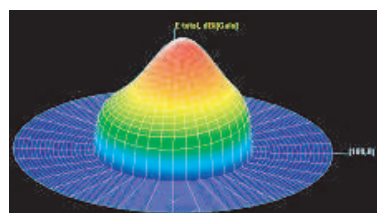
(a)
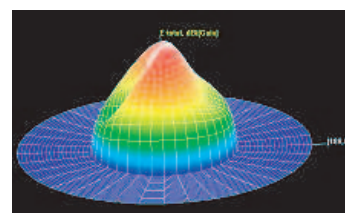
(b)
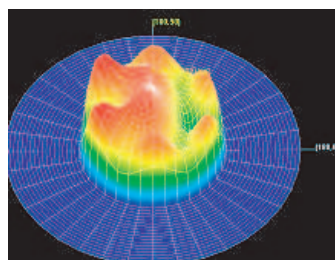
(c)

Figure 7: (a) 3D radiation pattern at 4.49 GHz. (b) 3D radiation pattern at 6.17 GHz. (c) 3D radiation pattern at 9.68 GHz.

REFERENCES

1. Choi, S. H., J. K. Park, S. K. Kim, and J. Y. Park, "A new ultra-wideband antenna for UWB applications," *Microwave and Optical Technology Letters*, Vol. 40, No. 5, March 5, 2004.

2. Palanisamy, V. and R. Garg, "Rectangular ring and H-shaped microstrip antennas-alternatives to rectangular patch antenna," *Electronics Letter*, Vol. 21, No. 19, September 12, 1985.

3. Deshmukh, A. A. and G. Kumar, "Formulation of resonance frequency for compact microstrip antennas," *IEEE Antennas and Propagation Society International Symposium*, 2006.

4. Kasabegoudar, V. G., D. S. Upadhyay, and K. J. Vinoy, "Design studies of ultra-wideband microstrip antenna with a small capacitive feed," *Internation Journal of Antenna and Propagation*, Vol. 2007, Article ID 67503, 8, 2007, doi:1155/2007/67503.

5. Kshetrimayum R. S., "Printed monopole antenna for multiband applications," *Intenational Journal of Microwave and Optical Technology*, Vol. 3, No. 4, September 2008.

6. Sheta, A. F., et al., "Multi-band operation of a compact H shaped microstrip antenna," *Microwave and Optical Technology Letters*, Vol. 35, No. 5, December 5, 2002.

7. Constantine, J., et al. "New multiband microstrip antenna design for wireless communication," *IEEE Antennas and Propagation Magazine*, Vol. 49, No. 6, 2007.

8. Salamat, C. D., M. Haneishi, and Y. Kimura, "L-probe fed multiband microstrip antenna with slots," *Proceedings of Asia-Pacific Microwave Conference 2006*, 2006.

9. Asrokin, A., M. K. A. Rahim, and M. Z. A. Abd Aziz, "Dual band microstrip antenna for wireless LAN application," *2005 Asia-Pacific Conference on Applied Electromagnetics Proceedings*, Malaysia, 2005.

Effective Electromagnetic Media for FDTD-PIC

Lars D. Ludeking and Andrew J. Woods

ATK-Mission Systems Group, Newington, Va 22122, USA

Abstract— Simulation and modelling using Finite Difference-Time Domain Particle-in-Cell (FDTD-PIC) techniques often requires the utilization of various electromagnetic field-material response properties to provide high-fidelity results. The various property models available have practical limitations, particularly when employed in problems which require the participation of ambient plasmas. The authors discuss some of the techniques and issues that must be addressed in obtaining robust simulations suitable for engineering design and demonstration while faithfully enforcing physical principles.

Simulations are used to provide insight into physical processes and guide the design and development of a variety of technologies. Historically, this approach has been applied in the domain of vacuum electronic devices; however, the methodology spans a much broader range of applications, which includes areas such as: bio-electric effects, bio-medical, radar cross sections, and high power microwave devices.

The trademarked ATK software package, commercially known as the "MAGIC Tool Suite", is a widely used FDTD-PIC software suite. The basic approach for FDTD-PIC is well known and the literature contains the basic equations and descriptions for the solution of Maxwell's equations. The FDTD method is very attractive because of its computational efficiency and relative ease of implementation.

As is well known, two of the principle challenges to the correct modelling of any electromagnetic field-material interaction lie in (1) the adequate representation of the model boundary conditions (BC), and (2) the representation of dynamic or frequency dependent material properties. Simple BC idealizations, such as the perfect conductor, the periodic, or mirror symmetric boundaries are readily treated. However, the open boundary termination is one which many researchers have devoted enormous amounts of energy in perfecting. The basic issue is the completion of the simulation domain with a one-sided wave equation. There are in fact two complementary aspects that are of value in 'MAGIC'. These are the scattered wave component and the incident wave component. In addition, there is the issue of modelling interior domains with permittivity and conductivity.

Simulations are widely used and often misused. Novice users of this approach often fail to understand fundamental issues and constraints in the application of the various models and how they may interact. In this paper, the authors will address some of the FDTD-PIC constraints associated with the following algorithms:

- The use of interior electric and magnetic conductivity models for representing effective absorbers and mimicking Field Effect Transistors (FETs).
- Conductor surface loss models and the side effects.
- Matched Phase Velocity method for bounding a domain and the introduction of an electromagnetic pulse into the interior.
- The optimized "free space" impedance matching method, or the "lazy man's version of PML"

  [What are its advantages and disadvantages?]
- The use of the Convolutional Perfectly Matched Layer (CPML) method and the caveats when applied to non-plasma free problems.

1. INTRODUCTION

Simulations with Finite-Difference-Time-Domain Particle-in-Cell (FDTD-PIC) computational methods are used to provide insight into a broad spectrum of electromagnetic field applications. Historically, these methods have been used in the research, design, and improvement of technologies that include microwave amplifiers, antennas, sensors, fiber optics, accelerator component design, beam propagation in vacuum & atmospheric gases, lasers, pulsed power sources, plasma switches, RF induced plasma heating, field emitter arrays, semiconductor devices, radiation reflection chambers, wave scattering, coupling analysis, bio-electric effects, radar cross section, and high power sources.

Alliant Techsystems Inc (ATK) is an aerospace and defense company. It is engaged in a variety of modern research and development projects. Included in the inventory of its specialized research tools are two proprietary engineering and scientific computational physics software suites which employ the FDTD-PIC methodology. These are the MAGIC Tool Suite and the LSP Suite. Both of these software packages are currently available as licensed commercial software products to the international research community. These FDTD-PIC software suites are both applications, which provide the capability to model and simulate a wide variety of vacuum electronics problems and beam-wave interaction issues. Both contain specialized packages of algorithms to address a broad set of problems and research issues. Both suites are available for execution on Windows OS and Linux/Unix OS platforms. The MAGIC Tool Suite is designed principally for use on Windows OS systems, and emphasizes vacuum electronics applications and moderate beam and plasma density interactions. Although the LSP suite is primarily designed for use in a Unix/Linux environment, it can be used in a Windows OS environment as well. It has been used primarily for dense plasma and intense beam wave interactions. Both packages include eigenmode solvers as well as dynamic time domain algorithms.

2. MAGIC TOOL SUITE

In this article, we will focus on the MAGIC Tool Suite [1], although much of the same features and capability is present in the LSP Suite as well. The MAGIC Tool Suite is a user-configurable electromagnetic particle-in-cell simulation code used for the modelling and simulation of beam wave interactions and particularly for electro-energetic modelling of processes that involve interactions between space charge and electromagnetic fields. As is common for all such codes, it begins from a specified initial state, typically vacuum with zero fields and no particles, although it can start with arbitrary fields and particles as desired. 'MAGIC' then simulates physical processes as they evolve in time. It makes use of Maxwell's time-dependent equations (specifically Faraday's law and Ampere's law) to obtain the time dependent electromagnetic fields. In addition, in the presence of particles, it solves the complete Lorentz force equation to obtain relativistic particle trajectories. The continuity equation is solved to provide current and charge densities for Maxwell's equations. This approach, commonly referred to as electromagnetic particleincell (PIC), provides selfconsistence, i.e., interaction between charged particles and electromagnetic fields.

The 'MAGIC' software is a general purpose FDTD-PIC engine, which is user configurable. It has been provided with algorithms to represent/generate structural geometries, material properties, incoming and outgoing waves, particle emission processes, and so forth. The software has been used in a wide range of plasma physics problems. Configurable elements of the software may be broadly categorized to include the following:

- Control language for user scripts

- Automatic and user controlled grid generation

- Materials and material properties

- Geometry generation and coordinate systems

- Maxwell time domain and eigenmode algorithms

- Particle and Lorentz algorithms for multiple sources and species

- Output diagnostics which include interior sampling and movies

3. INTERIOR MEDIA

In 'MAGIC', there are four basic classes of material media of primary significance, they are: (a) perfect conductors, (b) perfect (very low loss) dielectrics, (c) conductive zones (essentially lossy dielectrics), and surface loss models. (In addition, there are items such as resistors, inductors, current drivers, and membranes, as well as 1-d extensions to represent transmission lines).

The perfect conductor is trivial to enforce, and is of primary importance in providing a closed simulation zone and representing cavity and circuit structures. Used in conjunction with the surface loss model, it provides an accurate representation of resonant cavities. By empirically adjusting the surface loss, one can match the appropriate Q values for tubes, such as klystrons where Q and R/Q are essential to obtain proper performance. If one has a good experimental value for the surface loss character of the material employed, then one can predict the Q.

Dielectrics are frequently of interest in a variety of ladder circuit devices and travelling wave tubes, as well as in couplers and filters. Two important properties of dielectrics in the presence of ambient charged plasma (a) they must collect and trap charged particles; and (b) may suffer emission breakdown in presence of intense fields. The 'MAGIC' model allows for both conditions based on the user selection of property characteristics. In addition, the dielectric properties may be spatially varying and temporally varying (i.e., depending on ambient field intensity), thus allowing application in items such RF filters and limiters.

Conductivity is applied within specified spatial zones and results in an additional current as specified by Ohm's Law, $\mathbf{J} = \sigma\mathbf{E}$, which is applied to Ampere's Law as an additional current source. Conductive regions provide for a number of convenient features. First, they allow the representation of physically lossy dielectrics, when used in conjunction with the dielectric model. That is the models may be superimposed on the same physical simulation zones. They may be used as artificial load zones to provide an empirical method of matching a particular load characteristic or removing RF power. Additionally, in order to model interior absorbers with high fidelity in an RF chamber, we provide for an artificial magnetic conductivity which may be used in conjunction with the electric conductivity or in isolation. By using a suitable spatial variation it is possible to closely mimic the response of the absorptive materials used in chamber. This model is defined to resemble Ohm's law and is given by $\mathbf{J}_B = \sigma_B\mathbf{H}$ and is applied to Faraday's law. The magnetic conductivity, σ_B, is given in ohms/m. The attenuation of the magnetic field energy is at the same rate as that of the electric field energy, when $\sigma_B = \sigma_E Z_f^2$, where $Z_f (= (\mu_o/\varepsilon_o)^{1/2} = 377\,\text{ohms})$ is the vacuum impedance. The magnetic conductivity may be used to reduce stray charging fields to very low values [2]. This has been used for test chamber modelling [3], in which the current sensitivity measures 1 electron/1 nanosecond.

The surface loss model in 'MAGIC' is often used when modelling cavities with (or without) an ambient current beam. In its simplest instance, the basic model only provides attenuation of the magnetic field transverse to a conducting surface. It is based on generating an effective skin depth loss, assuming a constant RF frequency and knowledge of the skin depth for the particular conducting material. No attempt to model the behavior of the fields within the metal is actually made. Specifically, a surface consisting of a non-ideal metal possesses a non-zero, parallel electric field which is proportional to the parallel magnetic field. When this non-zero parallel electric field is accounted for in Faraday's Law, it results in the addition of an effective magnetic conductivity term for the parallel magnetic field at the surface, e.g.,

$$\partial_t\mathbf{B}_{surface} = -\nabla \times \mathbf{E}_{normal} - \eta(f_0)\mathbf{B}_{surface}, \tag{1}$$

where the factor $\eta(f_0)$ contains the skin depth information at a fixed frequency, f_0. This provides a very convenient mechanism for evaluation of the surface loss for cavities. It does however suffer from a serious liability when applied at a fixed frequency for beam-cavity problems. In this case, the beam transport tunnel will suffer an artificial charging because of the extra term introduced in Faraday's law when the beam is largely DC in character and carries little or no RF characteristic. Then the frequency used in the surface loss equation is incorrect. A more sophisticated (and costly) model may also be used. In this model, the weighted frequency $< f(t) >$ is locally and dynamically evaluated for each surface cell. This allows for a response to have a weak response on surfaces subject to essentially dc fields, and in addition captures the response effects for higher harmonics of the principle cavity/circuit frequency.

4. MAXWELL SOLVER TERMINATION

A major challenge for any Maxwell solver lies in the truncation of the simulation domain. Perfect conductor and periodic & mirror symmetries are easily treated. The open (or unconstrained) boundary termination is one which has received a great deal of attention. There are two crucial issues: first is the completion of the simulation domain with an effective one-sided wave equation. (This closes or completes Maxwell's equations). The second issue is the boundary truncation must be insensitive (transparent) to the passage of ambient particles through the boundary.

For the first issue, the closure of Maxwell's equations, there are several competing approaches, each with its own advantages and disadvantages. Three implemented in 'MAGIC' are: (1) the matched phase velocity method, (2) the optimized free space method, (3) and finally the CPML method.

The matched phase velocity method, while historically the oldest approach, serves in general to be the most robust for two important aspects of robust simulation modelling, beam wave propa-

gation and the boundary termination with wave injection. For boundary termination, two complementary aspects of this method are of value in MAGIC; these are the separation of the scattered wave component and the incident wave component.

The matched phase velocity method separates artificial numerical dispersion from physical dispersion and compensates for the computational artefact. It begins by writing the hyperbolic wave propagation equation in (Cartesian) operator form as follows:

$$(\partial^2/\partial t^2 - v^2 \partial^2/\partial x^2)E_x = 0$$
$$[(\partial/\partial t + v\partial/\partial x)(\partial/\partial t - v\partial/\partial x)]^n E_x = 0$$

For $n = 1$, the standard wave equation results. However, it is possible to operate on the wave field multiple times and still satisfy the wave equation. Thus in the finite difference representation, it is implicit that for a reasonable approximate value of the phase velocity, the first order equation is approximately zero, and one expects that multiple applications of the difference operator will provide improved results, leading to a better approximation to a perfect match. You will note that the wave equation is separated into a forward and backward wave operator. Ideally one or the other of the two operators operating on the wave field results in 0.

The first order ($n = 1$) solution to the operator equation, requires knowledge of the boundary value at the preceding time step and the values of the field one spatial step to the interior both for the current and the preceding time step. Application of higher order solutions requires information further into the simulation and farther back in time. Thus for $n = 1$, we need three time-space fields to obtain the edge value. For $n = 2$, we require eight time-space fields, and so forth to higher order.

An additional feature of this approach is that one separates the scattered and incident wave components. Because of the symmetry of the operations, it is a trivial extension to use it to introduce arbitrary EM waves into the simulation domain.

The second order ($n = 2$) matched phase velocity method has been implemented in 'MAGIC' and exhibits about 18 dB better performance in reduced reflection than does the first order.

The optimized free space method for boundary termination has properties similar to the electric and magnetic conductivity models. The legacy impedance-matching free space treatment [1, 4, 5] has been revised to provide for greater latitude in the magnetic and electric conductivities. The conventional free space approach adds electric and magnetic conductivity terms to the free space Maxwell's equations in a finite thickness boundary layer with carefully chosen values to preserve phasing. The electric conductivity is

$$\sigma_E = 4\pi c \varepsilon/\Delta,$$

where Δ is the free space layer width. The magnetic conductivity for a TE excitation is

$$\sigma_B^* = \sigma_E(377\beta)^2.$$

And where the phase velocity factor of the mode is given by

$$\beta = 1/\left[1 - (f_{cutoff}/f)^2\right]^{1/2} \tag{2}$$

This method serves quite well for absorption of waves of modest bandwidth problems. This method is intolerant of particle beam transport through the zone to the termination. Although the optimized method allows for a reasonable attenuation of the space charge of the beam, it is usually at the cost of increased reflection. It also fails to allow for injection of a wave into the simulation.

While the CPML [6–9] method provides the broadest "passive" absorption of outgoing waves with minimal reflection, it suffers from one minor and one major liability for beam wave propagation problems. First, the minor issue is that as a passive boundary it is not suitable for wave injection. This can of course be overcome by driving an interior split wave, but at the cost of an extra computational zone for the absorbed portion of the driven wave. Second, and more severe, is that it is unsuitable (and frequently unstable) to allow a particle beam to exit through the boundary zone. It appears that this boundary method attenuates the transverse and longitudinal beam charge by different amounts, depositing some of the charge on the numerical grid.

CPML adds electric Ψ_E and magnetic Ψ_H terms to the free space Maxwell's equations for transverse field components in similar layers as the conventional treatment. The remarkable property

that waves of all frequencies and angles of incidence are perfectly absorbed makes CPML theoretically effective in the near field for evanescent waves, as well as in the far field. The Ψ terms are obtained from auxiliary equations of the form:

$$\Psi^n = b\Psi^{n-1} + c\partial H_z/\partial y$$

where

$$b = \exp(-\sigma'\Delta t/\varepsilon) \quad \text{and} \quad c = (b-1)\sigma/(\sigma'\kappa).$$

The b and c coefficients above are derived using the convolution theorem in the inversion of the Fourier transform frequency domain equations in the stretched coordinate space with the modified electric and magnetic conductivities σ' and $\sigma^{*'}$. $\sigma' = \sigma + \kappa a$, where 'a' is a power law diminishing conductivity versus position in the layer and κ is an "attenuation amplifier" ($\kappa = 1$ in this work). Setting "a" to zero results in the Uniaxial PML (UPML) [7].

The resulting equations including finite difference forms are fully developed in [7], and are not repeated here. A note here on the difficulty encountered in getting results which agree with Gedney may be helpful to others, however. A misunderstanding of the development led to the inclusion of the conventional conductivity terms in the CPML region along with the new Ψ terms simultaneously. Subsequent examination of Gedney's code which he kindly provided [6] revealed that these terms should not have been included in the vacuum representation of the CPML region. Gedney included the conventional conductivity terms along with the Ψ terms to allow for a conducting host medium as well as the vacuum representation employed here.

5. CONCLUSIONS

The MAGIC Tool Suite (and the LSP Suite) offers a broad range of algorithms for examining beam-wave interactions as well as circuit problems that are free of ambient charged plasmas. The current versions of these tool suites have been focused on providing solid algorithms that generally do not interfere with each other and provide ample opportunity to configure novel structures and simulate novel mixtures of beam-wave and electro-energetic regimes. The user configurable approach provides great flexibility in examining many interesting experimental setups. Additional information on the MAGIC Tool Suite (http://www.magictoolsuite.com) and the LSP Suite (http://www.lspsuite.com) can be found on our website and helpdesk. The helpdesk allows one to download copies of the MAGIC User's Manual and contains other useful links and information for those with an interest in the FDTD-PIC.

ACKNOWLEDGMENT

REFERENCES

1. Goplen, B., L. Ludeking, D. Smithe, and G. Warren, "User-configurable MAGIC for electromagnetic PIC calculations," *Computer Physics Communications*, Vol. 87, 54–86, 1995.
2. Woods, A. and L. Ludeking, "Low noise static charging with MAGIC," *IVEC*, 2008.
3. Garner, S., N. Lehman, and R. A. Schill, Jr., "Spatial distribution of electron stimulated electron desorption from a metal surface," presented at the *IEEE International Power Modulator Conference 2008* for publication.
4. Gilbert, J. and R. MacGurn, "FLTSATCOM problem outer boundary," Mission Research Corp, Albuquerque, NM, AMRC-N-111, April 1979.
5. Goplen, B., "Boundary conditions for MAGIC," *Twenty-third Annual Meeting, APS Division of Plasma Physics*, October 12–16, 1981.
6. Gedney, S., Private Communication, June 2007.
7. Taflove, A. and S. Hagness, *Computational Electrodynamics: The Finite Difference Time Domain Method*, 3rd Edition, Artech House, 2005.
8. Alfadhl, Y. and X. Chen, "Numerical dosimetry of anatomically inhomogeneous animal models using an improved FDTD boundary implementation," *IEEE Antennas and Propagation Society International Symposium*, Vol. (3A), 819–822, 2005.
9. Gedney, S., "Perfectly matched layer absorbing boundary conditions," Ch. 7, A. Taflove and S. Hagness, *Computational Electrodynamics: The Finite Difference Time Domain Method*, 3rd Edition, Artech House, 2005.

Solving Guided Wave Modes in Plasmonic Crystals by Interfacial Operator and Coupling Interface Approach

Yu-Chen Shu[1, 4], **Chien C. Chang**[2, 4], **I-Liang Chern**[3, 4], **and Ying-Hong Liu**[1]

[1]Division of Mechanics, Research Center for Applied Sciences, Academia Sinica, Taipei 115, Taiwan, R.O.C.
[2]Institute of Applied Mechanics, National Taiwan University, Taipei 106, Taiwan, R.O.C.
[3]Department of Mathematics, National Taiwan University, Taipei 106, Taiwan, R.O.C.
[4]Taida Institute for Mathematical Sciences, Taipei 106, Taiwan, R.O.C.

Abstract— It has been difficult to solve eigenmodes of plasmonic crystals in two or three dimensions either analytically or numerically. In this study, we present an interfacial operator approach for solving guided wave modes of plasmonic crystals. They are formulated as an eigenvalue problem of the wavenumber along the axis of the crystal. In this formulation, the permittivity and permeability of the metallic component can be arbitrary functions of frequency. Moreover, a coupling interface method is introduced to facilitate accurate treatment of the interface conditions with an arbitrary shape between the metal and host materials. Numerical results are illustrated for different shapes of plasmonic crystals, layered, cylindrical and split-ring structures. The physical significance is discussed. Finally, it is demonstrated that the present method can resolve fine eigenmodes of the split-ring structure.

1. INTRODUCTION

Plasmonics is considered to be the strongest interplay of both optical and electronic data transfer along a tiny metal [1]. It has found important applications in modern nano-technology such as magneto-optic data storage, microscopy, solar cells, sensors for detecting biologically molecules, and plasmonic crystals [2]. Plasmonic crystals consist of periodic arrays of metalic and dielectric materials, and are useful to transmit sub-wave-length signals. Propagation of electromagnetic waves along the axis of a plasmonic crystal is now formulated as an eigenvalue problem of the axial wavenumber. In this formulation, the structure of the metal-dielectric composite material is considered to be periodic, and the permittivity and permeability of the metallic component can be arbitrary functions of frequency.

There are a few other methods [3–7] which have been applied to solve for the eigenmodes of plasmonic crystals. But they are not direct method. The present approach, the augmented coupling method, is a second-order finite difference approach to this eigenvalue problem. Its ingredients include: (i) an interfacial operator approach [8] (ii) an extension of the coupling interface method [9] which can treat elliptic interface problem with arbitrary shape under Cartesian grid structure. The discretization procedure is a dimension-by-dimension approach. The new ingredient in the present work is that the interface condition is introduced by interfacial variables and realized at uniformly distributed interfacial grid points, instead at the intersections of the interface and the underlying rectangular grid lines in the previous work. Modications on derivation of coupling equation and interpolation are needed.

2. GUIDED WAVE PROPAGATION IN METAL-DIELECTRIC COMPOSITE MATERIAL

Consider a waveguide along the axis of two-dimensional crystals. The crystal geometry is assumed to be periodic in the x-y plane with a unit cell. The governing equation for electromagnetic (EM) wave propagation in such media is the macroscopic Maxwell equations.

Since the waveguide is homogeneous in the z direction, we look for monochromatic modes of the form:

$$\mathbf{E}(x, y, z, t) = (E_x, E_y, E_z) e^{i(k_z z - \omega t)}, \tag{1}$$

$$\mathbf{H}(x, y, z, t) = (H_x, H_y, H_z) e^{i(k_z z - \omega t)}, \tag{2}$$

where $\mathbf{E}$ and $\mathbf{B}$ are the electric and magnetic fields; ω is the frequency; k_z is the wave number along the z-direction (axial wavenumber); (E_x, E_y, E_z) and (H_x, H_y, H_z) are functions of (x, y). At the cell boundary the Bloch boundary condition is applied to two fields:

$$\mathbf{E}(x + a_x, y + a_y, z, t) = e^{i(k_x a_x + k_y a_y)} \mathbf{E}(x, y, z, t), \tag{3}$$

$$\mathbf{H}(x + a_x, y + a_y, z, t) = e^{i(k_x a_x + k_y a_y)} \mathbf{H}(x, y, z, t), \tag{4}$$

where a_x, a_y are length and width of the unit cell; (k_x, k_y) is Bloch wave vector. Substituting (1) and (2) into the Farady and Ampere equations, we get

$$\nabla_2 \cdot \left(\frac{\varepsilon_0 \varepsilon \nabla_2 E_z}{\Lambda} \right) + \nabla_2 \times \left(\frac{k_z \nabla_2 H_z}{\omega \Lambda} \right) = -\varepsilon_0 \varepsilon E_z \tag{5}$$

$$\nabla_2 \cdot \left(\frac{\mu_0 \mu \nabla_2 H_z}{\Lambda} \right) - \nabla_2 \times \left(\frac{k_z \nabla_2 E_z}{\omega \Lambda} \right) = -\mu_0 \mu H_z \tag{6}$$

where $\Lambda = (\omega/c)^2 \varepsilon \mu - k_z^2$, $\nabla_2 = (\partial/\partial x, \partial/\partial y)$, and $c = 1/\sqrt{\varepsilon_0 \mu_0}$ is the speed of light in vaccum.

Suppose the interface between metal (Ω_-) and host medium (Ω_+) in x-y plane is Γ. Across the interface Γ, the following interface conditions should be satisfied:

$$[E_z]_\Gamma = 0, \tag{7}$$

$$[H_z]_\Gamma = 0, \tag{8}$$

$$\left[\frac{\varepsilon_0 \varepsilon}{\Lambda} \frac{\partial E_z}{\partial n} \right]_\Gamma + \left[\frac{k_z}{\omega \Lambda} \frac{\partial H_z}{\partial s} \right]_\Gamma = 0, \tag{9}$$

$$\left[\frac{\mu_0 \mu}{\Lambda} \frac{\partial H_z}{\partial n} \right]_\Gamma - \left[\frac{k_z}{\omega \Lambda} \frac{\partial E_z}{\partial s} \right]_\Gamma = 0, \tag{10}$$

where s is the tangent vector in the x-y plane.

3. METHOD OF INTERFACIAL OPERATOR

Because ε and μ are frequency-dependent functions in the present study, we shall solve the problem of plasmonic guided wave modes as an eigenvalue problem for k_z with given ω. In the interior of each region, the permittivity and permeability are spatially independent, thus Eqs. (5) and (6) are reduced to the Helmholtz equations with constant coefficients:

$$\left(\nabla_2^2 + \left(\frac{\omega}{c} \right)^2 \varepsilon \mu \right) E = k^2 E, \tag{11}$$

$$\left(\nabla_2^2 + \left(\frac{\omega}{c} \right)^2 \varepsilon \mu \right) H = k^2 H. \tag{12}$$

Here the sub-index z is dropped for notational simplicity. In order to express the interface conditions as an eigenvalue problem like Eqs. (11) and (12). Now we introduce two auxiliary interface variables:

$$J_E = \varepsilon_+ \frac{\partial E_+}{\partial n} - \varepsilon_- \frac{\partial E_-}{\partial n} = \left[\varepsilon \frac{\partial E}{\partial n} \right]_\Gamma, \tag{13}$$

$$J_H = \mu_+ \frac{\partial H_+}{\partial n} - \mu_- \frac{\partial H_-}{\partial n} = \left[\mu \frac{\partial H}{\partial n} \right]_\Gamma, \tag{14}$$

where $\varepsilon_\pm$ and $\mu_\pm$ are the relative permittivity and permeability functions in $\Omega_\pm$, respectively, and rearrange Eq. (9) as

$$M \left(\left[\frac{1}{\mu} \frac{\partial E}{\partial n} \right]_\Gamma + \frac{k}{\omega \varepsilon_0} \left[\frac{1}{\varepsilon \mu} \frac{\partial H}{\partial s} \right]_\Gamma \right) = k^2 J_E, \tag{15}$$

$$M \left(\left[\frac{1}{\varepsilon} \frac{\partial H}{\partial n} \right]_\Gamma - \frac{k}{\omega \mu_0} \left[\frac{1}{\varepsilon \mu} \frac{\partial E}{\partial s} \right]_\Gamma \right) = k^2 J_H, \tag{16}$$

where $M = \left(\frac{\omega}{c} \right)^2 \varepsilon_+ \varepsilon_- \mu_+ \mu_-$. Here, we have used $\left[\frac{\partial E}{\partial s} \right] = 0$, $\left[\frac{\partial H}{\partial s} \right] = 0$, which are due to the jump conditions (7) and (8). we shall call the left-hand side of Eqs. (15) and (16) the interfacial operators. It will be re-expressed in terms of E, H, J_E and J_H in the later section during discretization procedure. Eqs. (11), (12), (15), (16), together with the Bloch boundary condition, constitute a quadratic eigenvalue problem for k with unknowns (E, H, J_E, J_H), since both k and k^2 appear in Eqs. (15) and (16).

4. COUPLING INTERFACE METHOD

To accurately discretize the above equations, we adopt the coupling interface method [9]. This is a second-order accurate finite difference method under Cartesian grid, and it treats the interface as an immersed boundary. The advantages of this method include that: (1) it is a dimension-by-dimension approach thus has no limitation on the dimensionality; (2) its local truncation error is independent of the jumps of the coefficients and thus it is suitable to elliptic interface problems with large jumps in coefficients. Further, this method uses only first-order derivatives on the interface, and is thus naturally to incorporate with the interfacial operator approach, which introduces first-order derivatives as auxiliary variables. We will explain the two dimensional case here.

For two-dimensional guided wave problems, we partition the unit cell $[0, a] \times [0, a]$ into N^2 squares uniformly with mesh size $h = \frac{a}{N}$. The Cartesian grid points are $(x_i, y_j) := (ih, jh)$, $1 \leq i, j \leq N$. The abbreviation $E_{i,j} = E(x_i, y_j)$ is used. We shall also abbreviate the unknowns $(E_{1,1}, \ldots, E_{N,N})$ by $E_{1:N,1:N}$. On the interface Γ, a set of uniformly distributed grids based on arc length on Γ is adopted. They are labeled by $(\hat{x}_\ell, \hat{y}_\ell)$, $\ell = 1, \ldots, N_{\text{J}}$. At these points, the auxiliary interface variables $(J_{E,\ell}, J_{H,\ell})$ are defined.

In an interior Cartesian grid point, Eqs. (11) and (12) are discretized by the standard central finite difference method. At a Cartesian grid point which is adjacent to the interface, suppose (x_i, y_j) is the grid point at which we want to derive a finite difference equation. Let $P = (x_i + \alpha_x h, y_j)$ and $Q = (x_i, y_j + \alpha_y h)$ be the intersections of Γ and the x and y-axis from (x_i, y_j); and $R = (\hat{x}_R, \hat{y}_R) = (x_i + \alpha_{r,x} h, y_j + \alpha_{r,y} h)$ be the closest interface grid point to (x_i, y_j). Our goal is to derive a finite difference approximations for $\nabla_2^2 E\big|_{i,j}$ and $\nabla_2^2 H\big|_{i,j}$ in terms of the grid data and the interfacial variables. The discretization procedure is divided into the following steps.

4.1. Dimension-by-dimension Approach

First, we apply the previous one-dimensional method in each dimension. For instance, to derive finite difference approximation for $\frac{\partial^2 E}{\partial x^2}\big|_{i,j}$ we apply the one-dimensional method along the grid line $y = y_j$ and get

$$
\frac{\partial^2 E}{\partial x^2}\bigg|_{i,j} = [\rho_x^-, \rho_x^+] \begin{bmatrix} \beta_x + \beta_x^2 & \alpha_x (1 + 2\beta_{r,x}) \\ -(\beta_x + \beta_x^2) & -(1 + \alpha_x)(1 + 2\beta_{r,x}) \\ 0 & (1 + \beta_x)(1 + 2\beta_{r,x} - \beta_x) \\ 0 & -\beta_x (2\beta_{r,x} - \beta_x) \end{bmatrix}^t \begin{bmatrix} E_{i-1,j} \\ E_{i,j} \\ E_{i+1,j} \\ E_{i+2,j} \end{bmatrix}
$$
$$
- \frac{(\beta_x + \beta_x^2) h}{\hat{\varepsilon}_x} \left(\left[\varepsilon \frac{\partial E}{\partial x} \right]_R - \alpha_{r,y} \left[\varepsilon \frac{\partial^2 E}{\partial x \partial y} \right]_R \right) + O(h). \tag{17}
$$

where $\beta_{r,x} = 1 - \alpha_{r,x}$ and $\hat{\varepsilon}_x = (\beta_x + \beta_x^2)(\frac{1}{2} + \alpha_{r,x})\varepsilon_- + (\alpha_x + \alpha_x^2)(\frac{1}{2} + \beta_{r,x})\varepsilon_+$, $\rho_x^\pm = \frac{\varepsilon_\pm}{\hat{\varepsilon}_x}$. Similarly, in the y-direction, we can get $\frac{\partial^2 E}{\partial y^2}\big|_{i,j}$:

$$
\frac{\partial^2 E}{\partial y^2}\bigg|_{i,j} = [\rho_y^-, \rho_y^+] \begin{bmatrix} \beta_y + \beta_y^2 & \alpha_y (1 + 2\beta_{r,y}) \\ -(\beta_y + \beta_y^2) & -(1 + \alpha_y)(1 + 2\beta_{r,y}) \\ 0 & (1 + \beta_y)(1 + 2\beta_{r,y} - \beta_y) \\ 0 & -\beta_y (2\beta_{r,y} - \beta_y) \end{bmatrix}^t \begin{bmatrix} E_{i,j-1} \\ E_{i,j} \\ E_{i,j+1} \\ E_{i,j+2} \end{bmatrix}
$$
$$
- \frac{(\beta_y + \beta_y^2) h}{\hat{\varepsilon}_y} \left(\left[\varepsilon \frac{\partial E}{\partial y} \right]_R - \alpha_{r,x} \left[\varepsilon \frac{\partial^2 E}{\partial x \partial y} \right]_R \right), \tag{18}
$$

where $\beta_y = 1 - \alpha_y$, $\beta_{r,y} = 1 - \alpha_{r,y}$, and $\hat{\varepsilon}_y = (\beta_y + \beta_y^2)(\frac{1}{2} + \alpha_{r,y})\varepsilon_- + (\alpha_y + \alpha_y^2)(\frac{1}{2} + \beta_{r,y})\varepsilon_+$, $\rho_y^\pm = \frac{\varepsilon_\pm}{\hat{\varepsilon}_y}$.

4.2. Using the Interfacial Variables at Interface Grid Points

Next, we re-express the terms $\left[\varepsilon \frac{\partial E}{\partial x} \right]_R$ and $\left[\varepsilon \frac{\partial E}{\partial y} \right]_R$ in the above approximation in terms of normal and tangential derivatives of E at R:

$$
\left[\varepsilon \frac{\partial E}{\partial x} \right]_R = n_x J_{E,\ell} + s_x [\varepsilon]_R \frac{\partial E_-}{\partial s}\bigg|_R, \qquad \left[\varepsilon \frac{\partial E}{\partial y} \right]_R = n_y J_{E,\ell} + s_y [\varepsilon]_R \frac{\partial E_-}{\partial s}\bigg|_R, \tag{19}
$$

where $n = (n_x, n_y)$ and $s = (s_x, s_y)$ are the unit normal and tangential vectors of Γ at R, respectively. Here we have used $\left[\frac{\partial E}{\partial s}\right]_R = 0$.

4.3. Interpolation for the One-side Tangential Derivative

The one-side tangential derivative can be expressed by the near-by points, $\frac{\partial^2 E}{\partial x^2}\Big|_{i,j}$, $\frac{\partial^2 E}{\partial y^2}\Big|_{i,j}$ and cross second order derivatives. The cross derivatives in Eqs. (17), (18) and tangential derivatives are approximated by

$$\frac{\partial^2 E}{\partial x \partial y}\Big|_{i+1,j+1} = \frac{1}{h^2}\left(E_{i+2,j+2} + E_{i+1,j+1} - E_{i+2,j+1} - E_{i+1,j+2}\right) + O(h) \tag{20}$$

$$\frac{\partial^2 E}{\partial x \partial y}\Big|_{i,j} = \frac{1}{h^2}\left(E_{i,j} - E_{i-1,j} - E_{i,j-1} + E_{i-1,j-1}\right) + O(h). \tag{21}$$

Substitute Eqs. (19) $\sim$ (21) into Eqs. (17) and (18), we deduce a coupling equation for the second order partial derivatives $\frac{\partial^2 E}{\partial x^2}\Big|_{i,j}$ and $\frac{\partial^2 E}{\partial y^2}\Big|_{i,j}$. This coupling equation is solvable practically [10]. Finally, we get the approximations of $\frac{\partial^2 E}{\partial x^2}\Big|_{i,j}$ and $\frac{\partial^2 E}{\partial y^2}\Big|_{i,j}$ which are represented as the linear combination of near-by points and interfacial variables.

5. RESULT AND DISCUSSION

Table 1 shows the convergence result of our two dimensional method. It is a layer structure. But we enforce the periodic boundary condition in the y-direction. It shows that our method is second order in $\frac{k_z a}{2\pi}$ when $\frac{k_y a}{2\pi}$ is 0, 1, and 2. And it is slightly lower than second order accuracy when $\frac{k_y a}{2\pi}$ is 3. Lower accuracy for the eigenvalues of higher oscillatory eigenmodes is usually happened when we use a numerical method to find the eigenvalues. It also gives us the criterion to filter out some eigenmodes with bad accuracy.

Table 1: Convergence result of $\frac{k_z a}{2\pi}$ with different mesh size and wave number $\frac{k_y a}{2\pi}$. We use a two dimensional code to compute the layer structure by enforcing periodic condition in y-direction.

	$\frac{k_y a}{2\pi} = 0$		$\frac{k_y a}{2\pi} = 1$		$\frac{k_y a}{2\pi} = 2$		$\frac{k_y a}{2\pi} = 3$	
$h = 2\pi/40$	0.0427	0.0417	0.0923	0.0935	0.3405	0.3531	1.2290	1.2245
$h = 2\pi/80$	0.0063	0.0059	0.0214	0.0217	0.1057	0.1100	0.5829	0.5709
$h = 2\pi/160$	0.0009	0.0006	0.0050	0.0051	0.0289	0.0302	0.2178	0.1942
$h = 2\pi/320$	0.0001	0.0001	0.0012	0.0013	0.0074	0.0078	0.0659	0.0714
exact values of $\frac{k_z a}{2\pi}$	3.5447	3.5246	3.4007	3.3797	2.9266	2.9022	1.8880	1.8500

Figure 1 shows the dispersion relation of the cylindrical waveguides which is made of metals and dielectrics. We test two different configuations. One is a dielectric ring and the other is two dielectric rings imbeded in the metals. We found that when we add one more dielectric ring, there are one more dispersion curve with slower group velocity. It also reduces the attenuation constant of the previous one.

(a) MDM   (b) MDM   (c) MDMDM   (d) MDMDM
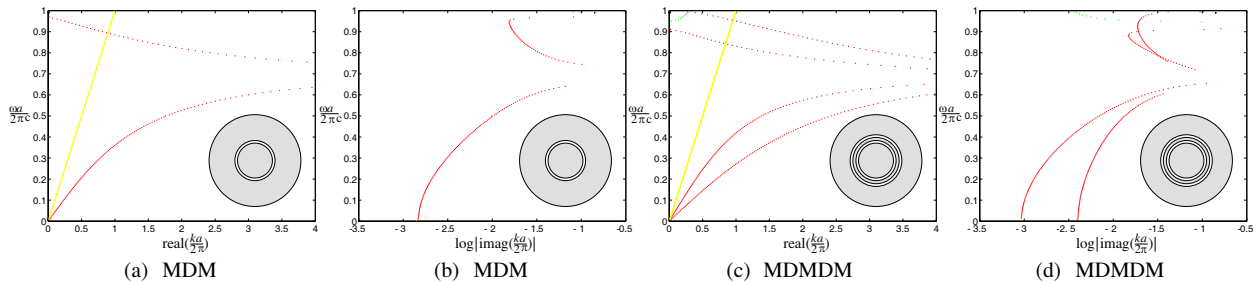
Figure 1: X: (a), (c) real$(ka/2\pi)$; (b), (d) $\log|\text{imag}(ka/2\pi)|$. Y: $\omega a/2\pi c$. We have two different configurations: MDM and MDMDM. The configuration of MDM is a dielectric ring located at $3.5 < \frac{2\pi\rho}{a} < 4$ and there is one more dielectric ring located at $4.5 < \frac{2\pi\rho}{a} < 5$ for MDMDM. We have chosen that $\frac{\omega_p a}{2\pi c} = 1$. The permittivities of dielectric and the metal are $\varepsilon_d = 1$, $\varepsilon_m = 1 - \frac{\omega_p^2}{\omega(\omega+i\omega_\tau)}$, where the damping coefficient $\frac{\omega_\tau a}{2\pi c} = 0.00296$.

Figure 2 shows some eigenmodes for split-ring structure. We selected the smoother eigenmodes which can be resolved accurately under the resolution. We found that the surface plasmons are localized on the edges of the split-ring structure.

(a) E field   (b) H field   (c) E field   (d) H field
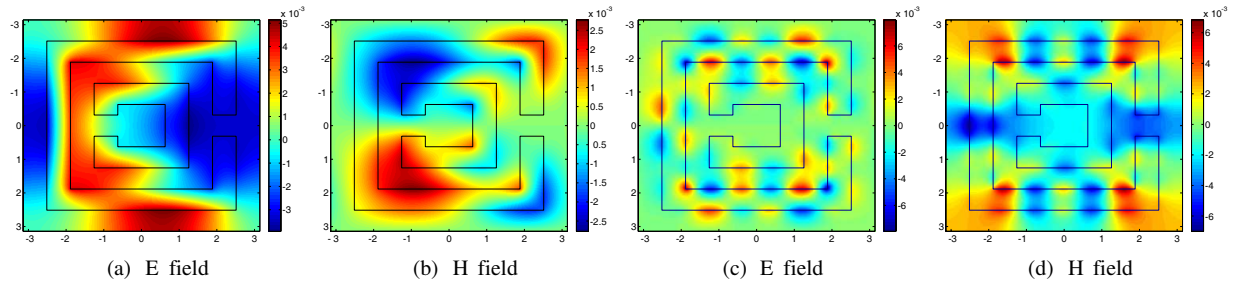
Figure 2: (a), (b): $\frac{ka}{2\pi c} = 0.7001$; (c), (d): $\frac{ka}{2\pi c} = 0.7102$. The length of the unit cell is a. The permittivities of dielectric and the metal are $\varepsilon_d = 1$, $\varepsilon_m = 1 - \frac{\omega_p^2}{\omega(\omega + i\omega_\tau)}$, where $\frac{\omega a}{2\pi c} = 0.7$ and $\frac{\omega_\tau a}{2\pi c} = 0$.

6. CONCLUSION

In the study, we presented the interfacial operator approach with the coupling interface method for solving guided wavemodes of plasmonic crystals. It is considered as an eigenvalue problem for the axial wavenumber. In this formulation, the metallic components can be very general in geometry, and are allowed to have arbitrary frequency-dependent permittivity and permeability. We use the layer structure to test the convergency of our method. The result shows that our method is second order. Two configurations of cylindrical waveguide are tested. By adding one more dielectric ring can produce a wave with slower group velocity and make the propagation length longer compared with the configuration with only one dielectric ring. Finally, it is demonstrated that the present method can resolve fine eigenmodes of the split-ring structure.

ACKNOWLEDGMENT

REFERENCES

1. Ritchie, R. H., "Plasma losses by fast electrons in thin films," *Physical Review*, Vol. 106, 874–881, 1957.
2. Barnes, W. L., A. Dereux, and T. W. Ebbesen, "Surface plasmon subwavelength optics," *Nature*, Vol. 424, 824, 2003.
3. Kuzmiak, V., A. A. Maradudin, and A. R. McGurn, "Photonic band structures of two-dimensional systems fabricated from rods of a cubic polar crystal," *Phys. Rev. B*, Vol. 55, 4298, 1997.
4. Sigalas, M. M., C. M. Soukoulis, C. T. Chan, and K. M. Ho, "Electromagnetic-wave propagation through dispersive and absorptive photonic-band-gap materials," *Phys. Rev. B*, Vol. 49, 11080, 1994.
5. Moroz, A. and A. Tip, "On-shell t-matrices in multiple scattering," *Phys. Lett. A*, Vol. 235, 195–199, 1997.
6. Ito, T. and K. Sakoda, "Photonic bands of metallic systems. II. Features of surface plasmon polaritons," *Phys. Rev. B*, Vol. 64, 045117, 2001.
7. Moreno, E., D. Erni, and C. Hafner, "Band structure computations of metallic photonic crystals with the multiple multipole method," *Phys. Rev. B*, Vol. 65, 155120, 2002.
8. Chang, C. C., R. L. Chern, C. C. Chang, and R. R. Hwang, "Interfacial operator approach to computing modes of surface plasmon polaritons for periodic structures," *Phys. Rev. B*, Vol. 72, 205112, 2005.
9. Chern, I. L. and Y. C. Shu, "A coupling interface method for elliptic interface problems," *J. Comput. Phys.*, Vol. 225, 2138–2274, 2007.
10. Chang, C. C., Y. C. Shu, and I. L. Chern, "Solving guided wave modes in plasmonic crystals," *Phys. Rev. B*, Vol. 78, 035133, 2008.

Extraction of Complex Permittivity of Multilayered Dielectric Sample Loaded in a Rectangular Waveguide

Uma Balaji

Electrical & Computer Engineering Department, California State University
Chico, CA 95929, USA

Abstract— Complex permittivity of dielectric materials loaded in the cross section of a rectangular waveguide can be determined from the S-parameter measurements of the discontinuity. The analytical electromagnetic tool used to solve the scattering problem of dielectric loaded waveguide is mode matching method. Minimization of an error function to extract the unknown complex permittivity of the material is done using a practical Quasi-Newton algorithm.

1. INTRODUCTION

The determination of permittivity of a dielectric material at high frequencies using transmission line technique is very popular. At frequencies starting from few GHz, waveguides are used as the transmission line and the dielectric sample is placed such that it completely or partially fills the cross-section of the waveguide. The S-parameter of the discontinuity is measured and the complex permittivity is extracted from this measurement by solving the inverse problem [1–3]. An error function that compares the measured and calculated S-parameters of the discontinuity is minimized in order to extract the complex permittivity. The theoretically calculated S-parameters of such a discontinuity are usually based on methods such as finite element method, moment method and mode matching method. In this paper the multilayered dielectric sample is placed in the cross section of the waveguide to form a series of discontinuities and mode matching method has been used to calculate the S-parameter of such series of discontinuities. The determination of the complex permittivity of the material filling the waveguide from the measured S-parameters is an inverse problem which is solved using a suitable optimization technique. The optimization methods are often based on genetic algorithms [1]. The optimization algorithm used in this paper is based on the practical Quasi-Newton algorithm as described in [4]. The error function that has been minimized is as given in [1].

2. THEORY

The method to calculate the theoretical value of S-parameter of the discontinuity from an empty waveguide to a single layer of dielectric filled waveguide as in Fig. 1 using mode matching method is discussed first. The analysis of such a discontinuity using mode matching method involves the following steps. The fields on both sides of the discontinuity are expanded in terms of a series of modes of incident and reflected waves. The magnitude of power carried by each of the modes is set to unity. The continuity conditions for the tangential components of electric and magnetic fields are imposed. Using the principle of orthogonality of modes, the equations of continuity conditions are transformed into matrices relating the expansion coefficients of incident and reflected waves at the discontinuity. The matrices are rearranged and inverted suitably to obtain the generalized scattering matrix which describes the discontinuity in terms of the dominant and higher order modes. Theoretically the generalized scattering matrix is of infinite dimension corresponding to the infinite number of modes. The matrix is truncated to a finite size for numerical computations. However for a discontinuity as in Fig. 1, it is sufficient to match only the fields of the dominant mode alone in order to obtain the S-parameters as the higher order modes do not couple with the fundamental mode. The computation of S-parameters from this method is given in the appendix.

In order to determine the S-parameters of multilayered dielectric sample, the following procedure is adopted. The discontinuity at each interface between two consecutive layers of the dielectric sample is analyzed using the mode matching method independently. While analyzing the discontinuity at the interface of two consecutive layers it is assumed that the material is ideal and hence the permittivity is real. However since each layer is of finite length, the S-parameter of the discontinuity from one layer to another layer of finite length is calculated considering the loss tangent and relative permittivity of that layer. This leads to the S-parameters of discontinuities at the interface of all the layers of multilayered dielectric sample including the losses that take place in all the layers of dielectric loaded waveguide. In order to incorporate the fact that the final layer of the
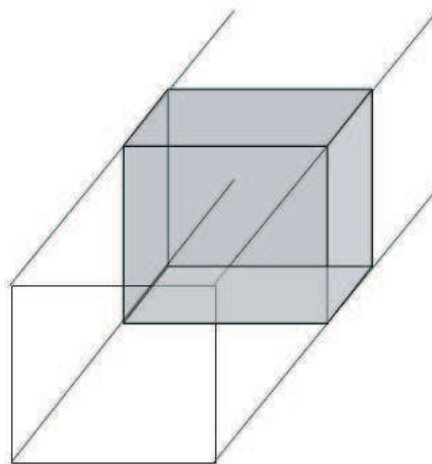
Figure 1: Dielectric sample in cross-section of waveguide.

sample ends in an empty section of waveguide, the S-parameters of a discontinuity from the final layer of dielectric sample to an empty waveguide is calculated. These S-parameters are all cascaded according to the placement of each layer to obtain the theoretical S-parameter of the multilayered discontinuity.

The inverse problem of extraction of complex permittivity is based on an initial guess for its value in the above algorithm that evaluates S-parameters of the discontinuity during optimization process. An error function that compares the theoretically computed S-parameter to the measured S-parameter is minimized during the optimization. The evaluation of S-parameter of the discontinuity is done over a narrow range of frequencies.

The S-parameter of such a structure with multilayered dielectric sample of unknown complex permittivity placed in the waveguide section is measured over the desired narrow range of frequencies. The error function E that is minimized by the optimization algorithm is calculated at these N discrete frequency points and is given by the function below.

$$E = \frac{1}{N} \sum_{i=1}^{N} \left\{ \begin{array}{l} (\mathrm{Re}[S_{11}]_{c,i} - \mathrm{Re}[S_{11}]_{m,i})^2 + (\mathrm{Re}[S_{12}]_{c,i} - \mathrm{Re}[S_{12}]_{m,i})^2 \\ +(\mathrm{Im}[S_{11}]_{c,i} - \mathrm{Im}[S_{11}]_{m,i})^2 + (\mathrm{Im}[S_{12}]_{c,i} - \mathrm{Im}[S_{12}]_{m,i})^2 \end{array} \right\}$$

The subscript c in the S-parameters indicates calculated values and the subscript m denotes the measured values. Minimization of the error function is performed using a practical Quasi-Newton algorithm [4].

3. RESULTS

A program to analyze the discontinuity of dielectric loaded waveguide was developed and tested based on the unitary property of the S-matrix when the losses in the network are zero. In order to do this the loss tangent of the dielectric sample was set to zero. The optimization algorithm was verified by finding the minimum of the Rosenbrock function. The minimum of this function which occurs at [1 1] was evaluated in 32 iteration by the program when a choice of initial guess was $[-1 \ -1]$.

The S-parameter of finite length of 2 mm discontinuity with teflon as dielectric sample of relative permittivity 2.2 and loss tangent 0.0009 in the WR 90 waveguide as in Fig. 1 was evaluated at eleven frequency points between 10–10.1 GHz and set as the measured value. The optimization algorithm was run with an initial guess of relative permittivity of 1.1 and loss tangent 1e-6. A relative permittivity of 2.2 and loss tangent of 0.0009 was obtained in 15 iterations. This is shown in Fig. 2 and Fig. 3.

The S-parameter of two layers of dielectric sample in the WR 90 waveguide of relative permittivity 2.2 and loss tangent 0.0009 followed by relative permittivity 3.5 and loss tangent 0.0001 was evaluated at eleven frequency points between 10–10.1 GHz and set as the measured value. The optimization algorithm was run with an initial guess of relative permittivity of 1.1 and loss tangent 9e-5 for the first layer and 1.1 and 1e-5 for the second layer. Convergence to the exact value was obtained in 25 iterations. This is shown in Fig. 4 and Fig. 5.
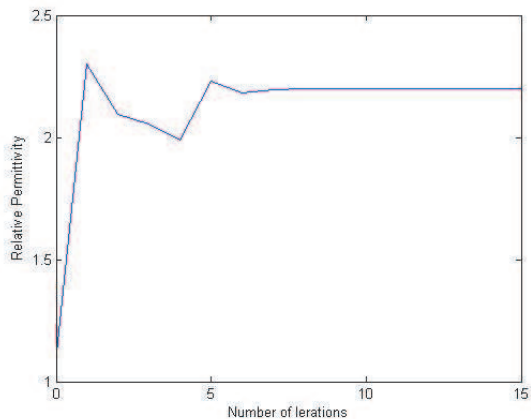
Figure 2: Convergence of relative permittivity of dielectric sample in Fig. 1.
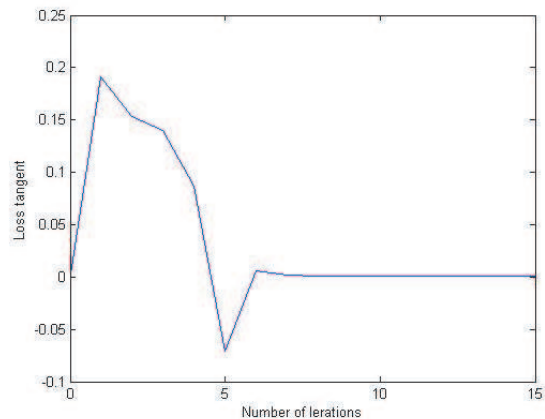
Figure 3: Convergence of loss tangent of dielectric sample in Fig. 1.
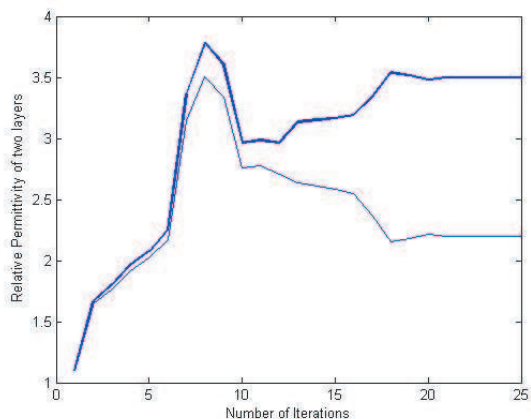
Figure 4: Convergence of relative permittivity for a two layered dielectric sample loaded in a rectangular waveguide.
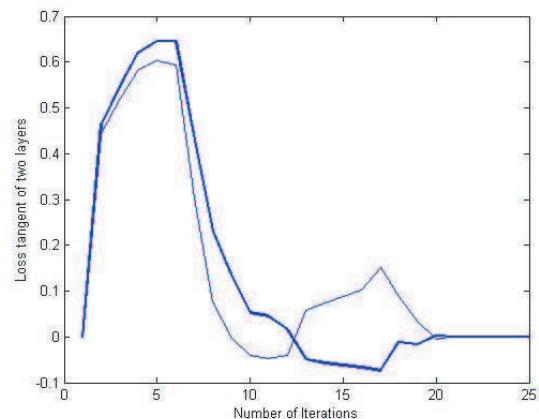
Figure 5: Convergence of loss tangent for a two layered dielectric sample loaded in a rectangular waveguide.

4. CONCLUSION

This paper has described a technique to determine complex permittivity of multilayered dielectric material. The optimization algorithm used in this paper has been found to give excellent results with the error function that has been used.

Appendix

The determination of S-matrix of discontinuity from empty waveguide to dielectric filled waveguide is given below.

$$L = \sqrt{\frac{k_{zm}^{I}}{k_{zi}^{II}}}$$

In above, k_{zm}^{I} is phase constant of fundamental mode in the empty waveguide and k_{zi}^{II} is the phase constants of the fundamental mode in dielectric filled waveguide.

The S-parameters of the discontinuity is obtained from

$$S_{11} = (LL + 1)^{-1}(LL - 1); \qquad S_{12} = 2(LL + 1)L$$
$$S_{21} = L(1 - S_{11}); \qquad S_{22} = 1 - LS_{12}$$

For determining the scattering matrix of the discontinuity of finite length, say l, of the dielectric sample with complex propagation constant γ of the fundamental mode, in an empty waveguide is

determined from the following equations:

$$R = [1 - S_{22}DS_{22}D]^{-1}$$
$$D = e^{-\gamma z}$$
$$S_{011} = S_{11} + S_{12}DRS_{22}DS_{21} = S_{022}$$
$$S_{012} = S_{12}DRS_{21} = S_{021}$$

REFERENCES

1. Faircloth, D. L., et al., "Complex constitutive parameter extraction for multilayered samples using S-parameter waveguide measurement," *Proceedings of IEEE Antennas and Propagation Symposium*, 4833–4835, 2006.
2. Barba, P., et al., "S-parameter extraction of off-set material samples in a waveguide," *Proceedings of IEEE Antennas and Propagation Symposium*, 4829–4832, 2006.
3. Balaji, U., "Complex permittivity extraction of dielectric samples in waveguide," *Proceedings of International Conference on Microwave and Millimeter Wave Technology, (ICMMT)*, Nanjing, China, April 21–24, 2008.
4. Antoniou, A., *Digital Filters, Design and Applications*, Second edition, McGraw-Hill, 1993.

Parallel Electric Field Integral Equation Solver for Arbitrary Shaped Conducting Bodies

Haythem H. Abdullah[1], **Jungsook Yang**[2],
Nader Bagherzadeh[2], **and Khalid F. Hussein**[1]
[1]Electronics Research Institute (ERI), Dokki, Cairo, Egypt
[2]University of California, Irvine, USA

Abstract— Parallel processing using distributed computers becomes important as computationally demanding applications in electromagnetic fields require highly efficient yet cost effective solutions. In this paper, we describe the development of parallel processing strategies of the Method of Moment applicable to a variety of electromagnetic problems such as scattering problems. A linear system of equations for the surface currents induced on a conducting scatterer is derived using the Method of Moment in conjunction with the RWG basis function. The Conjugate Gradient method is used to solve the linear equation. To reduce the total execution time of this procedure, we use distributed computers to process the most time consuming tasks, which are building the Z matrix and finding the surface currents by the Conjugate Gradient method. Our parallelization scheme is tested and run up to 10 distributed computers and the Message Passing Interface (MPI) is used for communicating among the nodes. From our experiment, we achieve more than 8-fold gain using 10 processors for solving the EFIE in parallel and prove the efficiency of our strategy.

1. INTRODUCTION

With the advent of high performance distributed computing, electromagnetic researchers have been quick to take advantage of increasing CPU performance per cost for solving the challenging problems. Among those computationally demanding applications are Radar Cross Section (RCS) computation of electrically large objects such as aircrafts [1, 2] and Specific Absorption Rate (SAR) inside a human body [3–6]. An appropriate method that can deal efficiently with such electromagnetic problems of conducting surfaces is the Electric Field Integral Equation (EFIE) technique that uses the Method of Moment [7, 8]. In an EFIE, an integral equation is formulated for the unknown currents on the scattering surface. It is then converted to a linear system of equations, which can be solved using well-known numerical techniques [9]. Once the currents distribution on the scatterer surface is obtained, the far field calculations can be processed in terms of the currents distribution on the scatterer surface. Triangular patch model, which is suitable for representing an arbitrarily shaped surface, is often used for the scatterer surface approximation.

For solving an EFIE of realistic size, run times on the order of hours, weeks, or longer are common. In this paper, a parallelizing scheme for solving EFIEs efficiently is discussed in detail and the performance gains that were achieved from our parallel processing experiment are shown.

2. THE PROBLEM FORMULATION

Rao-Wilton-Glisson (RWG) triangular patch basis function [7] is used to solve electromagnetic scattering problems. It is a set of basis functions having one function associated with each edge of a triangle to represent the surface currents distribution. It has the useful property of being free of line charges while being capable of modeling any arbitrarily shaped surface by the triangular patch model as shown in Fig. 1. Two triangles, T_n^+ and T_n^-, corresponding to the nth edge of the triangulated surface are depicted in Fig. 1(a). The $\pm$ sign is designated by a positive current reference direction for the nth edge. Points in $Tn^\pm$ may be denoted by $\overrightarrow{r}_n^\pm$, the position vector from the origin ($\mathcal{O}$), or $\overrightarrow{\rho}_n^\pm$, the position vector which is directed toward the free vertex $\overrightarrow{r}'$ in $T_n^\pm$. c represents the centroid of a triangle. Fig. 1(b) shows a source and observation triangle, T_n and T_m respectively, in a local coordinate system. l_k is the length of the arbitrarily labeled edge, where k is the point index of a source triangle in the local coordinate. $\overrightarrow{r}_k$ is a position vector toward a vertex of T_n, and $\overrightarrow{\rho}_k$ is toward the free vertex $\overrightarrow{r}'$ from $\overrightarrow{r}_k$. More details on the model and terms can be found in [7].

Based on this triangular patch model, an electric field integral equation is recast into an $N \times N$ system of linear equations by adopting the method of moment. $Z = [Z_{mn}]$ is an $N \times N$ matrix and $I = [I_n]$ and $V = [V_m]$ are column vectors of length N, where n and $m = 0 \ldots N$. The index n
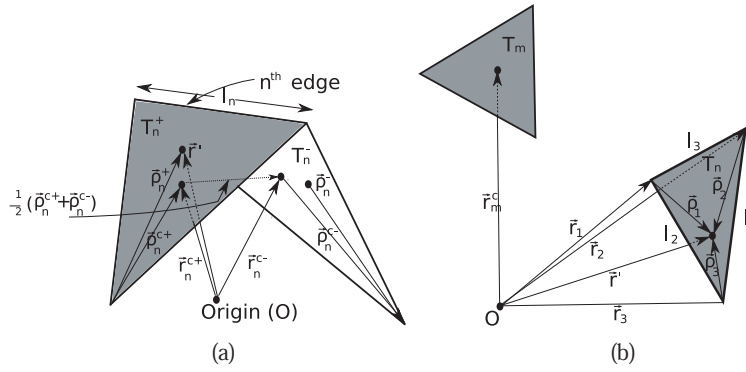
Figure 1: Triangular patch model. (a) Face pair and geometrical parameters associated with nth interior edge. (b) Local coordinates and edges for source triangle T_n with observation point $\vec{r}_m^x$ in T_m.

refers to the source edge, while the index m refers to the observation edge. Elements of Z and V with corresponding edges m and n are given by

$$[Z_{mn}][I_n] = [V_m] \tag{1}$$

$$Z_{mn} = l_m l_n \left(\frac{j\omega\mu}{4\pi} \alpha_{mn} + \frac{1}{j\omega\mu\epsilon} \beta_{mn} \right) \tag{2}$$

$$V_m = \frac{l_m}{2} \left(\overrightarrow{E}^i \left(\overrightarrow{r_m^{c+}} \right) \cdot \overrightarrow{\rho}_m^{c+} + \overrightarrow{E}^i \left(\overrightarrow{r_m^{c-}} \right) \cdot \overrightarrow{\rho}_m^{c-} \right) \tag{3}$$

where

$$\alpha_{mn} = \alpha_{m^+n^+} + \alpha_{m^+n^-} + \alpha_{m^-n^+} + \alpha_{m^-n^-} \tag{4}$$

$$\beta_{mn} = \beta_{m^+n^+} + \beta_{m^+n^-} + \beta_{m^-n^+} + \beta_{m^-n^-} \tag{5}$$

$$\alpha_{m^\pm n^\pm} = \frac{1}{4A_n^\pm} \left(\int_{T_n^\pm} G_m^\pm \left(\overrightarrow{r}' \right) \overrightarrow{\rho}_n^\pm d\acute{S} \right) \cdot \overrightarrow{\rho}_m^{c\pm} \tag{6}$$

$$\beta_{m^\pm n^\pm} = \frac{(\pm)(\pm)1}{4A_n^\pm} \left(\int_{T_n^\pm} G_m^\pm \left(\overrightarrow{r}' \right) d\acute{S} \right) \tag{7}$$

$$G_m^\pm \left(\overrightarrow{r}' \right) = \frac{e^{-jk|\overrightarrow{r_m^{c\pm}} - \overrightarrow{r}'|}}{|\overrightarrow{r_m^{c\pm}} - \overrightarrow{r}'|} \tag{8}$$

$$\overrightarrow{\rho}_n^\pm = \pm \left(\overrightarrow{r}_n^\pm - \overrightarrow{r}' \right) \tag{9}$$

$$\overrightarrow{\rho}_m^{c\pm} = \pm \left(\overrightarrow{r}_m^{c\pm} - \overrightarrow{r}' \right). \tag{10}$$

The first $\pm$ sign in Eq. (7) corresponds to the observation triangle while the second one is associated with the source triangle.

3. THE PARALLEL PROCESSING SCHEME

The time consuming nature of the application primarily results from two tasks: (1) the nested for-loops in building the Z matrix in which double integrals are evaluated; and (2) matrix-vector multiplications and vector inner products that are involved in finding the solution to the linear equation $ZI = V$. In this section, we propose the parallelization techniques for these two tasks. In (2), we use the conjugate gradient method to solve the linear equation, as it can be parallelized easily compared to direct methods. Description for parallelizing far field quantity computation is omitted because it takes very little time compared to these two tasks and the same technique that we propose in (1) can be applied for far-field parallel processing.

3.1. The Z Matrix Building Subroutine

Building the Z matrix consists of two time consuming tasks; computing the double integrals and filling the matrix. The integrals are expressed in nested loops. The outer loop goes over the observation triangles and the inner loop goes over the source triangles. As the integration process evaluated in the inner loop is independent from the outer loop's index, we divide the outer loop by the number of processors in the distributed system. Each distributed computer simultaneously executes the integrals in a designated subrange. After the task is finished, the integration result is synchronized among the distributed computers. Data synchronization refers to the status that the copies of the data in distributed computers are coherent with one another. MPI is used for exchanging messages in our implementation. After the integrals are determined, the elements of the Z matrix and vector V are evaluated using a double loop that goes over the combination of the shared interior edges. Since evaluation for each element is independent with one another, the Z matrix V and vector are divided by multiple rows as shown in Fig. 2(a) for a processor to complete one sub-block. After the partial filling on each node, they are synchronized. That is, all participating nodes obtain the same complete results of Z and V.

3.2. The Currents Subroutine

Once the elements of the Z matrix and vector V are determined, the Conjugate Gradient (CG) method [9] is used to solve the linear equation Eq. (1) for the unknown column vector I. The CG is one of the fastest and accurate methods for solving system of linear equations. Algorithm 1 describes the CG. Given the input $[Z_{mn}]$, $[V_n]$ and an error tolerance ϵ, it takes an iterative approach and finds successive approximations to the solution starting from the initial guess $[V_n]$. For each iteration, the error is evaluated and checked if $error > \epsilon$, and if the test results in true, the procedure terminates resulting in in which holds the currents distribution.

Algorithm 1 Conjugate Gradient Algorithm

1: /* Input: $[Z_{mn}]$, $[V_n]$, an error tolerance ϵ. Output: $[I_n]$ */

2: /* $m = 0 \ldots N$, $n = 0 \ldots N$, where N is the number of shared edges. */

3: $[I_n] \Leftarrow [V_n]$ /* initial guess */

4: $[P_n] \Leftarrow [Z_{mn}][I_n]$ /* matrix-vector multiplication */

5: $[r_n] \Leftarrow [V_n] - [P_n]$ /* residue */

6: $[oldr_n] \Leftarrow [r_n]$ /* old residue */

7: $error \Leftarrow max\_error$

8: **while** $error > \epsilon$ **do**

9: $[P_n] \Leftarrow [Z_{mn}][oldr_n]$

10: $rr \Leftarrow [r_n]^T \cdot [r_n]$ /* inner products */

11: $\alpha \Leftarrow \frac{[P_n]^T \cdot [oldr_n]}{rr}$

12: $[I_n] \Leftarrow [I_n] + \alpha[oldr_n]$

13: $[r_n] \Leftarrow [r_n] - \alpha[P_n]$

14: $\beta \Leftarrow \frac{[r_n]^T \cdot [r_n]}{rr}$

15: $[oldr_n] \Leftarrow [r_n] + \beta[oldr_n]$

16: $error \Leftarrow \alpha \left| max \left(\frac{[P_n]}{[I_n]} \right) \right|$

17: **end while**

18: /* The solution of the currents distribution exists in $[I_n]$ */

We parallelized the matrix-vector multiplications, vector inner products and error evaluation in the CG algorithm to reach the solution faster. As shown in Fig. 2, we assign sub-tasks to the participating computers. The vector $[I_n]$, $[P_n]$, $[r_n]$, and $[oldr_n]$ are divided into NP number of blocks where NP is the number of processors. Similarly, the rows of $[Z_{mn}]$ matrix are divided by NP blocks and each processor updates the elements of the assigned block. Inside the *while* loop, between line 8 and 9, $[oldr_n]$ is synchronized as the whole vector is needed for the following matrix-vector multiplication. For evaluating α, the vector inner products, $[r_n]^T \cdot [r_n]$ and $[P_n]^T \cdot [oldr_n]$,

are calculated for each processor regarding the assigned index range as shown in Fig. 2(b). The two partial inner products are collected and summed using the MPI messages at a manager node. It calculates α value using the expressions as shown in line 11 and synchronized by broadcasting. β evaluation is parallelized and synchronized the same way as α computation. After line 16, error is evaluated based on the partial data that each processor has, and a manager node collects them as shown in Fig. 2(c) and broadcasts the max *error* only.
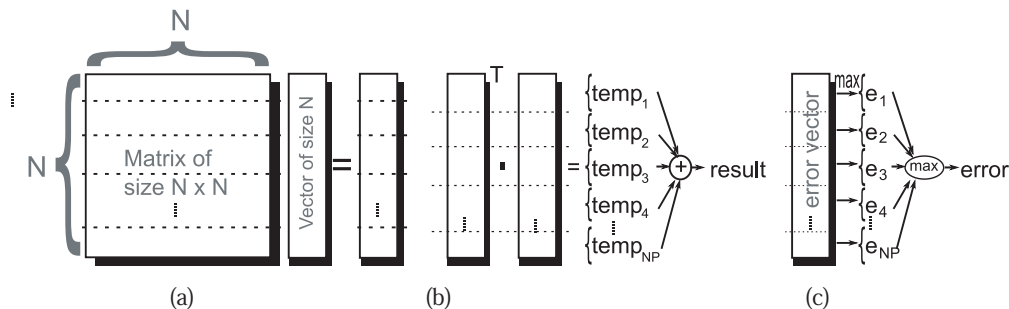

Figure 2: Data distribution schemes for parallel processing. (a) A matrix and vectors are divided by rows for parallel matrix-vector multiplication in distributed computers. (b) Vectors are divided by rows for parallel calculation of inner products. NP: Number of Processors. (c) Max error is collected from distributed computers.

4. RESULTS AND DISCUSSION

We use a distributed computing system that consists of 10 PCs. Each node has a 2.5 GHz Intel Pentium 4 CPU with 512 MB of memory. The nodes are connected by Gigabit Ethernet links. Our parallelization program is written in C++ using the MPI library. For our experiment, the calculations of Radar Cross Section (RCS) from a conducting sphere of radius 0.68λ is used. Two different surface discretization densities are used for the test. The *low-discr* uses 104 triangular patches and contains 152 shared edges, while the *high-discr* uses 1980 triangular patches and contains 2958 shared edges as shown in Table 1.

Table 1: Parameters of the triangulated surface model.

Discretization Case	Number of Patches	Number of Shared Edges
low-discr	104	152
high-discr	1980	2958

Figures 3(a) and 3(c) show the execution times of the Z matrix filling, currents, far-field subroutines measured using the two discretization modes respectively. The number of distributed computers is varied from 1 to 10. Initially, as the number of processors increases the efficiency of parallel processing clearly appears. As the number goes higher, the advantage of adding one more node becomes not as significant as seen earlier. Parallel processing appears to be more beneficial for the *high-discr*. Because it has higher computational loads compared to the *low-discr*, the gain from parallel processing far exceeds the communication overheads. Figs. 3(b) and 3(d) show the speed-up from parallel processing. The currents subroutine performs well on the *high-discr* as the heavy computational loads iterating multiple steps are efficiently parallelized, and achieves almost 10-fold speed-up using 10 nodes. The Z matrix subroutine shows maximum 7-fold speed-up on 10 nodes. The execution time of the far-field subroutine is much less than the Z matrix and currents subroutines nevertheless maximum 7-fold speed-up is gained. Parallel processing also benefits the *low-discr* but the average speed-up is lower than the *high-discr* as expected. The rough transition in the speed-up graph of *low-discr* shows that the communication time affects the total execution time more compared to the *high-discr*. Overall, the *high-discr* which has more computational loads takes advantage of parallel processing than the *low-discr*, and it achieves maximum 8-fold of speed-up gain.
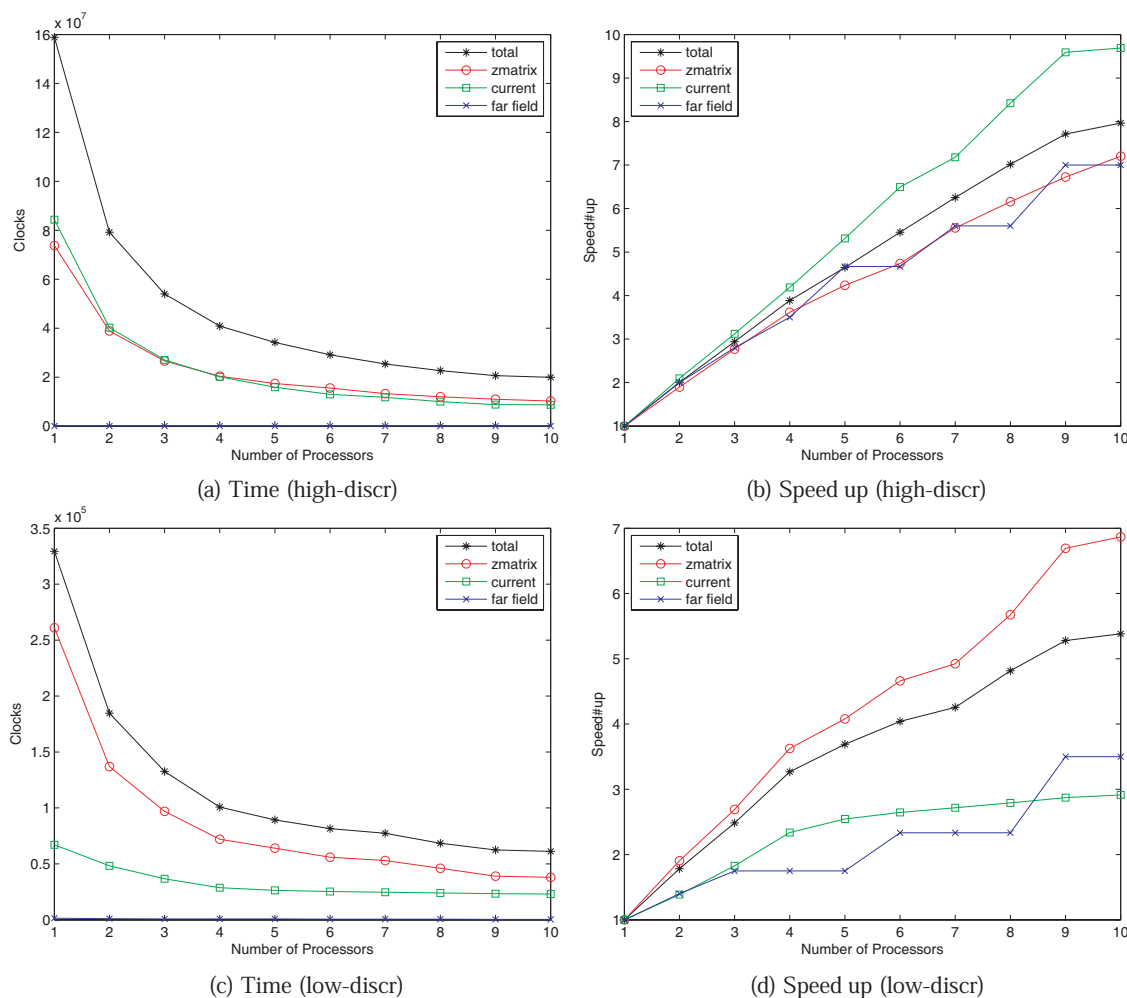
(a) Time (high-discr)

(b) Speed up (high-discr)

(c) Time (low-discr)

(d) Speed up (low-discr)

Figure 3: Execution time and speed-up.

5. CONCLUSION

In this paper, we demonstrated the high performance parallel EFIE solver on a distributed computing system by using MPI library. We adopted the RWG basis function in conjunction with the Method of Moment to make a linear system of equations for the surface current induced on a conducting scatterer. Parallelization strategies for building the Z matrix and solving the linear equation $ZI = V$ to obtain the currents. The execution time and speed-up obtained by running our program in a distributed multiprocessor computing environment were shown and analyzed. In conclusion, we achieved an efficient speed-up to solve the electromagnetic scattering problems, and the proposed parallelization technique can be applicable to general problems formulated by the Method of Moment.

REFERENCES

1. Zhang, G. H. and M. Y. Xia, "Parallel analysis of transient scattering by 3D conducting objects using time domain integral equation method," *Microwave, Antenna, Propagation and EMC Technologies for Wireless Communications, International Symposium*, 970–973, August 2007.

2. Zhang, G. H. and M. Y. Xia, "Parallel analysis of transient scattering by lossless dielectric bodies using time domain integral equation method," *Microwave Conference APMC*, 1–4, Asia Pacific, December 2007.

3. Tinniswood, A. D., P. S. Excell, M. Hargreaves, S. Whittle, and D. Spicer, "Parallel computation of large-scale FDTD problems," *Computation in Electromagnetic, Third International Conference on (Conf. publ. No. 420)*, 7–12, April 1996.

4. Tinniswood, A. D., C. M. Furse, and O. P. Ghandhi, "Computations of SAR distribution for two anatomically based models of the human head using CAD files of commercial tele-

phones and the parallelized FDTD code," *Antennas and Propagation Society International Symposium, Digest*, Vol. 2, 982–985, July 1997.

5. Eldeep H., H. Elsadek, H. H. Abdallah, M. Desouky, and N. Bagherzadeh, "FDTD accelerator for calculating SAR distribution inside human head due to radiation from wireless devices," *EMTS 2007 International URSI Commission B-Electromagnetic Theory Symposium*, 26–28, Ottawa, ON, Canada, July 2007.

6. Elsadek, H., H. Eldeeb, H. Abdallah, M. Desouky, and N. Bagherzadeh, "Specific absorption rate calculation using parallel finite difference time domain," *WORLDCOMP'08*, Las Vegas, July 2008.

7. Rao, S. M., D. R. Wilton, and A. W. Glisson, "Electromagnetic scattering by surfaces of arbitrary shape," *IEEE Trans. Antennas Propagat.*, Vol. 30, No. 5, 409–418, 1982.

8. Hussein, K. F. A., "Fast computational algorithm for EFIE applied to arbitrarily-shaped conducting surfaces," *Progress In Electromagnetics Research*, PIER 68, 339–357, 2007.

9. Hestenes, M. R. and E. Stiefel, "Methods of conjugate gradients for solving linear systems," *Journal of Research of the National Bureau of Standards*, Vol. 49, 409–436, 1952.

Improvement of Particle Swarm Optimization

K. Kawakami and Z. Meng

Department of Electrical Engineering, Fukuoka University, Japan

Abstract— A new technique titled "Particle Refresh" and a hybridization with conjugate gradient method are introduced to particle swarm optimization (PSO). The former charges power to inactive particle to improve the recovery ability of PSO after trapping on a local solution, and as a result, it becomes easy to choose suitable values for control-parameters to keep high performance for diverse objective functions. On the other hand, the point of the latter is how to determine the changeover timing between a conjugate gradient method and a PSO algorithm.

1. INTRODUCTION

There are cases when an optimization is required in the field of electromagnetic engineering, such as optimizing design parameters, resolving a nonlinear inverse scattering problem as an optimization problem, and so on. Usually, an objective function has multiple local solutions due to the function complexity or errors in approximately calculation, and a global applicable optimization method is necessary. Particle swarm optimization (PSO) [1] has achieved considerable success as a global optimization method with a wide range of applicability, requiring no prior information, being appropriate for diverse objective functions, and offering the ability of recovery even after trapping on a local solution. However, PSO has the following problems.

① Difficult to use

  Several parameters are needed for controlling the algorithm. Unless these parameters are set appropriately, search efficiency drops significantly. There are, however, no clear rules for setting the parameters, and almost all users have considerable difficulty in setting them without prior experience in parameter tuning.

② Slow convergence

  Global optimization methods do not utilize differential information, so that the convergence is usually slower than that of a gradient descent method.

In this paper, a new technique "Particle Refresh" and a hybridization with a conjugate gradient method are introduced to improve the conventional PSO algorithm.

2. THE ALGORITHM OF PARTICLE SWARM OPTIMIZATION

A PSO algorithm employs a swarm composed of multiple particles, and estimates each particle at each moment t ($t = 0, 1, 2, \ldots$) by using the value of the objective function at the particle's current position $\boldsymbol{x}_i(t) \in D$, where i and D denote the number of the particle and search space, respectively. Each particle has its velocity $\boldsymbol{v}_i(t)$. The velocity at moment t depends on the velocity at moment $(t-1)$, the position $\boldsymbol{x}_i(t)$, the best position $\boldsymbol{p}_i(t)$ found up to now by the article, and the best position $\boldsymbol{g}(t)$ found up to now by the swarm:

$$\boldsymbol{v}_i(t) = c_1 \boldsymbol{v}_i(t-1) + c_2 \left\{ \boldsymbol{p}_i(t) - \boldsymbol{x}_i(t) \right\} + c_3 \left\{ \boldsymbol{g}(t)\boldsymbol{x}_i(t) \right\}. \tag{1}$$

Here c_1 is a constant, c_2 and c_3 is selected randomly at the moment by

$$c_k = c_{\max} alea(0, 1), \quad k = 2, 3, \tag{2}$$

where $alea(0, 1)$ denotes a random number uniformly distributed in $[0, 1]$. Then the position of particle at moment $(t + 1)$ is defined by

$$\boldsymbol{x}_i(t + 1) = \boldsymbol{x}_i(t) + \boldsymbol{v}_i(t). \tag{3}$$

The algorithm of PSO can be described as follows:

① Select the size of the swarm n (how many particles), a threshold E (if possible) and a time limit T to stop the algorithm when the estimation of any particle e_i becomes $e_i \geq E$ or time t becomes $t \geq T$; let $t = 0$; randomly locate n particles in the search space and set an initial velocity to each particle according to the size of the search space.

② Get estimation e_i $(i = 1, 2, \ldots, n)$ for each particle by the objective function, and update $\boldsymbol{p}_i$ and $\boldsymbol{g}$.

③ Stop the algorithm and output $\boldsymbol{g}$ as its solution if one of the following "stop conditions" is satisfied, or go to the next step otherwise.

(a) the best estimation $e_{\text{best}} \geq E$;

(b) $t \geq T$;

③ Let $t = t + 1$ and update $\boldsymbol{v}_i$, $\boldsymbol{x}_i$ by (1), (3).

④ Go to step ②.

3. EFFECTS OF C_1 AND C_{MAX} ON ALGORITHM PERFORMANCE

Obviously, c_1 is an "inertial coefficient" of the particle. If $c_1 > 1$, a particle can power itself and the algorithm will not converge to a solution. Therefore, viewed in the light of algorithm convergence, c_1 should be smaller than one, and particles will converge to $(\boldsymbol{g} + \boldsymbol{p}_i)/2$ as a result of a damping motion. A PSO algorithm with a large c_1 turns to an advantage in recovery after trapping on a local solution and disadvantage in convergence rate. On the other hand, c_{max} is a "gravitational coefficient". The gravitation pulls particles to $\boldsymbol{g}$ or $\boldsymbol{p}$ to improve searching efficiency. However, the algorithm will diverge if c_{max} is too large.

Certainly, high performance can be obtained when the "inertia" and the "gravitation" work in cooperation with each other. Some textbooks recommend that user should select values for c_1 and c_{max} in pairs, such as "0.7,1.47", "0.8,1.62", and so on. To investigate the effects of c_1 and c_{max} minutely, two numerical experiments are considered.

The first one is to maximize the objective function given by (4):

$$\Omega = \prod_{i=1}^{2} \frac{\sin(x_i)}{(x_i)}, \quad -10 \leq x_i \leq 10, \quad i = 1, 2. \tag{4}$$

A three-dimensional graphical representation of (4) is given by Figure 1. We change c_1 from 0.1 to 0.6 and c_{max} from 1 to 2.4, and estimate the performance by the times of objective function calculated by the algorithm until the solution is found out. Because PSO is stochastic in nature and a single test does not validate the characteristic we made 1000 runs for each test of a pair values of c_1 and c_{max}, and show the average calculation times (ACT) in Figure 2. It is shown that there is a region for c_1 and c_{max} in which the PSO keeps high performance. If c_1 or c_{max} is set to a value out of the region, the algorithm will lose its performance or even fails in the searching.
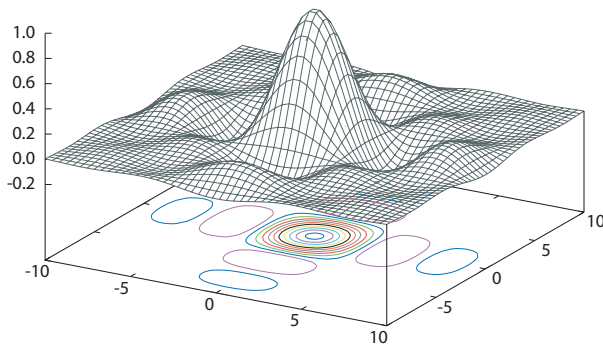
Figure 1: Three-dimentional representation of (4).
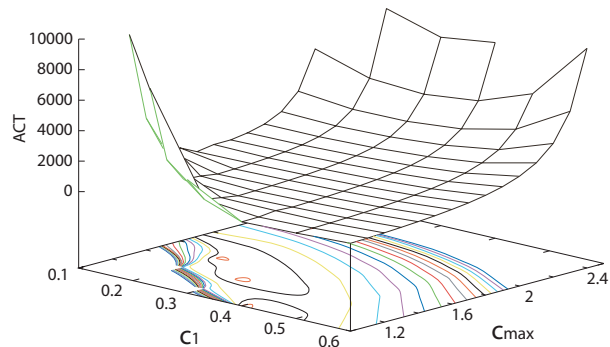
Figure 2: Estimation of the PSO applied to (4).

The second experiment is to minimize the objective function given by (5):

$$\Omega = x_1 \sin(4x_1) + x_2 \sin(2x_2), \quad 0 \leq x_i \leq 10, \quad i = 1, 2. \tag{5}$$

There are many local minima in the search space, as shown in Figure 3. Similar to that for experiment 1, we show the ACT in Figure 4. The region in which the PSO keeps high performance exists still, but becomes very narrow and almost does not double with the region shown in Figure 2 for (4). The results show that the conventional PSO may fail in the searching or lose its performance if c_1 and c_{max} are not set suitably, and the suitable values depend on the objective function. General values working well for any objective function do not exist.

Figure 3: Three-dimentional representation of (5).

Figure 4: Estimation of the PSO applied to (5).

4. REFRESH OPERATOR

Why are the suitable regions for (4) and (5) different from each other?

For an objective function like (4) that the algorithm does almost not trap on a local minimum, it becomes more important for PSO to focus the s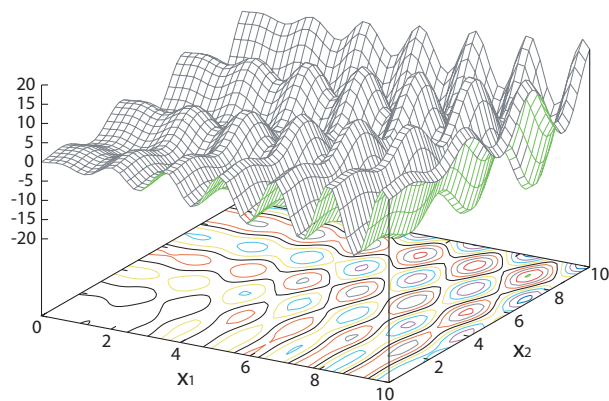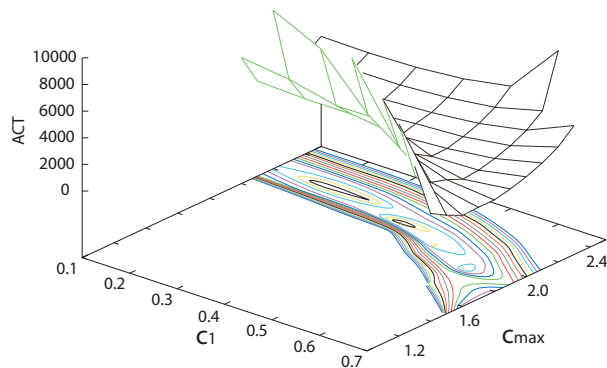earch attention to the neighborhood of the current best solution. Therefore, a "convergence-mode" where c_1 and c_{max} are small can get high performance in experiment 1.

However, if a convergence-mode PSO is applied to optimize an objective function with many local minima like (5), the algorithm will trap on a local minimum easily. As a result, the algorithm may fail or lose its performance. In fact, many particles lose their power and stop at a local minimum in many cases when the c_1 and c_{max} are set to small values in experiment 2. If all of the particles stop at a local minimum, the algorithm will fail in the searching.

Because a standstill particle is useless for the algorithm, we introduce a new operator titled "Particle Refresh" (PR) to PSO algorithm shown in Section 2 after the step ③:

③⁺ when $|v_i| < \varepsilon$, randomly re-locate the particle and clear p_i, where ε denotes a small enough value.

The operation will power standstill particles and reinforce the algorithm to avoid trapping on a local minimum.

The numerical experiments results applying a PSO with PR to (4) and (5) are shown in Figures 5 and 6.
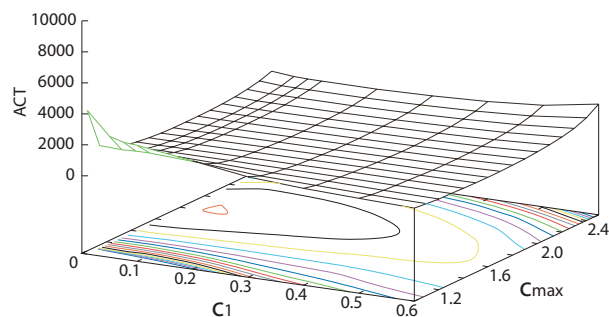
Figure 5: Estimation of the PSO with PR applied to (4).
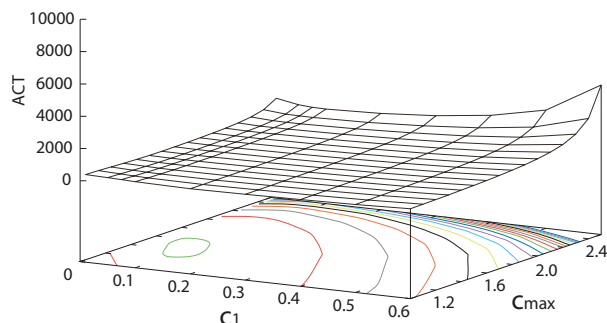

Figure 6: Estimation of the PSO with PR applied to (5).

Obviously, the suitable region of c_1 and c_{max} in Figures 5 and 6 become widely and partly double with each other, comparison with those in Figures 2 and 4. In other words, it becomes easy to choose suitable values for c_1, c_{max} to keep high performance for different type objective functions. The PSO becomes Easy-To-Use.

5. DISCUSSION ABOUT HYBRIDIZING WITH A GRADIENT METHOD

In order to improve the search performance, many hybridizations between a gradient descent method and an evolution algorithm, such as a conjugate gradient method (CGM) and a genetic algorithm, have been reported. One of the points for the hybridizations is how to determine the changeover timing between the two methods, because it is difficult to judge if the current best solution g is in the neighborhood of the global minimum. We try to hybridize a CGM (using numerically differential information) with a PSO. The algorithm is based on the PSO algorithm shown in Sections 2 and 4, but inserted the following step after the step ②.

②⁺ If one of the following conditions is satisfied, start CGM at g, update g with the solution of the CGM, and return to the PSO ③

- CGM has not been applied up to now and $g(t) = g(t-1) = g(t-2)$;
- CGM has been applied one or more than one times and $|g(t) - g(t-1)|$ is larger than an allowable margin of error;

The ACTs of 1,000 runs where the hybrid method is employed to (4) and (5) are shown in Table 1. The hybridization method shows higher performance than the PSOs in minimizing (5), but is almost not effective in maximizing (4). In fact, the above conditions for starting the CGM are almost not satisfied and the CGM had almost not been used in maximizing (4). We need a rational ground to construct the conditions for starting the CGM. A technique that can make use of the information of the objective function obtained by the algorithm at that time to roughly image the basins of attraction of local minima, as reported in [2], is necessary for a hybrid method.

Table 1: A comparison of ACT between the cases of using standard PSO, PSO with refresh operator and the hybridization method.

Objective Function	standard PSO	standard PSO + PR	standard PSO + PR + CGM
(4)	498	261	262
(5)	1521	535	365

6. CONCLUSIONS

A new operator titled "Particle Refresh" proposed in Section 4 is effective in improving particle swarm optimization (PSO) algorithm keeping high performance for different type objective functions. Hybridization between a PSO and a conjugate gradient method is tested and the results show that how to determine the changeover timing between the methods is very important. We are going to introduce a technique that can make use of the information of the objective function obtained by the algorithm to roughly image the basins of attraction of local minima, as reported in [2], to solve the problem.

REFERENCES

1. Clerc, M., *Particle Swarm Optimization*, ISTE Ltd., London, 2006.
2. Meng, Z., "Autonomous genetic algorithm for functional optimization," *Progress In Electromagnetic Research*, PIER 72, 253–268, 2007.

An Improved Adaptive Finite Element Method for the Simulation of Electromagnetic Field

Zhanghong Tang[1], **Jiansheng Yuan**[2], **and Gai Tao**[1]

[1]College of Materials Science and Engineering
Beijing University of Technology, Beijing 100124, China
[2]Department of Electrical Engineering, Tsinghua University, Beijing 100084, China

Abstract— Basing on the posteriori error estimation and Superconvergent Patch Recovery (SPR) technology, Elena Zhelezina proposed a fast adaptive mesh refinement strategy which can generate the refined mesh for the required accuracy of the solution with only one step of refining the initial mesh. However, this remeshing strategy still generates too many elements which will spend too much memory and CPU time to solve the FEM sparse matrices. In this paper, a new formula to estimate the global error is introduced, and then, a new strategy to calculate every element's refinement degree is presented. Results show that the improved method generates less than half of mesh nodes with only one step of mesh refinement comparing to Elena Zhelezina's method.

1. INTRODUCTION

Adaptive FEM mesh refinement is a very popular method in electromagnetic field computation nowadays since an accurate computation of large problems requires large number of unknowns, which cost large amount of computer resource. The adaptive FEM mesh refinement can significantly reduce the number of unknowns comparing to the traditional methods. Usually mesh refinement is controlled and governed by the posteriori local error estimation based on a post-processing recovery technique.

The idea of recovery methods is to post-processes an approximate solution and in return to get a better solution. There are two famous recovery techniques: recovery by equilibration of patches (REP) [1] and SPR [2, 3]. In this paper the latter method is used, which is based on least squares fitting locally. SPR is used to recover a gradient from the gradient of the finite element solution. Results from the applications demonstrate that it is robust and efficient. The SPR method recovers the gradient of solution (G^{spr}) with higher order approximation by smoothing the values evaluated by the finite element method (G^{fem}) at Gauss points (the so-called superconvergent points) of a patch of elements [2, 3].

Basing on the recovered or smoothed gradient G^{spr} by the SPR and evaluated gradient G^{fem} by the FEM, the indicator of the relative error of every element is defined as

$$\varepsilon_{s0}^e = e_e / \left\| G^{\mathrm{spr}} \right\|_2 = \left\| G^{\mathrm{fem}} - G^{\mathrm{spr}} \right\|_{2(e)} / \left\| G^{\mathrm{spr}} \right\|_2 . \tag{1}$$

and the global indicator of the error ε_s is computed as

$$\varepsilon_s = E_{\mathrm{total}} / \left\| G^{\mathrm{spr}} \right\|_2 = \left\| G^{\mathrm{fem}} - G^{\mathrm{spr}} \right\|_2 / \left\| G^{\mathrm{spr}} \right\|_2 . \tag{2}$$

By the adaptive analysis the relative error of the final solution gradient is expected to satisfies the inequality

$$\varepsilon_s < \varepsilon_0. \tag{3}$$

where ε_0 is a specified tolerance.

Suppose the error is equidistributed on the domain of calculation. If N_{elem} is the total number of elements, the permissible value of error of every element is

$$\varepsilon_{s\,\mathrm{max}}^e = \varepsilon_0 S_e / S. \tag{4}$$

where S_e is the area of an element and S is the total area of all elements.

Conventional adaptive refinement methods usually check the elements to be refined by comparing Equations (1) and (4). The elements which are selected will be refined by the same way and the field is solved again basing on the refined mesh and a step of iteration is done. The iteration will not

terminate until the relative error of the solution satisfies the inequality (3). Usually the conventional methods are slow because large number of iterations is required for the expected relative error; as a result, very large CPU time is needed because of large scaled sparse matrices solution.

Recently, Elena Zhelezina proposed a fast adaptive refinement method basing on error estimation which can construct the final refined mesh for the required accuracy with only one step of refinement [4]. The idea of this method is to calculate refinement degree of every element according to the estimated errors.

2. ELENA ZHELEZINA'S FAST ADAPTIVE REFINEMENT STRATEGY

Elena Zhelezina's fast adaptive refinement method calculates the permissible value of error of every element by (4) at first, and then, according to (1) and (4) the following inequality should be satisfied for every element

$$\varepsilon_{s0}^e \le \varepsilon_{s\,\max}^e. \tag{5}$$

or

$$\varepsilon_s^e = \left(\left\| G^{\mathrm{fem}} - G^{\mathrm{spr}} \right\|_{2(e)} / S_e \right) \left(S / \left\| G^{\mathrm{spr}} \right\|_2 \right) \le \varepsilon_0. \tag{6}$$

where ε_s^e is the indicator of the error of elements.

When the estimated global error ε_s is greater than the given tolerance ε, the mesh will be refined and the problem will be recalculated basing on the new mesh. In order to achieve an optimal adaptive mesh the element indicator ε_s^e is calculated for each finite element iteration. The elements, for which inequality (6) is not satisfied, are marked for the refinement. Elena Zhelezina suggested that the elements with larger error indicator should be refined more times. Furthermore, he defined elements' refinement degree (or refinement times, the number of refinement loops for the element) as follows:

If NC denotes the number of elements after one refinement of the element, the so-called child-element, and NI denotes the refine degree of element, after the whole refinement process the number of child-elements for the element is $(NC)^{NI}$. Basing on the suppose that the error is equidistributed on the domain, the error indicator of child-element can be approximated as

$$\varepsilon_s^{e-sub} \approx \varepsilon_s^e / (NC)^{NI}. \tag{7}$$

Since the tolerance for the child-element is ε_0 as well, the following estimation can be satisfied

$$\varepsilon_s^e / (NC)^{NI} \approx \varepsilon_0. \tag{8}$$

Therefore, the element's refinement degree can be predicted as

$$NI = [\log(\varepsilon_s^e / \varepsilon_0) / \log(NC)] + 1. \tag{9}$$

where [.] denotes an operator which takes the integer part of the real number.

By applying Equation (9), this method formed a "refinement map" for the initial coarse mesh with each element being assigned with the refinement degree. Computations on such a refined mesh produce results which satisfy the specified tolerance ε_0.

However, for some problems in which the distribution of field changes violently, the refinement degree of element predicted by (9) is very large and the difference of refinement degrees between two neighboring elements is large. Thus, many times of elements will be generated which leads to the solution of refined mesh become much more difficult — for some large problems, the matrix can't be solved on 32 bit machines.

3. IMPROVEMENT TO THE FAST MESH REFINEMENT STRATEGY

The fast mesh refinement strategy proposed by Elena Zhelezina is very efficient in most cases, since it reaches the satisfied solution after only one step iteration. However, results show that the refinement degree predicted by (9) is too large and too dense mesh with large number of elements will be generated. This strategy calculate the refinement degree for all elements which don't satisfy inequality (6) without updating the global error ε_s dynamically and as a result, after refinement the global error ε_s is much less than the specified tolerance ε_0, i.e., the final global error ε_s over-satisfies inequality (3). In this paper, a dynamic predicting strategy was presented basing on Elena Zhelezina's error estimation and assumption. The dynamic predicting strategy doesn't calculate

element's refinement degrees by (9), but increases elements' refinement degrees one by one and accordingly, the global error ε_s is updated. The predicting process will not be terminated until the global error ε_s satisfies inequality (3). Results show that the improved strategy decreases the number of refined elements and the CPU time of the whole calculation greatly.

According to (2) and (3), by the adaptive analysis the global error of the final solution gradient is expected to satisfies the inequality

$$E^2_{\text{total}} < \left(\varepsilon_0 \|G^{\text{spr}}\|_2\right)^2. \tag{10}$$

On the other hand, the global error can be expressed by the error of elements

$$E^2_{\text{total}} = \sum_e e^2_e. \tag{11}$$

According to (7), if the refinement degree of element k is NI, k, its error will reduce $(NC)^{NI,k}$ times, i.e.,

$$e'_k = e_k / NC^{NI,k}. \tag{12}$$

Therefore, after the refinement of element k, the global error changes to

$$E^2_{\text{total}} = \sum_e e^2_e - \left(e_k \left(1 - 1/NC^{NI,k}\right)\right)^2. \tag{13}$$

The dynamic predicting strategy presented in this paper increases the refinement degrees of elements one by one. Every time the refinement order of an element changes, the global error ε_s is also updated according to (13). The process is keeping doing until the global error satisfies the excepted tolerance.

The steps of dynamic predicting strategy can be described as follows:

1) initialize the NI of every element to zero;
2) sort the estimated error e_k of every element, larger is in front;
3) set the element which is in front of the sorted list to be the current element;
4) if the global error doesn't satisfy (10), increase the refinement degree of every element one at a time basing on the sequence sorted in step 2) and update the current element's error e_k according to (12) and the global error according to (13), else terminate the process;
5) move the current element to the next element in the sorted list and if the error of the current element satisfy (6), go to step 2), else go to step 4).

According to the dynamic predicting strategy, a similar "refinement map" is formed and then the solution is done basing on the refined mesh. Before forming and solving the FEM matrix, the refined mesh is smoothed by Laplace smooth method. Fig. 1 shows the result of an example by the adaptive refinement strategy presented in this paper. Mesh results shows that the refined mesh increased "smoothly" comparing to the initial mesh.

The dynamic predicting strategy presented in this paper increases the refinement degrees of elements one by one and changes the global error ε_s dynamically. Comparing to the strategy presented by Elena Zhelezina, the method in this paper generate much less elements and as a result, needs much less CPU time and memory to solve the problem.

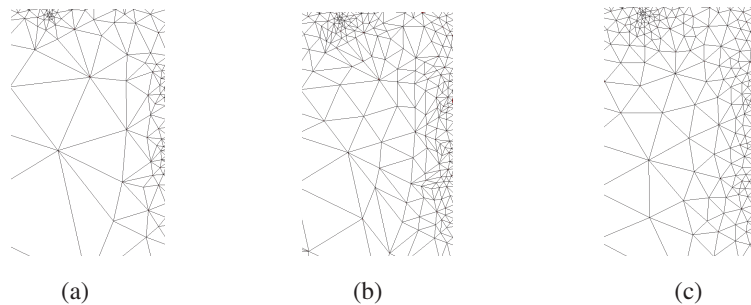| (a) | (b) | (c) |

Figure 1: Results of adaptive refinement presented in this paper, (a) initial mesh, (b) refined mesh, (c) refined mesh smoothed by Laplace smooth method.

4. RESULTS COMPARISON

Figure 2 shows an example of two-dimensional electrostatic problem tested by Elena Zhelezina. On the boundary Γ_1 and Γ_2, the potential was set to be 1 V and 0 V respectively, and for the other boundaries, the Neumann boundary conditions is set. In this example, the tolerance of relative error is set to 5%. The initial global relative error is 12% and the number of elements of initial mesh is 112. After adaptive refinement, the global relative error presented by Elena Zhelezina reaches 3.2% and the number of elements reaches 969. However, with the same initial mesh and the same tolerance, the method presented in this paper can get the solution with the global relative error 4.4% but the number of elements is only 448, which is less than half of the elements generated by Elena Zhelezina's method.
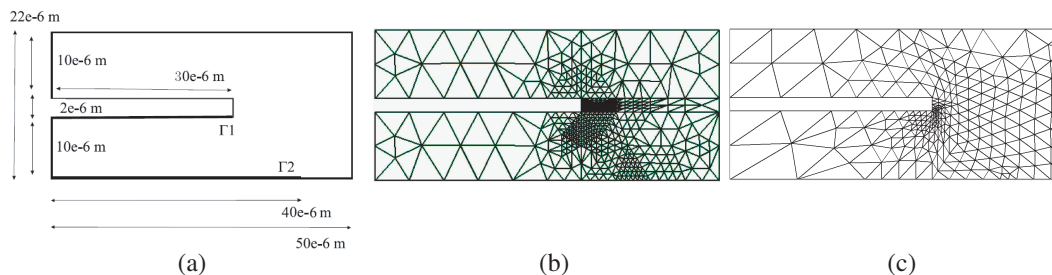

Figure 2: Results comparison of a simple example, (a) field domain, (b) adaptive refinement by Elena Zhelezina's method, (c) adaptive refinement by the method in this paper.

Furthermore, when the tolerance is set to 3.49% for the method in this paper, the global relative error reaches 3.17% (which is smaller than Elena Zhelezina's result 3.2%) but the number of elements is only 877.

Table 1 shows the adaptive mesh refinement results of three examples which calculate potential distribution on layers of integrated circuit packages (the tolerance is set to 5%). By some special technology, the three dimensional problem can be simplified to a series of two dimensional problems [5], which can be calculated by two dimensional adaptive mesh refinement. Results show that the dynamic predicting strategy presented in this paper can also reaches the excepted tolerance in only two steps, but generates much less unknowns and elements. In case 3 the number of unknowns generated by Elena Zhelezina's method is too large to be solved on 32 bit systems because of insufficient memory.

Table 1: Results comparison of adaptive refinement methods.

	Results by Elena Zhelezina's method			Results by the method in this paper		
	ε_0	Unknowns	Elements	ε_0	Unknowns	Elements
Case 1	9.08%	70,093	120,113	9.08%	70,093	120,113
	3.41%	670,311	1,257,834	5.63%	84,512	147,014
				4.70%	97,064	170,139
Case 2	12.62%	43,615	69,814	12.62%	43,615	69,814
	3.06%	408,011	766,980	5.52%	68,645	117,924
				4.52%	75,380	130,560
Case 3	17.72%	75,174	129,030	17.72%	75,174	129,030
	Insufficient Memory	12,673,238	25,181,216	5.76%	141,906	259,069
				4.55%	167,029	306,165

Table 1 also shows that the dynamic predicting strategy usually needs two steps of iterations to reach the excepted tolerance, which spends more CPU time to solve matrices. In fact, this disadvantage is easy to overcome by modifying the ε_0 in (10) to a small one, for example, $0.8 \times \varepsilon_0$.

5. CONCLUSION

The dynamic predicting strategy presented in this paper was successfully realized to improve the fast adaptive mesh refinement method introduced by Elena Zhelezina. Comparing to Elena Zhelezina's fast adaptive mesh refinement method, this method can also reach the excepted tolerance of global relative error with only one step of mesh refinement, but the method in this paper can generate much less unknowns and elements which save a lot of CPU time and memory.

ACKNOWLEDGMENT

REFERENCES

1. Boroomand, B. and O. C. Zhienkiewicz, "An improved REP recovery and the effectivity robustness test," *International Journal for Numerical Methods in Engineering*, Vol. 40, 3247–3277, 1997.
2. Zienkiewicz, O. C. and J. Z. Zhu, "The superconvergent patch recovery and a posteriori error estimates. Part 1: The recovery technique," *Int. J. Numer. Meth. Eng.*, Vol. 33, 1331–1364, 1992.
3. Zienkiewicz, O. C. and J. Z. Zhu, "The superconvergent patch recovery and a posteriori error estimates. Part 2: Error estimates and adaptivity," *Int. J. Numer. Meth. Eng.*, Vol. 33, 1365–1382, 1992.
4. Zhelezina, E., "Adaptive finite element method for the numerical simulation of electric, magnetic and acoustic fields," PhD thesis, 2005.
5. Tang, Z., T. Su, J. Fang, and J. Yuan, "A fast method for identifying parallel plate field domains in high speed PCBs and IC packages," *Microwave and Wireless Components Letters*, IEEE, Vol. 18, No. 6, 380–382, 2008.

A Domain Map Finite Element Method for Solving Open Boundary Electromagnetic Field Problem and Its Application

Zhanghong Tang[1], Yueqin Dun[2], Jiansheng Yuan[2], and Gai Tao[1]

[1]College of Materials Science and Engineering
Beijing University of Technology, Beijing 100124, China
[2]Department of Electrical Engineering, Tsinghua University, Beijing 100084, China

Abstract— A domain map finite element method was applied to the three dimensional (3-D) numerical mode-matching (NMM) method to solve the 3-D electric logging problem. This method divides the infinite field domain of every layer into inner and outer field domains, in inner domain the matrices are formed by standard FEM and in outer domain, the Kelvin transformation is applied and the matrices are formed by a new infinite element format. Matrices of two different parts are coupled in the domains' common interface to form the final matrices of the original problem. This method generates smaller size of generalized matrices than original 3-D NMM method. In addition, this method changes the distribution of eigenvalues which needs less eigenvalues under the same error. Results show that the new method spends only 1/7 CPU time and uses only 1/3 memory of the original method.

1. INTRODUCTION

Resistivity Log is a widely used method for calculating the resistivity distribution of the ground layers, which is an inversion problem and every step of inversion calculation requires constructing different models for the forward calculation, therefore, the speed of forward calculation decides the speed of inversion calculation. As a result, a high speed of forward calculation is needed for inversion problem, especially for the real time inversion calculation in workplace.

NMM is a very efficient semi-analytical and semi-numerical method and is widely used in electromagnetic field well-logging [1]. The NMM method is used to model various multiregion vertically stratified inhomogeneous media. The efficiency of the NMM method is based on the idea that it can reduce the $n - D\,(n = 2,\,3)$ problem to a series of $(n - 1) - D$ problem, and then, the field in all layers can be obtained by a recursive (analytical) scheme.

When the media is axisymmetric, the 2-D problem (the source is centered) or the 2.5-D problem (the source is off-centered), can be calculated by the 2-D NMM method. Furthermore, when the media of every layer is inhomogeneous horizontally (on the x-y plane), a 3-D NMM method is introduced [2]. This method reduces the original 3-D problem into a series of 2-D eigenvalue problems, which can be discretized to generalized matrix eigenvalue equations by the 2-D FEM analysis (on the x-y plane). The field of all layers is obtained by a recursive analysis basing on the generalized matrix eigenvalue equations of all layers.

Comparing to the 3-D standard FEM, the 3-D NMM method is efficient because that it reduces the original 3-D problem into the 2-D problems which can be solved by FEM and speeds up the solution greatly. However, when discretizing the 2-D open boundary problem to form the generalized matrix eigenvalue equations, the 3-D NMM method adopted 2-D standard FEM, which leads to large number of FEM nodes and truncation error, since it places an artificial boundary far away from the near field domain with a Dirichlet or Neumann boundary condition applied to the artificial boundary.

In this paper, a domain map finite element method was introduced to discretize the 2-D open boundary problem. This method is based on the Kelvin transform, which transforms the outer infinite field domain into a bounded field domain. It divides the whole infinite field domain into two parts: the inner and outer domains. The inner domain is close to and includes the source and the outer domain is the rest. In the inner domain, the standard FEM equations are established and in the outer domain, the Kelvin transform is applied at first and for the transformed domain, similar equations are established. At last, the total equations are formed by coupling the nodes which are lie on the common boundary of these two domains.

Since no truncation on the domain, this method could have smaller error comparing to the standard FEM. Furthermore, this method needs less elements to discretize the transformed domain and so it is faster to solve the open boundary problem than the 2-D standard FEM. Results also show that there are more small eigenvalues of the generalized matrix eigenvalue equations formed

by this method than that of formed by standard FEM. These small eigenvalues can be truncated without lossing the accurate. As an improvement of [2], this method is applied into the resistivity well-logging successfully.

2. FORMULATION OF DOMAIN MAP FINITE ELEMENT METHOD

The 2-D eigenvalue equation in [2] can be expanded in terms of basis functions and the resulting generalized matrix eigenvalue equation is of the form

$$\boldsymbol{A} \cdot \boldsymbol{v}_j = k_{jz}^2 \boldsymbol{B} \cdot \boldsymbol{v}_j \tag{1}$$

where k_{jz}^2 are eigenvalues, $\boldsymbol{v}_j$ are corresponding eigenvector. In [2], the matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ are given by standard FEM calculation:

$$\boldsymbol{A} = \sum_{e\in\Omega} \iint_e \sigma^e \left(\nabla N^e\right)^T \cdot \nabla N^e d\Omega, \quad \boldsymbol{B} = \sum_{e\in\Omega} \iint_e \sigma^e N^e \cdot (N^e)^T d\Omega \tag{2}$$

In (2), $N^e = [N_1^e \ N_2^e \ldots]^T$, $\nabla N^e = \left[\frac{\partial N^e}{\partial x} \ \frac{\partial N^e}{\partial y}\right]^T$, and σ^e and N_i^e are the conductivity and shape functions of every element.

Generally the field domain Ω in (2) is open bounded and usually a large bounded domain is used to replace the original one, which leads to large number of FEM nodes and truncation error. In this paper, the domain map finite element method is firstly introduced into the 3-D NMM method. Results show that the improved method can solve the open bounded electromagnetic field much faster and more accurate.

The matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ in Equation (1) can be calculated together by the following function

$$F(N) = \sum_{e\in\Omega} \iint_e \left(\sigma(x,y)\left((\nabla N^e)^T \cdot \nabla N^e + N^e \cdot (N^e)^T\right)\right) d\Omega \tag{3}$$

The idea of domain map finite element method to calculate $F(N)$ is as follows: divide Ω by a closed boundary Γ into Ω_i and Ω_o, the inner domain Ω_i include the source and important areas and the outer domain Ω_o is the rest part, as shown in Fig. 1. As a result, the Ω in (3) can be rewritten as the sum of Ω_i and Ω_o. Since the inner domain can be calculated by standard FEM, only the outer domain is considered.

The outer domain Ω_o can be mapped to a bounded domain Ω_t by a coordinate transformation $\mathfrak{R}$ (the details are shown in the next section). Fig. 1 shows the transformation from Ω_o to Ω_t. Consider the continuity conditions for electric potentials, the boundary Γ should be transformed to itself, i.e., $\mathfrak{R}\Gamma = \Gamma$.
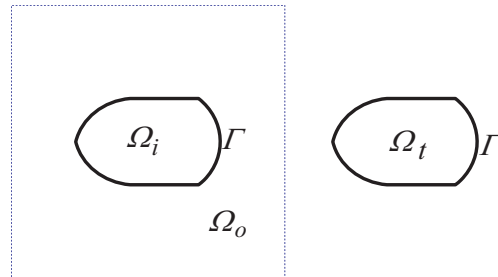

Figure 1: Map domain Ω_o to Ω_t by a coordinate transformation $\mathfrak{R}$.

Under the coordinate transformation $\mathfrak{R}$, the old coordinate (x, y) becomes the new one (x', y') and the derivative of shape functions to the coordinates is as follows [3]

$$\nabla N = J \cdot \nabla' N \tag{4}$$

where $\nabla N = \left[\frac{\partial N}{\partial x} \ \frac{\partial N}{\partial y}\right]^T$, $\nabla' N = \left[\frac{\partial N}{\partial x'} \ \frac{\partial N}{\partial y'}\right]^T$, and J is the Jacobian matrix of coordinate transformation.

According to (4), the derivative of shape functions in (3) can be written as

$$\sigma(x, y) \left(\nabla N^e\right)^T \cdot \nabla N^e = \sigma(x', y') \left(\nabla' N^e\right)^T \cdot J^T \cdot J \cdot \nabla' N^e \tag{5}$$

where $\sigma(x', y') = \sigma(\mathfrak{R}(x), \mathfrak{R}(y))$, $\nabla' N^e = \left[\frac{\partial N^e}{\partial x'} \ \frac{\partial N^e}{\partial y'}\right]^T$.

According to (5), the integral $F(N)$ of discretized domain Ω_t in the new coordinate system can be written as

$$F_o\left(N\right) = \sum_{e \in \Omega_t} \iint_e |J|^{-1} \left(\sigma(x', y') \left(\nabla' N^e\right)^T \cdot J \cdot J^T \cdot \nabla' N^e + \sigma(x', y') N^e \cdot (N^e)^T\right) d\Omega'. \tag{6}$$

The final matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ can be calculated by combining the integral of (6) and the integral of inner domain.

Note that the domain map finite element method is based on a transformation $\mathfrak{R}$ which should ensure the transformed domains satisfy continuity conditions. In this paper, a simple and efficient transformation-the Kelvin transformation which satisfies such requirement is adopted.

3. KELVIN TRANSFORMATION AND THE STIFF MATRIX

Under 2-D Cartesian coordinate, the Kelvin transformation is as follows [4]

$$\begin{cases} x' = R^2 x/r^2 \\ y' = R^2 y/r^2 \end{cases}, \qquad r^2 \geq R^2 \tag{7}$$

where $r^2 = x^2 + y^2$ and R is the radius of interface (2-D: circle).

Under Kelvin transformation the Jacobian matrix can be written as

$$J = \left(r'/R\right)^2 J^0 \tag{8}$$

where $J^0 = \frac{1}{(r')^2} \begin{bmatrix} (y')^2 - (x')^2 & -2x'y' \\ -2x'y' & (x')^2 - (y')^2 \end{bmatrix}$, $(r')^2 = (x')^2 + (y')^2$.

Note that J is orthogonal and $\det(J) = -\left(r'/R\right)^4$. Function (6) can be simplified as follows

$$F_o\left(N\right) = -\sum_{e \in \Omega_t} \iint_e \left(\sigma(x', y') \left(\nabla' N^e\right)^T \cdot \nabla' N^e + \sigma(x', y') \left(R/r'\right)^4 N^e \cdot (N^e)^T\right) d\Omega'. \tag{9}$$

and the final stiff matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ can be calculated basing on (9).

4. APPLICATION

The domain map finite element method is successfully applied to the 3-D NMM method for the forward problem of resistivity well-logging. In this application, the electrodes and sources move inside the borehole and the potentials on electrodes are to be solved. Basing on this fact, the field domain (the whole x-y plane) is divided into two parts by borehole and the interface of Kelvin transformation is the face of the borehole.

The inner part is the part inside the borehole which includes the electrodes, some of which are current sources. Since every electrode is equi-potential, special process should be done to ensure the potentials on the electrode are equal. It is well know that when conductor is placed inside the electric field, the surface of the conductor is equi-potential. Basing on the fact that conductor has high conductivity, a simple and effective method is adopted to process this problem in this paper: act the electrode as the medium with high conductivity (for example, 1.0e+7 s/m), and then place a point current source to the center of the electrode to ensure the value of current from the point current source is equal to the total current from the surface of electrode current source. This method is much easier and faster to process the electrode current source basing on the original 3-D NMM method which can only process the point current source than some other method [5].

The outer part is outside the borehole. Fig. 2 shows an example that the borehole is horizontal and there are three different ground layers. In Fig. 2(a), Γ_1 and Γ_3 are the interfaces of different ground layers and after Kelvin transformation, the domain is divided by two different circles, as shown in Fig. 2(b).

The meshes [6] of the inner and outer domain are shown in Figs. 2(c) and (d). In Fig. 2(c), the gray area is electrode which is acted as a special medium and in Fig. 2(d), the center of the disc is correspondent to the infinite point and the gray areas are the domains of Ω_1 and Ω_3 in Fig. 2(a).
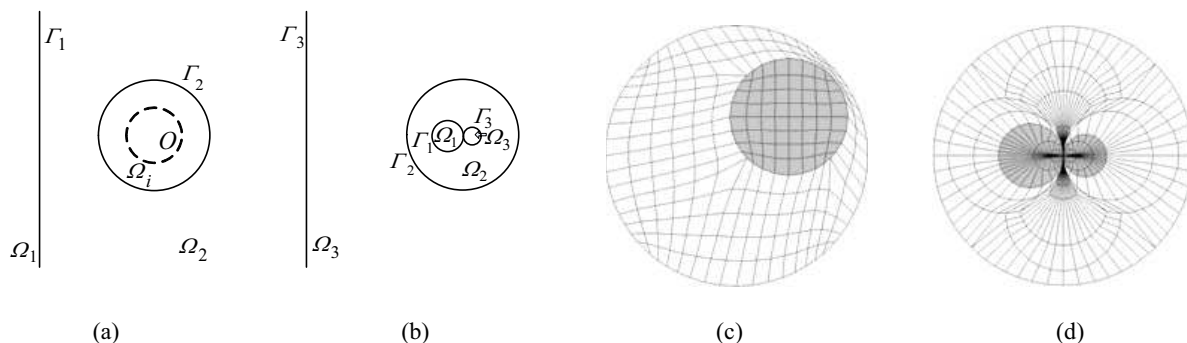
(a)　　　　(b)　　　　(c)　　　　(d)

Figure 2: Domains map and the mesh generation, (a) original regions, (b) mapped domains under Kelvin transformation, (c) mesh of inner domain, (d) mesh of mapped domain.

5. NUMERICAL RESULT

A numerical example is given to verify this method by comparing the speed and results with other methods. Fig. 3(a) shows a typical group of electrodes' arrange of an array laterlog tool (symmetric about x-y plane). By controlling electrical potential balance of electrodes in different working frequencies, the tool can work in different logging modes with different detecting depths. In these modes, electrode A_0 is the main current supply electrode. Fig. 3(a) also shows a shallow logging mode which means that electrode A_1 is the shielding electrode of A_0 and the current returns to A_2, A_3, A_4, A_5 and A_6.
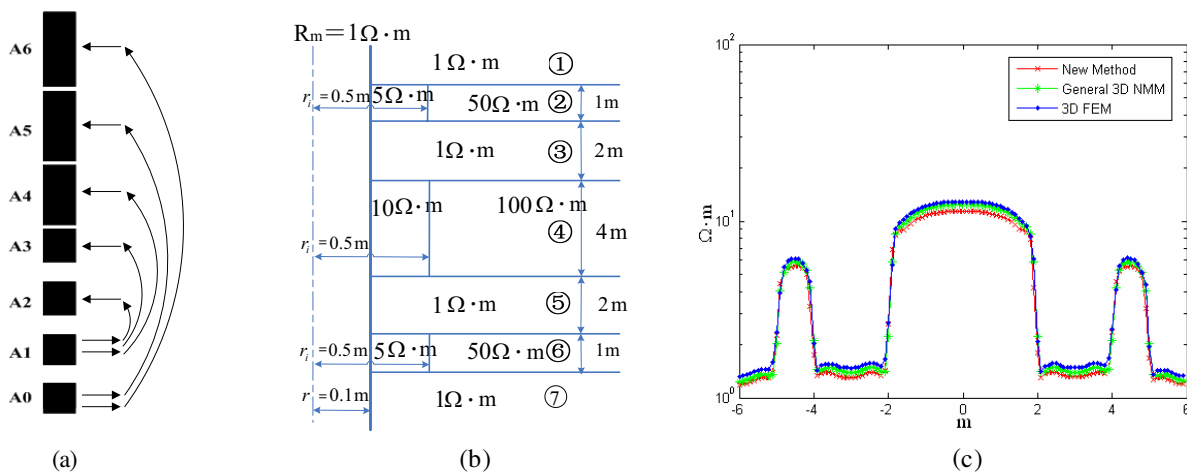

(a)　　　　(b)　　　　(c)

Figure 3: Numerical result, (a) original regions, (b) resistivity of ground layers, (c) apparent resistivity of different methods for the same logging mode.

For the mode shown in Fig. 3(a), the balance equations is as follows

$$\begin{cases} I_0 = 0.5, \quad \sum_{i=0}^{6} I_i = 0 \\ u_0 = u_1, \quad u_2 = u_3 = u_4 = u_5 = u_6 \end{cases}. \tag{10}$$

Figure 3(b) shows the model of ground layers. In this example, the borehole is vertical and its radius is 0.1 m. The tool is off-axis in the borehole and the radius of electrode is 0.045 m. Fig. 3(c) shows the results of a mode calculated by general 3-D FEM, general 3-D NMM and the improved 3-D NMM method presented in this paper. Results show that these three methods have good agreement.

The comparisons of the speed and memory usage of these methods are shown in Table 1. Results show that the method presented in this paper is almost 7 times faster than the general 3-D NMM and the usage of memory is only about one third of that used by the general 3-D NMM with almost the same accuracy.

Comparison results also show that the truncation of eigenvalues by the improved 3-D NMM method can be much more than that of general 3-D NMM method without reducing the precision, which also speed up the time of solving eigenvalue equations.

Table 1: Speed comparison of different methods.

	3-D FEM	3-D NMM	Method in this paper
Time (s)	475.5	16.5	2.4
Memory (Mb)	3,862	236	85

6. CONCLUSIONS

By introducing the Kelvin transformation, the open boundary electromagnetic problem in 2-D can be solved efficiently without adding artificial boundary. In this paper, the Kelvin transformation is also successfully applied to the regions which are inhomogeneous in infinite point. As an application of this method, a 3-D well-logging problem is solved based on the 3-D NMM which was created by the general 2-D FEM previously. Results show that this method is much faster than the original method and saves much memory. This great improvement makes it possible to do real-time inverse work for the 2.5-D well-logging problem.

ACKNOWLEDGMENT

REFERENCES

1. Chew, W. C., *Waves and Fields in Inhomogeneous Media*, Van Nostrand Reinhold, New York, 1990.
2. Fan, G.-X. and Q. H. Liu, "3-D numerical mode-matching (NMM) method for resistivity well-logging tools," *IEEE Trans. Antennas Propagat.*, Vol. 48, No. 5, 1544–1553, 2000.
3. Silvester, P. P., *Finite Elements for Electrical Engineers*, Cambridge University Press, 1996.
4. http://en.wikipedia.org/wiki/Inversion_%28geometry%29.
5. Nie, Z., F. Yang, X. Nie, Y. Yuan, and S. Cheng, "An efficient numerical modeling of the DC resistivity logging: Theory and applications," *Journal of Electronics*, Vol. 14, No. 4, 169–179, 1997.
6. Thompson, J. F., B. K. Soni, and N. P. Weatherill, *Handbook of Grid Generation*, CRC-Press, 1998.

Study of Optical Propagation in Hybrid Periodic/Quasiregular Structures Based on Porous Silicon

José Escorcia-García[1] and Miguel E. Mora-Ramos[2]
[1]Centro de Investigación en Ingeniería y Ciencias Aplicadas
Universidad Autónoma del Estado de Morelos
Ave. Universidad 1001, Cuernavaca 62209, México
[2]Facultad de Ciencias, Universidad Autónoma del Estado de Morelos
Ave. Universidad 1001, Cuernavaca 62209, México

Abstract— In this work we report the propagation of the light waves within Bragg-Fibonacci (BM-FN) hybrid structures and the appearance of photonic bandgaps in the reflectance spectrum of these systems, which are studied theoretically with the use of oblique incidence transfer matrix technique. In all cases, the discussion is made regarding the possible application of these structures as photonic quasicrystals. For the simulation we have adopted the typical parameters of a Porous-Silicon dielectric multilayer. It is found that a hybrid configuration of the type BM-FN-BM is a promising candidate for resonant microcavities with strong mode localization.

1. INTRODUCTION

Light propagation in quasiregular dielectric heterostructures has been a subject of study in recent years (see, for instance, references [1–11]).

The design of hybrid periodic-quasiperiodic systems is among the proposals for the obtention of one-dimensional (1D) quasicrystals with original properties. For instance, the selective spatial localization of atom displacements has been studied in one-dimensional hybrid quasi-regular/periodic atomic chains [12, 13]. In the case of dielectric multilayers, the hybrid systems are said to show complementary optical responses [6]. On the other hand, by combining two Fibonacci quasi-periodic structures and a periodic structure to form a heterostructure, a broad omnidirectional reflection band is obtained [14].

In this work we provide a theoretical analysis of the optical propagation in hybrid periodic/Fibonacci dielectric multilayers. Periodic parts of the structures are represented by $\lambda/4$ Bragg mirrors. Fibonacci generations of A and B layers of the same widths of Bragg ones are the corresponding quasi-regular constituents in the systems. The geometrical configurations considered are of the type Bragg-Fibonacci-Bragg (BM-FN-BM) and Fibonacci-Bragg-Fibonacci (FN-BM-FN). We shall show that the BG-FN-BG array is the most convenient if we are looking for sharp-shaped microcavities within the photonic band gap.

2. MODEL SYSTEMS AND SIMULATION TOOL

Hybrid structures here studied are designed using Bragg reflectors fulfilling the $\lambda_0/4$ condition. In our particular case we have chosen $\lambda_0 = 800\,\text{nm}$, and the corresponding refractive indices are $n_A = 1.2$ and $n_B = 1.8$. Consequently, the layer widths are $d_A = \lambda_0/4n_A = 166.67\,\text{nm}$, and $d_B = \lambda_0/4n_B = 111.11\,\text{nm}$ (*set I*). A second choice exchanges the values of the refractive indices: $n_A = 1.8$ and $n_B = 1.2$. Thus, $d_A = 111.11\,\text{nm}$, and $d_B = 167.67\,\text{nm}$ (*set II*). This exchange does not affect the optical response of the Bragg mirrors but it does significantly alter the features of light propagation throughout the Fibonacci substructure. The values of the refractive indices here selected are among those that can be experimentally achieved by using layers made of porous silicon (p-Si) [11].

The theoretical simulation is made using the transfer matrix formalism. This tool is derived for one-dimensional problems that can be resolved via the solution of a master equation which ultimately reduces to a linear system of first order ordinary differential equations [15]. The analysis is made in the case of the 1D Maxwell equation for the TE modes propagating in the structure. Normal incidence is assumed, although generalization to oblique angles of incidence is straightforward [16].

3. RESULTS AND DISCUSSION

In Figure 1 it is shown the outcome of the calculation of the reflectance spectrum for hybrid Bragg-Fibonacci p-Si dielectric heterostructures. Two periods Bragg mirrors together with second

generation Fibonacci multilayer (ABA) are the components of the structures. To distinguish the results, a *prime* is include as a superscript. Its presence indicates that the parameters for the A, B layers are those of the *set I*. Unprimed labels indicate the use of *set II* in the calculation. The reflectance from both hybrid geometries are included.
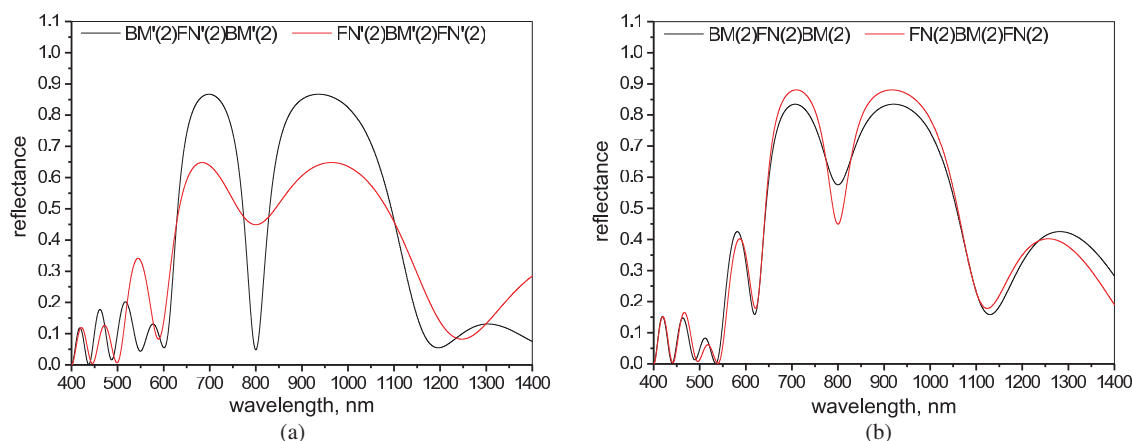

Figure 1: (a) Reflectance spectra for a hybrid porous silicon Periodic/Fibonacci dielectric heterostructure. Periodic substructures consist of a two periods Bragg mirror while the quasiregular are second generation Fibonacci multilayers. Black curve represents a Bragg-Fibonacci-Bragg system while the red curve is the resulting reflectance spectrum of a Fibonacci-Bragg-Fibonacci one. The layer widths and refractive indices considered are $d_A = 166.67$ nm, and $d_B = 111.11$ nm, and $n_A = 1.2$, $n_B = 1.8$. (b) The same as in Figure (a) but with input parameters given by $d_A = 111.11$ nm, and $d_B = 166.67$ nm, and $n_A = 1.8$, $n_B = 1.2$.

Differences between the spectra of Figures (a) and (b) can be attributed to a combination of geometric and dielectric contrast effects. In both figures, the black curve corresponds to the BM-FN-BM distribution (11 layers). Curves in red represent the reflectance of a FN-BM-FN multilayer (10 layers). This latter is precisely the design proposed by Maciá for a resonant microcavity [6]. A double-A-layer *defect* appears in each heterostructure. But in the BM-FN-BM geometry it is located almost at the right end, following a three periods Bragg reflector. On the contrary, in the FN-BM-FN structure, such defect is very closed to the light incidence point (left end of the system), following only one period Bragg reflector. Furthermore, in the whole structure, the amount of layers A is greater than that of layers B. This means that in the case of using the *set II* of parameters, there will be a larger number of layers bearing a higher refractive index. The opposite is true in the case of using the *set I*.

Being the total number of layers so small, a photonic transmission bandgap (PBG) located around λ_0 will correspond at most to an 87–88% of reflectance. The higher dielectric contrast associated to the use of the *set II* of parameters makes that there is only a slight difference between both geometries in Figure 1(b), regarding the formation of a PBG. In the reflectance spectra of the heterostructures using the *set I* (Figure 1(a)), the position of the AA-defect with higher index of refraction determines that the BM-FN-BM geometry will have the most defined PBG.

The most important feature appearing when a comparison between Figures 1(a) and 1(b) is made is the presence of a very well defined resonant microcavity for the BM-FN-BM geometry with the *set I* of parameters. It is precisely centered at λ_0 and represents a strong localization of this particular mode within the PBG of the structure. The remarkable aspect of this property is that it is achieved in a heterostructure made of only eleven layers.

Increasing the number of layers in the systems serves to confirm the previously discussed features. Of course, now the combined effect of geometry and refractive index distribution becomes more complex; but the basic explanation remains in the spirit of the above stated. Figure 2 presents the calculated reflectance when the order of the Fibonacci generation is kept as 2 while the constituent Bragg mirrors are of four periods. This means that the BM-FN-BM systems consists of 19 layers and the FN-BM-FN ones have 14 layers. Now the PBG reach the 100% of reflectance in the BM-FN-BM configuration and, the λ_0-centered microcavity is precisely defined in this particular geometry. A sharply defined one is again associated with the use of the parameters of *set I*.

The possibility of modifying the number of layers by changing the orders of the BM and FN substructures as well as the combination of *set I* and *set II* parameters within one geometry opens a way for many distinct variations. To illustrate this, Figure 3 includes a case in which the order
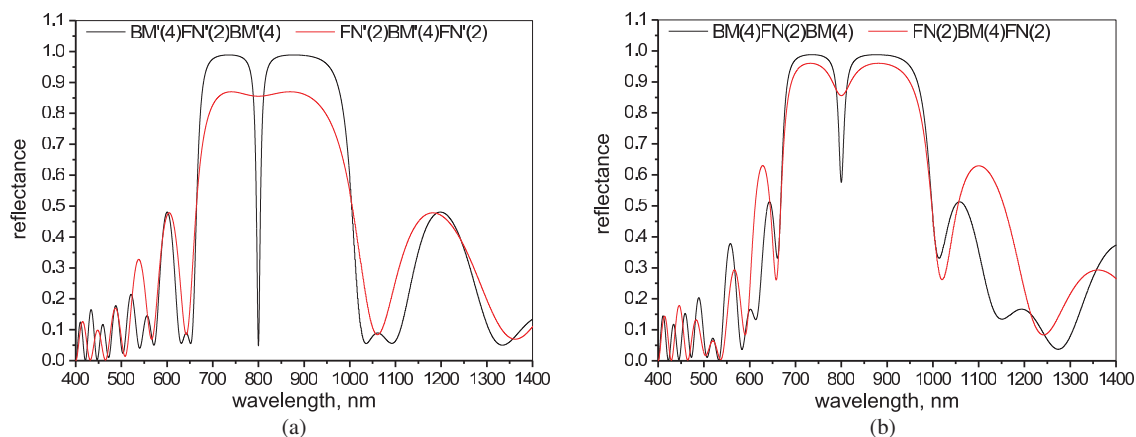
Figure 2: The same as corresponding Figure 1; but in this occasion using a four periods Bragg mirror periodic substructure.

of the FN generation is increased to 3 (FN(3)→ $ABAAB$), keeping the four periods of the Bragg mirrors, BM(4), (3(a)). Another situation (3(b)) is when the BM(4) substructure is made using parameters of *set I*, while the FN(3) consists of layers with refractive indices defined according to *set II*. Once again, we focus our attention in the BM-FN-BM geometry. Black curve in Figure 3(a) shows a reflectance spectrum with a sharply-defined microcavity centered at λ_0. On the other hand, the black curve in Figure 3(b) corresponds to a quite distinct reflectance spectrum in which it is seen the λ_0-centered microcavity together with a tendency of the PBG to widen and, most interestingly, the presence of two microcavity-like structures around 690 nm and 950 nm which are related to full transmission modes in the structure. The only difference between 3(a) and 3(b) multilayers is the change in the values of the layers width and refractive indices associated to the third order FN "defect". Therefore, the significantly distinct features of the spectra in both cases can be attributed to the role played by the dielectric contrast in the system.
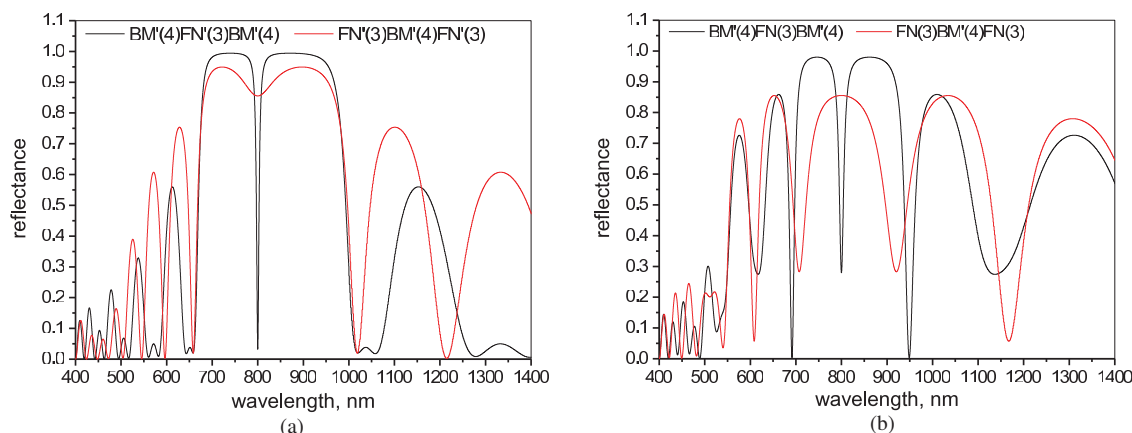

Figure 3: (a) Reflectance spectra for a hybrid porous silicon Periodic/Fibonacci dielectric heterostructure. Periodic substructures consist of four periods Bragg mirrors. Quasiregular ones are third generation Fibonacci multilayers. Curves distribution and input parameters are the same of Figure 1(a). (b) Light reflectance from Four-periods-Bragg/Third-generation-Fibonacci hybrid porous silicon multilayers. Bragg mirrors have the same values of layer widths and refractive indices of Figure 1(a). Fibonacci substructures are designed with the input values used in calculating the spectra of Figure 1(b).

4. CONCLUSIONS

In this article we have presented some theoretical results on the propagation of light in hybrid periodic/Fibonacci dielectric heterostructures. With the idea of a possible practical realization the refractive indices, used as input parameters in the designs presented, are compatible with those experimentally achieved using porous silicon. The main conclusion of the work is that resonant microcavities with strong mode localization can be realized in systems with a small number of layers. Contrary to the suggestion made by Maciá [6], we have found that the most favorable

configuration to obtain such kind of microcavities is that where the quasiregular substructure is located in between two periodic ones. This might be true for porous silicon although some other values of refractive indices achievable for this material, and different layer widths, need to be investigated. Due to the great amount of possible combinations for d_A, d_B, n_A, and n_B that can be explored, we believe that the study of light propagation in the hybrid periodic/quasiregular dielectric heterostructures is worth to be continued.

ACKNOWLEDGMENT

REFERENCES

1. Gellermann, W., M. Kohmoto, B. Sutherland, and P. Taylor, "Localization of light waves in Fibonacci dielectric multilayers," *Phys. Rev. Lett.*, Vol. 72, No. 5, 633–636, 1994.

2. Liu, N.-H., "Propagation of light waves in Thue-Morse dielectric multilayers," *Phys. Rev. B* Vol. 55, No. 6, 3543–3547, 1997.

3. Musikhin, S. F., V. I. Il'in, O. V. Rabizo, L. G. Bakueva, and T. Y. Yudintseva, "Optical properties of quasiperiodic and aperiodic PbS-CdS superlattices," *Semiconductors*, Vol. 31, No. 1, 46–50, 1997.

4. Macia, E., "Optical engineering with Fibonacci dielectric multilayers," *Appl. Phys. Lett.*, Vol. 73, No. 23, 3330–3332, 1998.

5. Yang, X., Y. Liu, and X. Fu, "Transmission properties of light through the Fibonacci-class multilayers," *Phys. Rev. B*, Vol. 59, No. 7, 4545–4548, 1999.

6. Macia, E., "Exploiting quasiperiodic order in design of optical devices," *Phys. Rev. B*, Vol. 63, 205421-1–8, 2001.

7. Qiu, F., R. W. Peng, X. Q. Huang, Y. M. Liu, M. Wang, A. Hu, and S. S. Jiang, "Resonant transmission and frequency Trifurcation of light waves in Thue-Morse dielectric multilayers," *Europhys. Lett.*, Vol. 63, No. 6, 853–859, 2003.

8. Dal Negro, L., M. Stolfi, Y. Yi, J. Michel, X. Duan, L. C. Kimerling, J. LeBlanc, and J. Haavisto, "Photon band gap properties and omnidirectional reflectance in Si/SiO$_2$ Thue-Morse quasicrystals," *Appl. Phys. Lett.*, Vol. 84, No. 25, 5186–5188, 2004.

9. Jiang, X., Y. Zhang, and S. Feng, "Photonic band gaps and localization in the Thue-Morse structures," *Appl. Phys. Lett.*, Vol. 86, 201110-1–3, 2005.

10. Carretero, L., M. Perez-Molina, P. Acebal, S. Blaya, and A. Fimia, "Matrix method for the study of wave propagation in one-dimensional general media," *Opt. Express*, Vol. 14, No. 23, 11385–11391, 2006.

11. Agarwal, V. and M. E. Mora-Ramos, "Optical characterization of polytype Fibonacci and Thue-Morse quasiregular dielectric structures made of porous silicon multilayers," *J. Phys. D: Appl. Phys.*, Vol. 40, 3203–3211, 2007.

12. Montalbán, A., V. R. Velasco, J. Tutor, and F. J. Fernández-Velicia, "Phonon confinement in one-dimensional hybrid periodic/quasiregular structures," *Phys. Rev. B*, Vol. 70, 132301-1–4, 2004.

13. Montalbán, A., V. R. Velasco, J. Tutor, and F. J. Fernández-Velicia, "Selective spatial localization of the atom displacements in one-dimensional hybrid quasi-regular (Thue-Morse and Rudin-Shapiro)/periodic structures," *Surf. Sci.*, Vol. 601, 2538–2547, 2007.

14. Wen, D. J., H. Peng, and W. H. Zhou, "Broad omnidirectional reflection band forming using the combination of Fibonacci quasi-periodic and periodic one-dimensional photonic crystals," *Chinese Phys. Lett.*, Vol. 20, 1963–1965, 2003.

15. Mora, M. E., R. Pérez, and Ch. Sommers, "Transfer matrix in one-dimensional problems," *J. Physique*, Vol. 46, 1021–1026, 1985.

16. Kumar, N. and S. P. Ojha, "Photonic crystals as infrared broadband reflectors with different angles of incidence: A comparative study," *Progress In Electromagnetics Research*, PIER 80, 431–445, 2008.

Magnetic Proximity Effect in Isolator Crystal Pairs

Yoav Linzon[1], **Marcello Ferrera**[1], **Luca Razzari**[1, 2], **Alain Pignolet**[1], **and Roberto Morandotti**[1]

[1]Universite du Quebec, Institute National de la Recherche Scientifique
Varennes, Quebec J3X 1S2, Canada
[2]Dipartimento di Elettronica, Universita di Pavia
via Ferrata 1, 27100 Pavia, Italy

Abstract— We experimentally study the polarization dynamics (orientation and ellipticity) of near infrared light transmitted through magnetooptic Yttrium Iron Garnet isolator crystal pairs using a modified balanced detection scheme. When the pair separation is in the sub-millimeter range, we observed a proximity effect in which the saturation field is reduced by up to 20%. Calculations using a 1D model suggest that the proximity effect originates from magnetostatic interactions between the dipole moments of the isolator crystals.

1. INTRODUCTION

Optical isolators are important polarization components which are controlled by static external magnetic fields. They induce nonreciprocal polarization phase shifts originating from the so-called magneto-optical (MO) Faraday rotation [1]. The unidirectional nonreciprocal polarization control is crucial for the reduction of reflection-related instabilities in active electro-optical devices [1–6]. A fully functional isolator consists of a magnetically active isolator crystal situated between two crossed polarizers, and a rotation equal to $\pm 45°$ of an input linear polarization is required at the wavelength of operation, in the presence of an external magnetic field $\pm H_{ext}$. Ferrimagnetic iron garnets are popular materials to serve as isolators in the visible and near infrared, as they possess large induced magnetizations, leading to the highest known Faraday rotations in the spectral range used for today's optical telecommunication systems [7]. Specifically, in an Yttrium Iron Garnet (YIG, $Y_3Fe_5O_{12}$) crystal subject to an external magnetic field, the latent Faraday rotation angle per unit thickness, at the principal telecom wavelength ($1.55\,\mu m$), is typically $\sim 0.016°/\mu m$ at saturation, when the light propagates along the crystal's easy axis of magnetization, [001] [7, 8]. As a result, a YIG crystal operating in the telecom spectral region typically requires a length of several millimeters to be effective as an optical isolator. In propagation of light through such thick MO samples, the (usually) desirable circular birefringence is accompanied by substantial magnetically-induced circular dichroism [8], which introduces an ellipticity ("smearing") ε of the polarization state [9].

While the optical response of MO single crystals has been thoroughly characterized in past studies [2–8], the polarization dynamics associated to the **combination of several separate MO components**, located in close proximity and subject to a uniform magnetic field, has not been addressed before. Such a study, however, is important in view of recent potential applications involving the integration of several optoelectronic devices on the same chip [2–6, 10]. In such devices the overall response of a set of magneto-optical components, including magnetostatic proximity effects [11], plays a significant role, and can lead to surprising behaviors. Here we report our study of the polarization dynamics (in terms of orientation and ellipticity, as a function of the external DC magnetic field) of light transmitted through YIG isolator crystal pairs in series, using a modified balanced detection scheme. For small crystal separations, we observed a magnetostatic proximity effect [11], in which the saturation field is reduced significantly, i.e., by up to 20%.

2. SECTION 1: EXPERIMENTAL SETUP AND TECHNIQUE

The experimental setup is represented in Fig. 1. A linearly polarized laser diode beam at a wavelength of $1.55\,\mu m$ is collimated with a telescope and passed through a chopper (C) which synchronizes a lock-in amplifier (LIA) at 1 kHz repetition. The combination of a quarter wave plate (QWP) and a polarizer (P) enables the preparation of linearly polarized light with a constant intensity for any chosen orientation of P. The input polarization is always set initially to a value of $+45°$ in the first quadrant [$s > 0$ and $p > 0$, see Fig. 1(a)]. After propagation through the magnet, the output beam enters a polarizing beam splitter (PBS) which separates the s and p components, both of which are detected by a pair of identical photodiode detectors connected to the LIA operating in differential mode. Although PBSs are generally not ideal components, i.e., they transmit a small

amount of p and reflect some residual s-polarization, here these artifacts were measured to be negligible ($< 1\%$). As optical detectors always measure intensity rather than amplitude, the signal read by the amplifier when both channels are open is $s^2 - p^2$. When either one or both of the polarization components change sign, a vectorial correction must be applied, as described below. Each component can be blocked individually with a mechanical shutter (S1 or S2) to read the other component independently. The total optical energy is proportional to $s^2 + p^2$, a quantity which is nearly constant in all scans of the magnetic field. In our experiments, a GMW magnet (model 3470) was used. With a separation of 2 cm between the poles, the magnet delivers a nearly uniform magnetic field of up to 4 kg in the volume between the poles. Inside the magnet, either a single crystal [Fig. 1(b)] or a pair of crystals [Fig. 1(c)] have been used. Each MO sample is a 2.8 mm-long single crystal YIG rod of 5 mm diameter, polished and anti-reflection coated for operation at the laser's wavelength. The easy axis of magnetization [001] is nearly parallel to the rod axis, which is also the direction of the optical beam and the magnetic field [see Figs. 1(b), (c)]. The transmission is 88% with a single crystal and 80% with the pair setting.
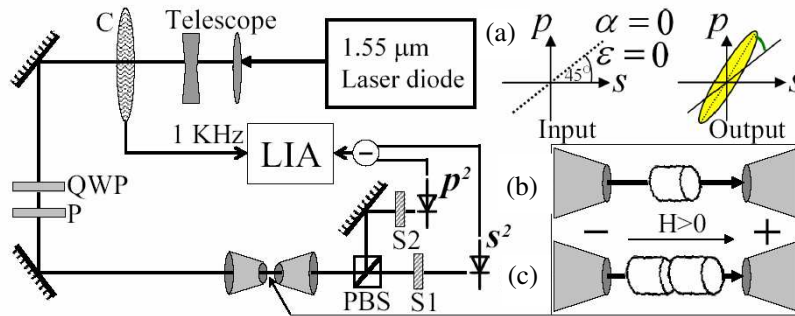

Figure 1: Experimental schematics, as described in the text. (a) Definition of input (left) and output (right) polarizations. (b)-(c): Volume within the magnet with (b) single and (c) double crystal settings.

A general output state of polarization has both an orientation angle α with respect to the input state and an ellipticity angle ε [9] [see Fig. 1(a)]. We define a normalized balanced detection signal as $D = \left(s^2 - p^2\right) / \left(s^2 + p^2\right)$. In the presence of ellipticity the entire energy can never be set parallel to a single linear polarization state (as opposed to the input state): some residual energy always remains in the other orthogonal state, corresponding to the minor axis of the ellipse. Closing S1 and maximizing the signal by rotating P yields a signal $p^2_{\max}$ that corresponds to alignment of the ellipse's major axis along p. In the same setting, closing S2 and opening S1 yields a corresponding signal $s^2_{\min}$ relating to the ellipse's minor axis aligned along s. The square root of the ratio between these signals is thus proportional to the ratio between the polarization ellipse axes, and we can thus define a normalized signal $\eta = \sqrt{\left|p^2_{\max}\right| / \left|s^2_{\min}\right|}$ which is measured in conjunction with D. A straightforward calculation yields the following relations between the measured signals, D and η, and the physical parameters, α and ε:

$$D \cdot \left(\eta^2 + 1\right) = \left(\eta^2 - 1\right) \sin 2\alpha \mp 2\eta \cos 2\alpha,$$
$$\varepsilon = \arctan \eta, \tag{1}$$

The sign in the right-hand term of the first relation corresponds to different quadrants of the output polarization, in the s-p plane. The actual quadrant of the polarization must thus be taken into account before applying Eq. (1). In the absence of circular dichroism, there is no ellipticity, such that $\eta = 0$, and accordingly, $D = \sin 2\alpha$. With $\eta \neq 0$, however, each equation has two real solutions, of which only one has physical meaning. This balanced detection measurement is sensitive to the sign of α, but not to that of ε.

3. SECTION 2: EXPERIMENTAL RESULTS

Experimental results are shown in Fig. 2 below. Considering first the single crystal geometry [Fig. 1(b)], the polarization dynamics (α and ε calculated from Eq. (1), as a function of the magnetic field) are shown in Figs. 2(a) and 2(b). Here the polarization is always in the first quadrant of the s-p plane, corresponding to the minus sign in Eq. (1). The sample exhibits slight hysteresis,

with a coercive field of ± 20 G. The slopes obtained from linear fits to $\alpha(H)$ and $\varepsilon(H)$, in fields far below saturation, are 4.1×10^{-2} °/(G · mm) for the Faraday rotation (Verdet coefficient) and 4.5×10^{-3} °/(G · mm) for the magnetic circular dichroism, in accordance with the literature [7, 8]. The orientation and ellipticity both saturate in a magnetic field of ± 820 G at the values of $\sim \pm 45°$ and $8°$, respectively. Turning to the double-crystal geometry [Fig. 1(c)], with a 1 mm separation between the crystals, the polarization dynamics are shown in Figs. 2(c) and 2(d). Above half the saturation field, the output polarization is in the second or fourth quadrant, and the plus sign is used in Eq. (1). The slopes of $\alpha(H)$ and $\varepsilon(H)$ far below saturation are approximately twice those obtained in the single crystal case, implying that the Verdet and the dichroism coefficients, normalized by the total thickness, are equal in both of the cases. The orientation and ellipticity again saturate at ± 820 G, and their values are $\sim \pm 90°$ and $15°$, respectively.
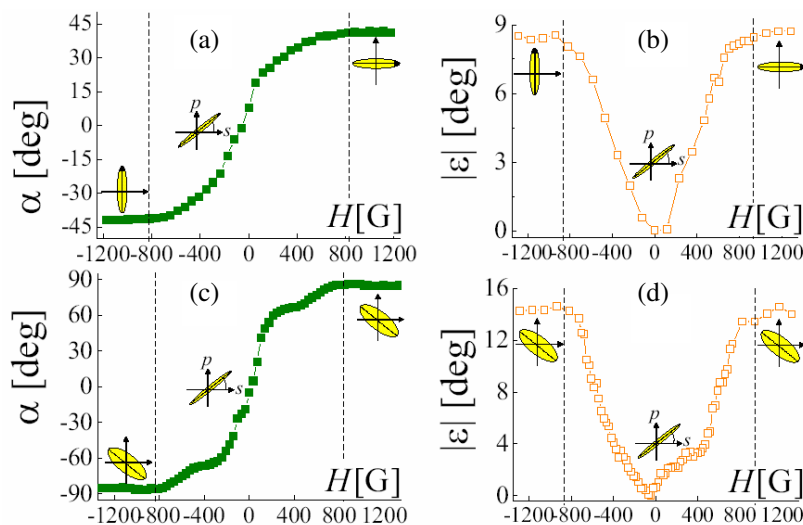

Figure 2: Polarization dynamics as a function of the magnetic field in the single crystal [(a),(b)] and in distant crystal pair [(c),(d)] geometries.

As the distance between the crystals is made small, however, we observed a strong dependence of the saturation field of both the orientation and ellipticity angles on the crystal separation. Fig. 3(a) shows the dynamics of ε for distant (squares) and adjacent (circles) crystal pair settings. While the saturation field is ± 820 G when the two crystals are set apart, similar to the case of a single crystal, in adjacent crystal settings the saturation field is reduced to ± 650 G, namely to 80% of the initial value. A detailed characterization of the saturation field (H_s) as a function of the pair facet-to-facet separation (d) [Fig. 3(b)] reveals that the proximity effect decays on a sub-millimeter separation scale. As both the Faraday rotation and magnetic circular dichroism are proportional to the total magnetization, this exponential dependence of the decay on the distance between the two crystals suggests that the proximity effect originates from effective magnetostatic interactions
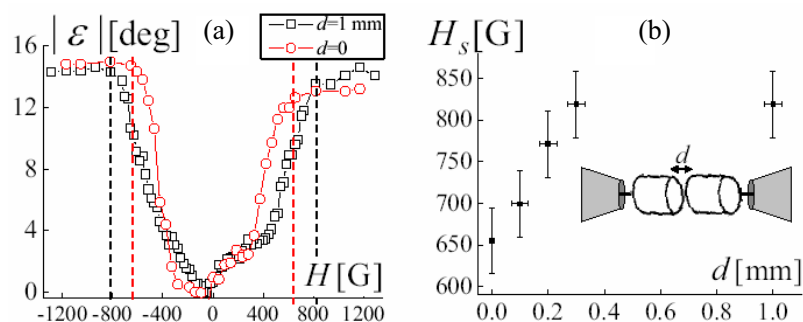

Figure 3: Effect of the proximity between the crystals on the saturation field. (a) Ellipticity dynamics in far ($d = 1$ mm, squares) and near ($d = 0$, circles) pair settings. (b) Saturation field H_s (in terms of both ellipticity and orientation) as a function of the separation d.

between the crystal magnetic moments in the presence of an external field [11]. Calculations using a 1D model (not shown) further support our interpretation of the proximity effect, and yield both an exponential decay dependence of the saturation field on the crystal separation and a maximum reduction of 20% in adjacent isolator settings, in agreement with the experimental results.

4. CONCLUSIONS

In conclusion, we have studied the polarization dynamics in YIG isolator crystal pairs using a modified balanced detection setup. A proximity effect was observed in which the saturation field is reduced by up to 20% for nearby crystals. The decay rate suggests that this effect originates from magnetostatic interactions between the nearby magnetic moments in the presence of an external magnetic field [11]. This substantial reduction of the saturation field is potentially useful for the realization of integrated magneto-optical devices [2–6, 10], since a lower saturating field would essentially mean lower device power consumption.

REFERENCES

1. Dillon, J. F., "Magnetooptics," *Journal of Magnetism and Magnetic Materials*, Vol. 100, 425–439, 1991.
2. Sato, T., J. Sun, R. Kasahara, and S. Kawakami, "Lens-free in-line optical isolators," *Optics Letters*, Vol. 24, 1337–1339, 1999.
3. Espinola, R. L., T. Izuhara, M. C. Tsai, and R. M. Osgood, "Magneto-optical nonreciprocal phase shift in garnet/silicon-on-insulator waveguides," *Optics Letters*, Vol. 29, 941–943, 2004.
4. Levy, M. and R. Li, "Polarization rotation enhancement and scattering mechanisms in waveguide magnetophotonic crystals," *Applied Physics Letters*, Vol. 89, 121113, 2006.
5. Holmes, B. M. and D. C. Hutchings, "Demonstration of quasi-phase-matched nonreciprocal polarization rotation in III-V semiconductor waveguides incorporating magneto-optic upper claddings," *Applied Physics Letters*, Vol. 88, 061116, 2006.
6. Shoji, Y., T. Mizumoto, H. Yokoi, I.-W. Hsieh, and R. M. Osgood, "Magneto-optical isolator with silicon waveguides fabricated by direct bonding," *Applied Physics Letters*, Vol. 92, 071117, 2008.
7. Dillon, J. F., "Magnetic and optical properties of rare earth garnets," *Journal of Magnetism and Magnetic Materials*, Vol. 84, 213–221, 1990.
8. Scott, G. B., D. E. Lacklison, H. I. Ralph, and J. L. Page, "Magnetic circular dichroism and Faraday rotation spectra of $Y_3Fe_5O_{12}$," *Physical Review B*, Vol. 12, 2562–2571, 1975.
9. Huard, S., *Polarization of Light*, Wiley-Interscience, New York, 1997.
10. Mekis, A., "Silicon photonics: Lighting up the chip," *Nature Photonics*, Vol. 2, 389–390, 2008.
11. Gubbiotti, G., S. Tacchi, G. Carlotti, P. Vavassori, N. Singh, S. Goolaup, and A. O. Adeyeye, "Magnetostatic interaction in arrays of nanometric permalloy wires: A magneto-optic Kerr effect and a Brillouin light scattering study," *Physical Review B*, Vol. 72, 224413, 2005.

New Measuring Method of Examination of Planar Optical Waveguides

D. V. Svistunov

St.-Petersburg State Polytechnic University, Russia

Abstract— A new method of mode spectrum measurement in planar waveguides is presented. The described end-fire mode spectroscopy technique exploits the difference of refraction angles of different modes skew-incident to the output face of the waveguide. This method provide reliable measurements of the whole mode spectra in any planar waveguides including step-index waveguides with thick cover layers and deep-buried graded-index waveguide structures. Furthermore, the technique allows conducting direct measurements of the maximal refractive index in arbitrary graded-index waveguides as well. Results of the comparative experiments performed by the end-fire mode spectroscopy and other techniques prove a feasibility of the developed method.

1. INTRODUCTION

Parameters of the optical waveguide (WG) define the operating performances of the whole photonic device. In order to optimize these parameters, appropriate technology conditions should be chosen. Necessary works include preliminary computer simulation of the process, fabrication of trial samples and their examination.

Important part of examination of a planar optical WG is measurement of the WG mode spectrum. Usually this procedure is performed by well-known m-line spectroscopy technique (e.g., see [1, 2]). However, in cases of planar WG structures with thick cover layers or so-called buried graded-index WGs this method does not provide reliable measurements. Thick cover layers or large burying depths do not allow tunneling the modes to the external prism and forming the corresponding spatial m-lines. In these cases, some modes (first of all, the lower-order modes) may be simply missed in examinations by m-line spectroscopy [3, 4]. The greater the burying depth the fewer number of modes can be measured by this method. To avoid missing the modes, a layer-by-layer etching of the sample surface could be applied [5, 6]. However, this procedure have many chances to cut a part of the refractive index profile occupied by the mode fields, and that should lead to distortion of the original WG mode spectrum. The use of nonlinear optical effects like second harmonic generation [4] can be successful only for limited number of optical materials demonstrating high values of the corresponding coefficients.

The paper presents a developed measuring technique named the end-fire mode spectroscopy which provides reliable measurement of the whole mode spectrum [7]. This method is suitable for examination of planar WGs having arbitrary refractive index profiles including the case of buried WG structures with any burying depths. Furthermore, here is shown that this technique allows also conducting direct measurements of another important characteristic — the maximal refractive index in graded-index WGs, unlike conventional techniques that involve the set of measured mode indices and employ computing of the maximal value in the refractive index profile using different approximations.

2. METHOD CONTENT

The end-fire mode spectroscopy technique of mode index measurement is based on the registration of light beams radiated from the abrupt output edge of a planar WG, with each beam corresponding to the individual mode. Due to the different values of mode propagation constants, modes of different orders demonstrate different refraction angles at the output face of WG if they are directed to this face under the same inclination angle. It is this feature that is exploited by the technique. Both excitation and output of WG modes are performed at WG faces by the end-fire coupling method. This coupling method allows reliable launching and output of the whole mode spectrum in any planar WG. Therefore, the proposed technique can be applied to the analysis of mode spectra of arbitrary planar WGs, including ones with symmetric step-index profiles and deep-buried graded-index WGs.

The experimental scheme is shown on Figure 1. A collimated incident beam is focused on the input WG face by the cylindrical lens. The whole mode spectrum can be launched by this manner

in few-mode WGs. In the case of a thick multimode WG, a group of modes is excited simultaneously and can be registered. Further scanning the input sample face along the Y axis allows launching and registration of other mode groups until the whole WG mode spectrum is measured.

Skew incidence of the light beams to the sample output face can be provided by different manners, and two variants of the scheme have been proposed [7]. The former one uses the samples having mutually tilted opposite (input and output) WG faces as it is shown on Figure 2. The latter one allows testing WGs having the usual rectangular form due to the used procedure of launching WG modes which is performed by focusing the probe beam on the polished side face of the WG. Whereas this alternative variant is non-destructive, we used in experiments the former scheme because it is more convenient for measuring. Let us consider just the used former scheme. The Figure 2 demonstrates that application of a cylindrical lens for launching the modes enables one to obtain light beams being collimated in their longitudinal XZ-section as in the WG as on light emission from the sample output face. In the longitudinal YZ plane, which is normal to the WG surface, the output beams have large divergence (see Figure 1). Therefore, these beams (each associated with the individual WG mode) appear on the cross screen as separate narrow light strips. The mode orders of the corresponding WG modes are identified by the inclination of a certain output beam to the output WG face: the lower the mode order, the bigger the output angle. The fundamental mode forms the light beam having maximal value of the output angle $\psi_{\max}$.
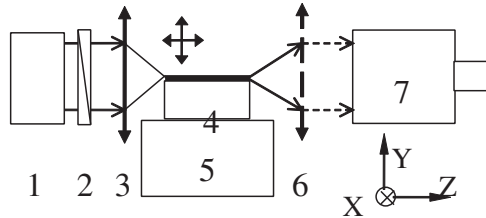

Figure 1: Experimental scheme. 1 — laser with collimator, 2 — polarizer, 3 and 6 — cylindrical lenses, 4 — examined WG sample, 5 — goniometer sample mount, 7 — sighting tube or photoreceiver matrix on the goniometer alidade.
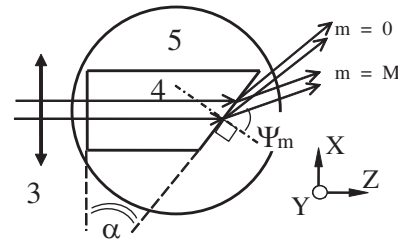
Figure 2: Planform projection of the goniometer mount with the WG sample and beam tracing in longitudinal XZ-section under normal light incidence to the input WG face. Designations are the same as in Figure 1.

Determination of beam traces in XZ-section leads to the following expression for calculating the mode indices N_m:

$$N_m \sin\{\alpha - \arcsin[(\sin i)/N_m]\} = \sin \psi_m \tag{1}$$

where m is the mode order, i and ψ_m are the incident and output angles of the spatial light beams measured in the XZ-plane, and α is the angle between the input and output WG faces. The solution of this equation can be written as

$$N_m = \sqrt{(\sin \psi_m / \sin \alpha + \sin i / \tan \alpha)^2 + \sin^2 i} \tag{2}$$

Evidently, normal incidence of the probe beam to the input WG face is the simplest scheme that is most suitable for measuring. Just this case is shown on Figure 2. For this scheme, both Equation (1) and formula (2) are reduced to

$$N_m = \sin \psi_m / \sin \alpha \tag{3}$$

Following important requirement is to be met: the generatrix line of the launching cylindrical lens and the WG surface must be parallel. Otherwise, an uncertain variation of the direction of the planar beam occurs leading to a mistake in the measurement.

It can be shown from formula (3) that the angle variation $\Delta\psi_m$ related to the index variation ΔN_m increases with a rise in the tilt angle α. So, in order to provide high method sensitivity we must maximize α. However, when α approaches the value $\alpha_{\lim} = \arcsin(1/N_m)$, total inner reflection appears at the output WG face, and the mode is reflected back to the WG (to the side sample face) instead of being emitted out through the output WG face. Proper choice of the angle α should be carried out considering the maximal refractive index which is expected to be obtained for the WG by the applied method of WG fabrication. Slightly overestimated value of this refractive

index is substituted for N_m to the formula for α_{lim}. Calculated in such a manner this angle limit is the optimal value of α.

Let us consider a procedure of direct measurement of the maximal refractive index in a graded-index WG. A principle of this procedure as well as a view of the obtained output light pattern can be explained involving a well-known ray approximation of WG mode propagation that is illustrated by Figure 3 where several rays associated with the same mode are plotted. It is seen that ray 3 reaches the output WG face at the so-called turning point of the ray curve and leaves the WG being directed parallel to the sample surface. According to the considered approach, this point corresponds to the depth in the refractive index profile where the index value is equal to the mode index in graded-index WG. So, just these rays are to be registered in mode spectra measurements. Other rays plotted are tilted to the abrupt WG edge in the YZ plane, and they define a divergence of the output light beam in this plane. Opposite inclinations to the output face demonstrating by arbitrary pair of rays 2 and 4 mean that the output beam is symmetric regarding the ray 3. Furthermore, the rays 1–3 leave the sample at points of different depths and, consequently, of different refractive index values. Therefore, projections of these emitted rays on the XZ plane have different directions in this plane. So, we see the total view of the whole output mode beam as an axisymmetric curved light strip when considered on the cross screen apart from the sample. As we must execute mode spectrum measurements registering the ray 3 which is the symmetry axis of the pattern, we find this ray as the top of the curved output beam associated with the certain WG mode.
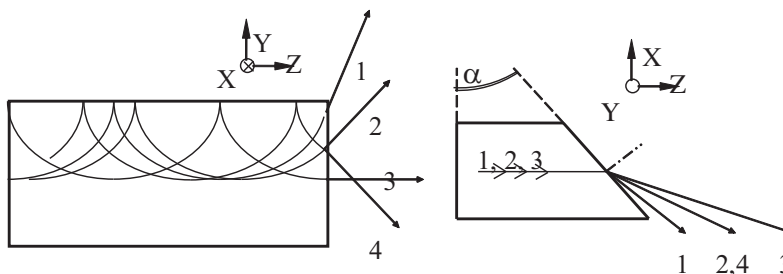

Figure 3: Ray approximation of mode propagation in graded-index planar WG. Plotted rays are associated with the same WG mode.

In measuring the maximal refractive index we should register the ray 1 which leaves the WG at the point of maximal refractive index in graded-index layer. The measured output angle of this boundary ray of the emitted light beam can be used for determination of the maximal index value. Ray tracing performed for this ray shows that the maximal refractive index N_0 can be determined solving the equation

$$N_0 \cos\{\arcsin[(N_m/N_0)\cos\alpha]\} = \sin\varphi \tag{4}$$

where N_m is the mode index of the considered WG mode, and φ is the output angle of the boundary ray. This angle φ is to be measured in the plane that include both boundary ray and a normal to the output sample face as intersecting lines, and this plane is tilted to the XZ plane. However, in the measuring scheme shown on Figure 1 the goniometer measures the angles lying just in the XZ plane. So, we obtain in measuring the values of the angle φ_{xz} which is the projection of the angle φ to the XZ plane. Relation between these angles is presented by the expression

$$\tan\varphi = \tan\varphi_{XZ}\sqrt{[1+(d/2f)^2]\,[(N_0/N_m)^2 - \cos^2\alpha]}/\sin\alpha \tag{5}$$

where f is the focus length of the output cylindrical lens (see Figure 1), and d is the size of the output curved light strip measured along the Y axis behind this lens which collimates the output beam. Let us note that such collimation of the output beam is needed only in measurements of the maximal refractive index when one must register high-divergent boundary rays. Measurement of the mode spectrum is executed by registering the central part of the output light pattern, and it does not matter in this case is the output cylindrical lens applied or not. Considering the formula (5) and also a well-known relation $\sec^2\varphi = 1 + \tan^2\varphi$, one obtain from (5) the final equation containing the only unknown N_0. Numerical solution of the obtained equation give the desired value of the maximal refractive index in the graded-index WG.

Thus, both the mode spectrum and the maximal refractive index in the planar WG can be measured in a single procedure by the end-fire spectroscopy. Due to application of the end-fire

coupling method for both excitation and output of the modes, this measuring technique is capable of providing reliable measurements of the whole mode spectrum in planar WGs of any type, including deep-buried ones.

3. EXPERIMENTAL RESULTS

Comparative measurements of the WG mode spectra were performed by the end-fire mode spectroscopy technique and m-line spectroscopy. Figure 4 presents a photo of the typical pattern formed by output light beams on the cross screen placed apart from the sample when the examination is performed by the developed technique. The whole spectrum of TE modes was launched simultaneously in this experiment, and the output cylindrical lens was removed. The light strips in the pattern really demonstrate some curvature due to the graded-index profile in the examined WG layer. So, the central part of the pattern (the top of the parabolic-like light curves) is to used in measuring the output angles ψ_m. The insert in Figure 4 shows the enlarged view of these central parts of the curves formed by different modes, and it is seen that the light strips are well separated and can be easily registered.

Figure 4: Typical view of output light beams on the cross screen apart from the planar WG and the enlarged central part of this pattern (see insert). Both photos are negatives.
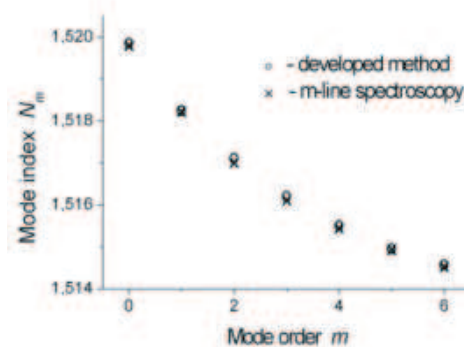
Figure 5: The WG mode spectrum measured both by the developed method and m-line spectroscopy.

As an example, we present the results of examination of the planar WG which was fabricated on a substrate of commercial sodium-containing glass K8 by ion exchange in a potassium nitrate melt at 400°C. Results of comparative measurements of the spectrum of TE modes are presented in Figure 5. The operational light wavelength was equal to 633 nm and the angle between the opposite WG faces was measured with the goniometer by the autocollimation method as $\alpha = 38°57'08'' \pm 5''$. A good agreement between the mode index values measured by the proposed technique and m-line spectroscopy is evident. The difference between the results obtained by these methods does not exceed 10^{-4}, which is similar to the error level considered quite acceptable in mode index measurements.

Direct measuring the maximal refractive index in a graded-index WG was conducted using the same sample. For comparison, the maximal refractive index was computed according to conventional methods using the measured mode spectrum. The White-Heidrich (WH) method [8] gives the result as $N_{0,\text{WH}} = 1.52204$, and the Chiang method [9] is resulted in $N_{0,Ch} = 1.52138$. So, these widespread computing techniques give different results for the same WG. Basing on the previous experience we can guess that application of the WH technique is more appropriate to the case of the used WG fabrication technology. The direct measurement executed by the presented technique resulted in the value $N_0 = 1.5223$ when the highest-order mode was registered. This value is even closer to the result of the WH method than the solution of the Chiang method. It means that the developed technique provides direct measuring the maximal refractive index in graded-index WGs with good accuracy.

Thus, application of the end-fire mode spectroscopy allows conducting reliable direct measurements of the set of important optical characteristics of arbitrary planar WGs (the mode spectrum and the maximal refractive index) in a single procedure.

4. CONCLUSIONS

The obtained results of comparative measurements performed by the end-fire mode spectroscopy and other methods prove a feasibility of the developed technique. This technique enables conducting direct measuring the mode spectrum and the maximal refractive index in graded-index planar WG in a single procedure. Due to the used manner of launching and output of the modes, the end-fire mode spectroscopy allows characterization of arbitrary planar WGs without any restrictions on burying depth and cover layer thickness.

ACKNOWLEDGMENT

REFERENCES

1. Tien, P. K. and R. Ulrich, "Theory of prism-film coupler and thin-film light guides," *J. Opt. Soc. Am.*, Vol. 60, No. 10, 1325–1337, 1970.
2. Barnosky, M. K., *Introduction to Integrated Optics*, Plenum Press, NY-London, 1974.
3. Rogozinski, R., "Determination of refractive index profiles of planar buried waveguides on the basis of a set of modal propagation constants," *Opt. Commun.*, Vol. 219, No. 1, 199–214, 2003.
4. Marangoni, M., R. Osellame, R. Ramponi, and E. Giorgetti, "Second harmonic generation from radiation to guided modes for the characterization of reverse-proton-exchanged waveguides," *Optics Express*, Vol. 12, No. 2, 294–298, 2004.
5. Glebov, L. B., V. G. Dokucgaev, and I. S. Morozova, "Simple method of reconstruction of refractive index profile in planar waveguides," *Optics and Spectroscopy*, Vol. 66, No. 5, 1110–1114, 1989.
6. Monir, M., H. El-Refaei, and D. Khalil, "Single-mode refractive index reconstruction using an NM-line technique," *Fiber and Integrated Optics*, Vol. 25, No. 2, 69–74, 2006.
7. Svistunov, D. V., "End-fire mode spectroscopy technique of examination of planar waveguides," *J. Optics A: Pure and Applied Optics*, Vol. 10, No. 8, 085301 (4pp), 2008.
8. White, J. M. and P. F. Heidrich, "Optical waveguide refractive index profiles determined from measurement of mode indices: A simple analysis," *Applied Optics*, Vol. 15, No. 1, 151–155, 1976.
9. Chiang, K. S., "Construction of refractive-index profiles of planar dielectric waveguides from the distribution of effective indexes," *J. Lightwave Technol.*, Vol. 3, No. 2, 385–391, 1985.

Time-domain Analysis of Wideband Optical Pulse SHG in Layered Dispersive Material

M. A. Alsunaidi[1] and F. S. Al-Hajiri[2]

[1]King Fahd University of Petroleum & Minerals, Saudi Arabia
[2]King Faisal University, Saudi Arabia

Abstract— A time-domain nonlinear model representing the propagation of ultra-short optical pulses in dispersive layered structures containing second order nonlinearity is presented. The model accounts for material dispersion through Lorentz model in addition to the spatiotemporal coupling between the pulsed fundamental and second harmonic beams. Numerical results for the generated second harmonic pulse clearly demonstrate the pulse break-up phenomenon.

1. INTRODUCTION

Time-domain analysis of Second Harmonic Generation (SHG) in modern optical devices provides an invaluable insight into the understanding and potential utilization of this phenomenon and wave-device interaction in general. For CW inputs or long pump pulses, it is quite sufficient to match the phase velocities of the propagating fundamental and second harmonic beams for optimum conversion. As the bandwidth of the pump (fundamental) pulse increases, the group velocity mismatch and wave packet spreading need to be taken into consideration. These effects become significant for pulse durations less than 100 femtoseconds. The earlier attempts to model and simulate pulsed SHG in second order materials involved 1-D approximations with high-order time derivatives and input depletion ignored to simplify the analysis [1–3]. These attempts, in general, have ignored material dispersion.

In this paper, the analysis of ultrashort optical pulse propagation in dispersive second order non-linear materials is introduced. The formulation of the problem involves incorporating the frequency-dependent material dispersion in the time-domain SHG model. For simulation purposes, explicit and uncoupled FDTD solution equations are obtained. The present model fully accounts for all temporal and spatial variations and takes the depletion of the input beam into consideration.

2. FORMULATIONS

The material is assumed to have the general Lorentzian dispersion relation of the form

$$\varepsilon_r(\omega) = \varepsilon_\infty + \frac{(\varepsilon_s - \varepsilon_\infty)\,\omega_o^2}{\omega_o^2 + 2j\omega\delta + (j\omega)^2} \tag{1}$$

where ε_s and ε_∞ are the static and optical values of the dielectric constant, respectively, ω_o is the material resonance frequency and δ is the damping factor. The frequency dependence in the dispersion relation is accommodated in the time domain model through an auxiliary differential equation relating the electric flux density to the field intensity [4]. For the fundamental wave, the frequency domain equation for the electric flux density is given by

$$D^f(\omega) = \varepsilon_o\varepsilon_r(\omega)f(\omega) + \varepsilon_o\chi^{(2)}(fs)(\omega) \tag{2}$$

where the second term in the right hand side represents the contribution of the second order nonlinear polarization. Using Equation (1), Equation (2) becomes

$$\begin{aligned}
\omega_o^2 D^f + 2\delta(j\omega)D^f + (j\omega)^2 D^f &= \varepsilon_o\varepsilon_s\omega_o^2 f + 2\varepsilon_o\varepsilon_\infty\delta(j\omega)f + \varepsilon_o\varepsilon_\infty(j\omega)^2 f \\
&\quad + \varepsilon_o\chi^{(2)}\omega_o^2 fs + 2\varepsilon_o\chi^{(2)}\delta(j\omega)fs + \varepsilon_o\chi^{(2)}(j\omega)^2 fs
\end{aligned} \tag{3}$$

Taking the inverse Fourier transform of Equation (3), the following time-domain differential equation is obtained.

$$\begin{aligned}
\omega_o^2 D^f + 2\delta\frac{\partial D^f}{\partial t} + \frac{\partial^2 D^f}{\partial t^2} &= 2\varepsilon_o\varepsilon_\infty\delta\frac{\partial f}{\partial t} + \varepsilon_o\varepsilon_\infty\frac{\partial^2 f}{\partial t^2} + \varepsilon_o\varepsilon_s\omega_o^2 f + \varepsilon_o\chi^{(2)}\omega_o^2 fs \\
&\quad + 2\varepsilon_o\chi^{(2)}\delta\frac{\partial(fs)}{\partial t} + \varepsilon_o\chi^{(2)}\frac{\partial^2(fs)}{\partial t^2}
\end{aligned} \tag{4}$$

Similarly, for the second harmonic wave, with

$$D^s(\omega) = \varepsilon_o\varepsilon_r(\omega)s(\omega) + \varepsilon_o\chi^{(2)}f^2(\omega) \tag{5}$$

it can be shown that

$$
\omega_o^2 D^s + 2\delta\frac{\partial D^s}{\partial t} + \frac{\partial^2 D^2}{\partial t^2} = 2\varepsilon_o\varepsilon_\infty\delta\frac{\partial s}{\partial t} + \varepsilon_o\varepsilon_\infty\frac{\partial^2 s}{\partial t^2} + \varepsilon_o\varepsilon_s\omega_o^2 s
$$
$$
+ \varepsilon_o\chi^{(2)}\omega_o^2 f^2 + 2\varepsilon_o\chi^{(2)}\delta\frac{\partial f^2}{\partial t} + \varepsilon_o\chi^{(2)}\frac{\partial^2 f^2}{\partial t^2} \tag{6}
$$

The fundamental and second harmonic flux densities are related to the electric field intensities, respectively, by

$$\frac{\partial^2 D^f}{\partial t^2} = \frac{1}{\mu o}\nabla^2 f \tag{7}$$

$$\frac{\partial^2 D^s}{\partial t^2} = \frac{1}{\mu o}\nabla^2 s \tag{8}$$

The resulting dispersive SHG model equations are strongly-coupled highly-nonlinear set of partial differential equations. These equations provide a time-dependent self-consistent solution for the propagation of the input wave as well as the subsequent generation of the second harmonic wave for any input conditions. Considering the sensitivity of the SHG problem to phase differences between the propagating waves, the equations need to be solved in their complete form.

3. SOLUTION METHOD AND NUMERICAL RESULTS

The model equations are approximated using an explicit FDTD scheme. The solution algorithm starts by first finding all D^{n+1} components using Equations (7) and (8). Next the updated value for f is calculated. Finally, Equation (6) is used to update the second harmonic field. To study the effectiveness of the proposed model, a symmetric AlGaAs-based dielectric slab waveguide is considered to test the proposed FDTD algorithm. It consists of a 0.44-µm thick guiding layer sandwiched between two 3-µm thick AlAs layers. The dispersion relation of the material is obtained from published experimental data [5] and fitted to Lorentz model of Equation (1). With $\varepsilon_\infty = 1.24$, $\varepsilon_s = 2.88$, $\omega_o = 0.65 \times 10^{16}\,\text{rad/s}$ and $\delta = 6.0 \times 10^{13}\,\text{rad/s}$, a very good agreement is achieved, as shown in Figure 1. The non-linear susceptibility is taken as $\chi^{(2)} = 113\,\text{pm/V}$. The excitation field is a Gaussian-modulated CW signal at a fundamental wavelength of 1.064 µm. The transverse profile of the excitation corresponds to the first guided mode at the given operating frequency. Due to group velocity mismatch (GVM) phenomenon, the power exchange between the propagating
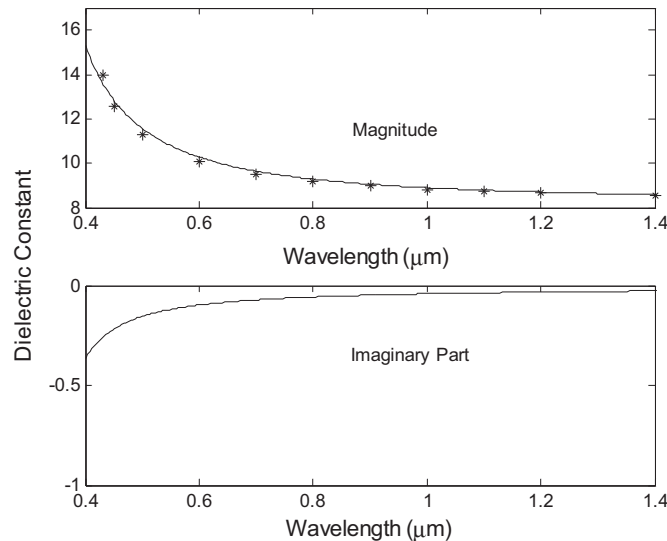

Figure 1: The material dispersion relation. Solid lines: Lorentz fit, *: experimental data [5].

input pulse and the generated second harmonic pulse diminishes along the propagation direction as shown in Figure 2. The initially generated second harmonic pulse near the input side of the waveguide eventually breaks up after propagating a curtain distance along the waveguide. The higher frequency components travel slower and lag the fundamental pulse. Eventually, they stop interacting with the fundamental wave and propagate independently along the waveguide. This behavior is very clearly exhibited as the pulse duration is made shorter. The power exchange process becomes similar to the CW case for longer pulses (see Figure 3). The generated second harmonic pulse breaks up after propagating a curtain distance along the waveguide. This process takes around 50 μm to complete for a 15 fs pulse input, as shown in Figure 4, but longer distance for wider pulses. As the pulse waists become shorter, the dispersion effects become more pronounced. Inversely, as the pulse waist becomes larger, the results converge to the CW case.
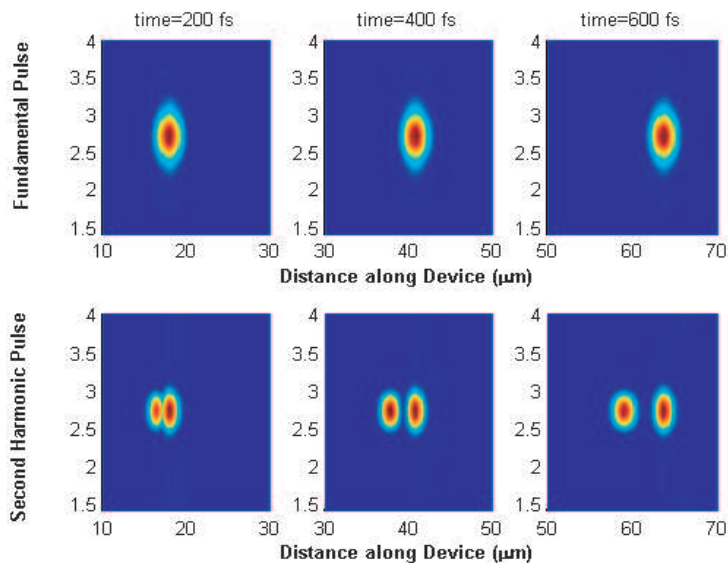

Figure 2: The fundamental and second harmonic fields inside the dispersive nonlinear waveguide. Initial pulse waist is 15 fs.
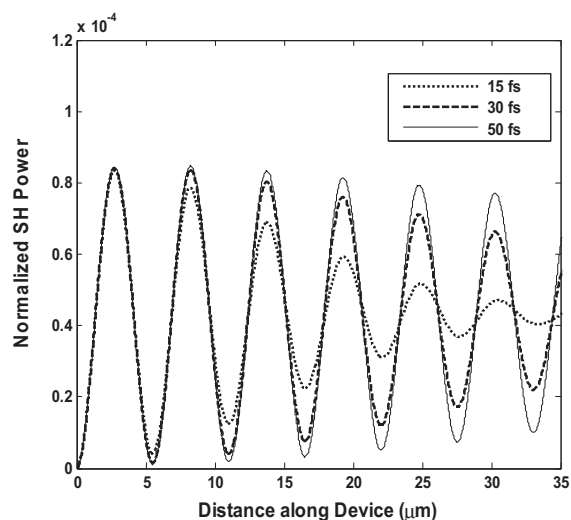

Figure 3: Normalized second harmonic power along the nonlinear, dispersive and lossless waveguide for different initial pulse waists.
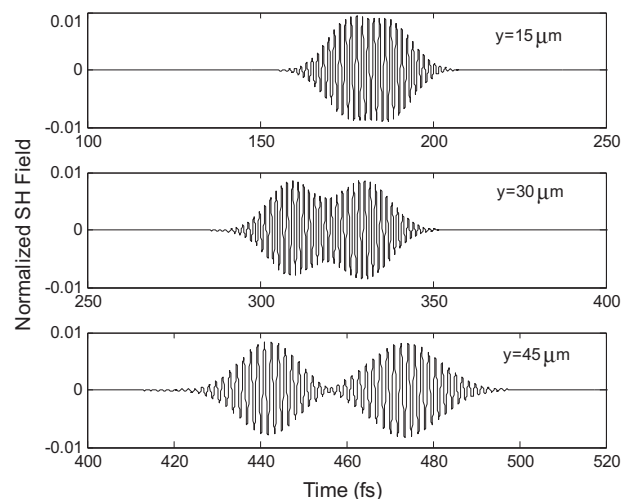
Figure 4: Second harmonic pulse time profile in a nonlinear, dispersive and lossless waveguide at three locations along the propagation direction. Initial pulse waist is 15 fs.

4. CONCLUSIONS

The effect of material dispersion on the efficiency of pulsed second harmonic generation in optical structures containing second order nonlinearities is modeled. The frequency-dependent material response is incorporated into the time-domain model using the ADE approach. The application of this FDTD algorithm for modeling short-pulse SHG in dispersive nonlinear optical waveguides has been demonstrated. Numerical results for the generated second harmonic pulse clearly demonstrate the pulse break-up phenomenon.

ACKNOWLEDGMENT

REFERENCES

1. Sidick, E., A. Knosen, and A. Dienes, "Ultrashort-pulse second-harmonic generation. I. Transform-limited fundamental pulses," *J. Opt. Soc. Am. B*, Vol. 12, 1704–1712, 1995.
2. D'Orazlo, A., D. de ceglia, M. de Sario, V. Petruzzelli, and F. Prudenzano, "Second order nonlinear interactions in periodic waveguides," *ICTON 2004*, 126–130, 2004.
3. Masoudi, H., "A Time-domain beam propagation method for analyzing pulsed optical beams in second order nonlinear waveguides," *Microwave and Optical Technology Letters*, Vol. 28, 253–257, 2001.
4. Taflove, A., *Computational Electrodynamics: The Finite-difference Time-domain Method*, Artech Hous, Norwood, MA, 1995.
5. Gehrsitz, S., F. K. Reinhart, C. Gourgon, N. Herres, A. Vonlanthen, and H. Sigg, "The refractive index of $Al_xGa_{1-x}As$ below the band gap: Accurate determination and empirical modeling," *J. Appl. Phys.*, Vol. 87, 7825–7837, 2000.

Intense Terahertz Radiation from GaAs Photoconductive Antenna Array

Wei Shi, **Hong Xue**, **Xiangrong Ma**, **and Zhenzhen Zhang**

Department of Applied Physics, Xi'an University of Technology, Shaanxi 710048, China

Abstract— In this paper, we fabricated an intense photoconductive antenna array with eight antenna units. Each unit has separate electrodes and could be operated independently. The spacing between electrodes and the width of the electrodes are 150 μm and 10 μm, respectively. The substrate is commercial semi-insulating GaAs. The electrode pattern is made using a conventional photolithography technique. Ohmic contacts are obtained by using a standard mixture of Ni/Au-Ge/Au for the metallization and the thickness is 700 nm. The antenna array is mounted in an integrated circuit shell with ceramic substrate. Those antennas are tested by TDS-Z2 system made in Zomega Terahertz Corporation. A fiber laser (IMRA F-100) has been used in the system as the excitation source for the generation of THz radiation. The pump beam has a power of 90 mW, wavelength of 800 nm and the 1/e beam diameter on the emitter surface is about 20 μm, and so only one antenna unit is illuminated in our experiment. The field waveform is measured using an electrooptic sampling method employing a 1-mm-thick ZnTe crystal. To understand the behavior of interdigital antenna array, THz transmission through the ceramic substrate is measured. Only 23% of THz wave transmitted through the substrate. If a hole is drilled through the ceramic substrate to make THz wave go through, the THz amplitude of antenna unit will increase more than 4 times, and the THz amplitude from the antenna unit is 4/3 times larger than the THz amplitude from the stripline antenna at the same electrical field.

1. INTRODUCTION

The development of Terahertz (THz) source is of great importance for a large variety of scientific and technological application [1]. Recently, various methods for generating terahertz electromagnetic radiation have attracted considerable attention. However, from 1984, Auston first demonstrated THz radiation can be generated by femtosecond laser illuminating a biased photoconductive antenna [2], this method is always a standard method until now. And photoconductive antenna is widely used as THz source in THz TDS system, which is used in many fields, such as explosives and chemical detection [3], biology threat detection [4], quality control and non-destructive testing [5], material analysis and testing [6] etc. Various antenna geometries have been used, these include Hertzian dipole, resonant dipole antennas, spiral antennas, antenna array, and so on. The dipolar antennas and resonant dipole antennas are the most common structures. In this paper, we fabricated a photoconductive antenna array with interdigital electrodes on commercial SI-GaAs substrate and compared the performance of a antenna unit of the array with a conventional stripline antenna. And the absorption of ceramic substrate of antenna mount of interdigital antenna array is measured. The results indicate that the antenna unit can generate a higher THz wave and the interdigital antenna array is expected to emit THz wave with amplitude more than 10 times higher than that of the stripline antenna at same electrical field.

2. EXPERIMENT

The photoconductive antenna array with interdigital electrodes is composed of eight photoconductive antenna units. The substrate of antenna is commercial high-resistivity ($> 107\,\Omega\,\mathrm{cm}$) liquid-encapsulated-Czochralsky-(LEC) grown, (100)-oriented, semi-insulating (SI) GaAs wafer. The schematic of the entire antenna array structure is shown in Fig. 1. Each unit has separate electrodes and could be operated independently. The spacing between electrodes and the width of the electrodes are 150 μm and 10 μm, respectively. The electrode pattern is made using a conventional photolithography technique. Ohmic contacts are obtained by using a standard mixture of Ni/Au-Ge/Au for the metallization and the thickness is 700 nm. The antenna array is mounted in an integrated circuit shell with ceramic substrate, and the electrodes are connected with pins by golden wire.

The performance of the emitter is compared with a conventional stripline antenna, which has a 100 μm gap between electrodes, the substrate is also commercial SI-GaAs and the fabricate processing is same to the interdigital antenna array. Those antennas are tested by TDS-Z2 system made in Zomega Terahertz Corporation. A fiber laser (IMRA F-100) has been used in the system as
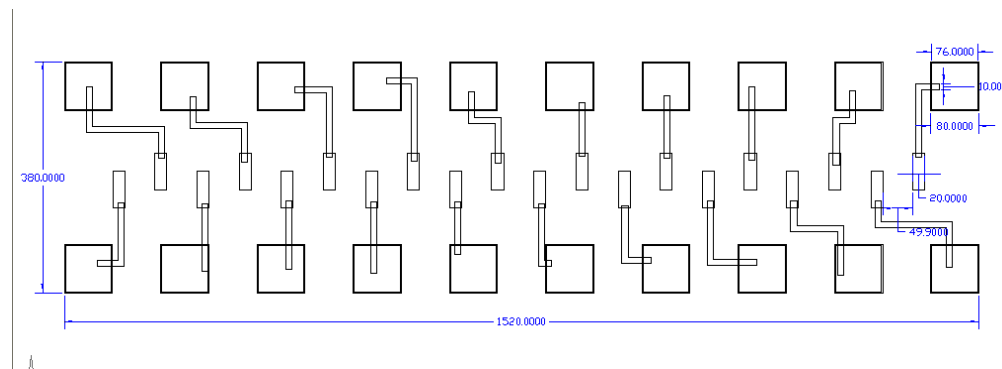
Figure 1: Schematic of interdigital antenna array structure.

the excitation source for the generation of THz radiation. The pump beam has a power of 90 mW, wavelength of 800 nm and the 1/e beam diameter on the emitter surface is about 20 µm. Because the beam size is much smaller than the antenna array, only one antenna unit is illuminated in our experiment. The field waveform is measured using an electrooptic sampling method employing a 2-mm-thick ZnTe crystal. THz field is measured with system optimized for each antenna individually.

3. RESULTS

To compare the performance of the two emitters at same electrical field, the bias voltages of interdigital electrodes antenna array (150 µm gaps) and the conventional stripline antenna (100 µm gap) are 90 V and 60 V, respectively. Figs. 2(a) and (b) show the time-domain waveform and spectra of conventional stripline antenna and one antenna unit of the antenna array at the same electrical field.
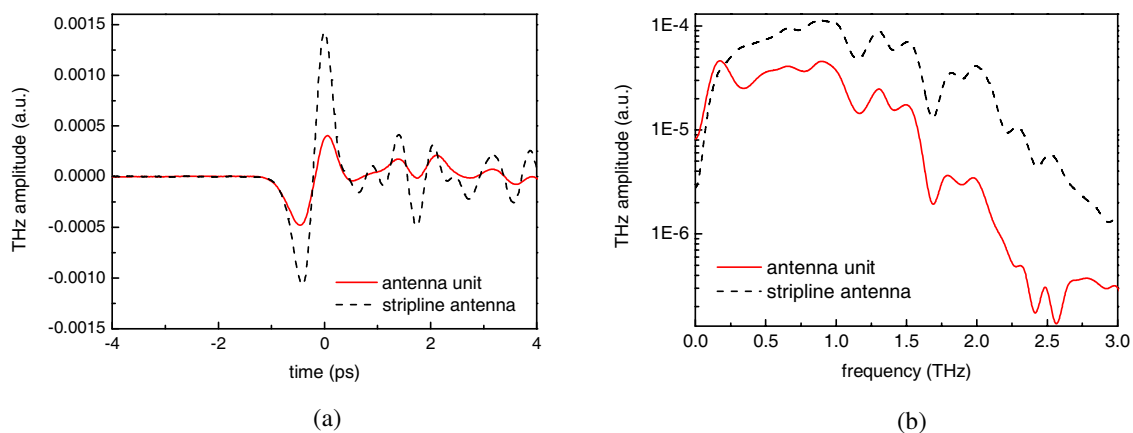

Figure 2: (a) THz time domain waveform from stripline antenna and one antenna unit at same electrical field. (b) Spectra of THz wave from stripline antenna and one interdigital antenna unit at same electrical field.

Stripline antenna emits stronger THz field, whose THz amplitude is nearly 3 times higher than the THz amplitude from the antenna unit. The THz spectrum emitted from stripline antenna is broader, extending to almost 3 THz. While the spectrum from antenna unit shows a cut-off frequency around 1.7 THz. Above 1.7 THz, little THz signal exists. Both the two emitters are made of SI-GaAs and are tested by the same TDS system, the antenna unit should also generate THz wave with 3 THz bandwidth. The reduction is caused by the absorption of ceramic substrate [7].

To understand the behavior of interdigital antenna array, THz transmission through the ceramic substrate is measured. The time-domain waveform is shown in Fig. 3(a). Only 23% of THz wave transmitted through the substrate. Their spectra are shown in Fig. 3(b). Compared with reference wave, THz spectrum shows a strong reduction above 1.7 THz. If a hole is drilled through the ceramic substrate to make THz wave go through, the THz amplitude of antenna unit will increase more than 4 times, and the THz amplitude from the antenna unit is 4/3 times larger than the THz amplitude from the stripline antenna at the same electrical field.

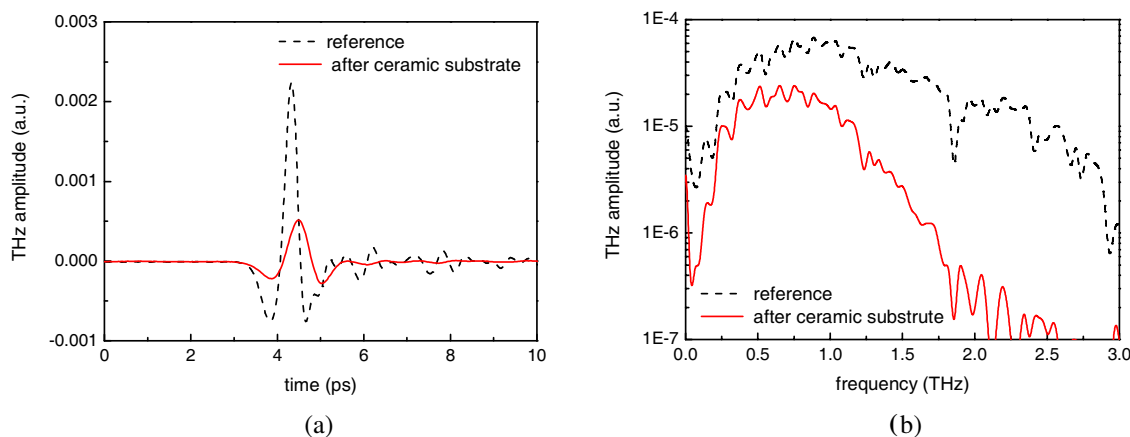(a)   (b)

Figure 3: (a) THz time-domain waveform before and after transmission of ceramic substrate. (b) Spectra of THz wave before and after transmission of ceramic substrate.

Since the THz field from photoconductive antenna is decided by the following equations [8]:

$$E_{\text{THz}}(t) = -\frac{1}{4\pi\varepsilon_0 c^2}\frac{A}{z}\frac{d}{dt}J_s(t) \tag{1}$$

An expression for the surface current can be expressed as:

$$J_s = \frac{\sigma_s(t)E_b}{\frac{\sigma_s(t)\eta_0}{1+n}+1} \tag{2}$$

According the two equations,

$$E_{\text{THz}}(t) \propto AE_b \tag{3}$$

From (3), there are eight antenna units on the chip, if the pump beam size is large enough to illuminate all units and the laser power is also high enough, the THz amplitude of the interdigital antenna array will be improved 8 times at same electrical field. So the amplitude of THz from the whole interdigital antenna array is more than 10 times larger than the THz amplitude from conventional stripline antenna at same electrical field.

4. CONCLUSIONS

In summary, we have fabricated a powerful antenna array with 8 interdigital electrodes, and compared the performance of one antenna unit of the array with that of conventional stripline antenna. However, the THz amplitude of one antenna unit is only 1/3 of that of stripline antenna because 77% THz wave is absorbed by the ceramic substrate. So if drill a hole through the substrate to release the THz wave and illuminate all antenna units using enough powerful pump beam, the THz amplitude will be more than 10 times higher than that from conventional stripline antenna at same electrical field.

ACKNOWLEDGMENT

REFERENCES

1. Dreyhaupt, A., S. Winnerl, T. Dekorsy, and M. Helm, "High-intensity terahertz radiation from a microstructured large-area photoconductor," *Applied Physics Letter*, Vol. 86, 121114, 2005.
2. Auston, D. H., K. P. Cheung, and P. R. Smith, "Picosecond photoconducting Hertzian dipoles," *Applied Physics Letter*, Vol. 45, No. 3, 284, 1984.
3. Liu, H.-B., Y. Chen, G. J. Bastiaans, and X.-C. Zhang, "Detection and identification of explosive RDX by THz diffuse reflection spectroscopy," *Optics Express*, Vol. 14, 415–423, America, January 2006.

4. Kemp, M. C., P. F. Taday, B. E. Cole, J. A. Cluff, A. J. Fitzgerald, and W. R. Tribe, "Security applications of terahertz technology," *Proc. SPIE*, Vol. 5070, 44–52, 2003.

5. Rutz, F., S. Wietzke, M. Koch, et al., "Non-destructive testing of glass-fibre reinforced polymers using terahertz spectroscopy," *9th ECNDT*, 1–7, Berlin, Germany, September 2006.

6. Zhang, X.-C. and D. H. Auston, "Optoelectronic measurement of semiconductor surfaces and interfaces with femtosecond optics," *J. Appl. Phys.*, Vol. 71, 326–338, American, January 1992.

7. Rutz, F., M. Koch, L. Micele, and G. de Portu, "Ceramic dielectric mirrors for the terahertz range," *Applied Optics*, Vol. 45, 8070–8073, America, November 2006.

8. Benicewicz, P. K., J. P. Roberts, and A. J. Taylor, "Scaling of terahertz radiation from large-aperture biased photoconductors," *Optical Letter*, Vol. 11, 2533–2546, America, December 1994.

Electromagnetic Radiation from Organic Light-emitting Diodes

A. Epstein, **N. Tessler**, and **P. D. Einziger**

Department of Electrical Engineering, Technion — Israel Institute of Technology
Haifa 32000, Israel

Abstract— An analytical prototype model for the electromagnetic radiation emitted from a nanometric organic light-emitting diode device is presented and thoroughly investigated herein. The results are obtained via asymptotic evaluation of the resultant radiation integral in conjunction with coherence considerations, resulting in closed-form analytical expressions. For the sake of simplicity and clarity, we focus on a two-dimensional canonical configuration excited by impulsive (line) sources. The resultant expressions can be most effectively utilized by engineers for improved design, as they enable the calculation of the device's physical parameters, such as electrical to optical conversion efficiency and emission angular distribution, as a function of device structure. It should be pointed out that the incorporation of both rigorous electromagnetic analysis and coherence effects is addressed in our report, to the best of our knowledge, for the first time. This results in a precise model capable of repeating and interpreting experimental and simulated data.

1. INTRODUCTION

Organic light emitting diodes (OLEDs) have been intensively investigated for the past two decades as potentially promising candidates for the fabrication of thin and flexible displays as well as other novel optoelectronic devices [1–4]. The relatively simple and cheap manufacturing procedures involved in OLED production, the prospects for wide viewing angle and high luminescence as in inorganic LEDs and the fast response compared to liquid crystal displays (LCDs) have made OLED technology a very attractive one [1, 4]. In recent years a major technological effort has been made in order to design durable efficient OLEDs with wide viewing angle [3, 5–7].

Electromagnetic modeling of OLEDs is crucially important for analysis and synthesis of their radiation pattern and radiation efficiency, thereby determining the devices' performance. Indeed, vast research in the past few years aimed at optical modeling of OLEDs in order to achieve an understanding of how the layer dimensions and material composition alter their electromagnetic properties [5–10].

Most of these attempts rely on the early work of Chance, Prock and Silbey [11] (CPS) which showed that a radiatively decaying excited molecule can be modeled by a classical oscillating dipole, and presented integral expressions for the radiative and non-radiative decay rates of such dipole when embedded in layered media, from which the radiation efficiency is readily achieved. The CPS model uses the Hertz vectors formalism and later the dyadic Green function method [12] to express the possible source excitations and apply the layered media constraints on the electromagnetic fields. Numerically evaluating the integral expressions, they found good agreement between simulation and experimental results for fluorescence lifetime of molecules near metallic interfaces.

The intimate connection between the radiation pattern and coherence properties of the radiation source is widely studied by Wolf, and well summarized in [13]. Intuitively speaking, it is clear that sharp interference patterns in emission angular distribution as well as interference fringes on a distant screen are caused by some consistent phase difference between two sources. As temporal or spatial coherence decrease this phase difference is no longer consistent in time or space, thus averaging on a large ensemble of sources diminishes these effects dramatically.

In this paper we present a rigorous electromagnetic analysis of two-dimensional OLED radiation. The basis of our analysis is the decomposition of the source term into its plane-wave spectrum [12, 14, 15], a method which preserves the fundamental physical intuition of plane-wave optics. Asymptotic evaluation of the resultant radiation integral leads to closed-form analytical expressions, which, along with the plane-wave interpretation, enables clear identification of dominant factors which determine the device's radiation pattern properties. The coherence effects are explicitly incorporated in the electromagnetic model and their impact is readily observed in the simulated results, outlined in Sections 2 and 3, respectively.

2. FORMULATION

Let us consider a two-dimensional device with $M + N + 2$ layers, with a line source embedded at a certain plane, $z = z'$, sandwiched between layers (-1) and $(+1)$, as depicted in Fig. 1. The homogeneous layer formed by combining layers (-1) and $(+1)$, containing the line source, is termed the active layer. Each layer is characterized by its permittivity, permeability, and conductivity marked ε_n, μ_n and σ_n, respectively, for the n-th layer. Furthermore, the n-th and $(n+1)$-th layers are separated by the plane $z = d_n$ for $n > 0$ and $z = d_{n+1}$ for $n < 0$. Note that $\varepsilon_{-1} = \varepsilon_1$, $\mu_{-1} = \mu_1$, and $\sigma_{-1} = \sigma_1$. For the sake of completeness, we treat here both transverse electric (TE) and transverse magnetic (TM) modes, excited via electric line source and magnetic line source (Fig. 1), respectively. Throughout the entire paper we use e and m left superscripts or subscripts to denote electric or magnetic cases, respectively. Both sources are assumed to be time harmonic, with time dependence of $e^{j\omega t}$. The wave number and wave impedance of the n-th layer are given as

$$k_n = \omega \sqrt{\mu_n \varepsilon_n \left[1 - j\sigma_n/(\omega \varepsilon_n)\right]} = (\omega/c)\left(\mathrm{n}_n - j\kappa_n\right), \qquad Z_n = \sqrt{\mu_n/\left\{\varepsilon_n \left[1 - j\sigma_n/(\omega \varepsilon_n)\right]\right\}}, \quad (1)$$

where c, n and κ denote the velocity of light in vacuum, refractive index and extinction coefficient, respectively. To satisfy the radiation condition we require $\Im\{k_n\} \leq 0$, leading to $\Im\{Z_n\} \geq 0$. Furthermore we define the two-dimensional space vector, $\vec{\rho} = \rho_t \hat{t} + z\hat{z} = (-\rho \sin\theta)\hat{t} + (\rho \cos\theta)\hat{z}$, where ρ_t and $\hat{t}$ are its transverse coordinate magnitude and direction, and θ is the angle between the z-axis and $\vec{\rho}$. The transverse coordinate is different for the electric and magnetic cases due to the different symmetry they induce. In the electric line source scenario, there is no change along the x direction, therefore $^e[\partial/\partial x] = 0$, $^e\rho_t = y$ and $^e\hat{t} = \hat{y}$. Analogously, for the magnetic line source scenario we have symmetry along the y-axis, thus, $^m[\partial/\partial y] = 0$, $^m\rho_t = x$ and $^m\hat{t} = \hat{x}$. The source vector in both cases is $\vec{\rho}' = z'\hat{z}$.
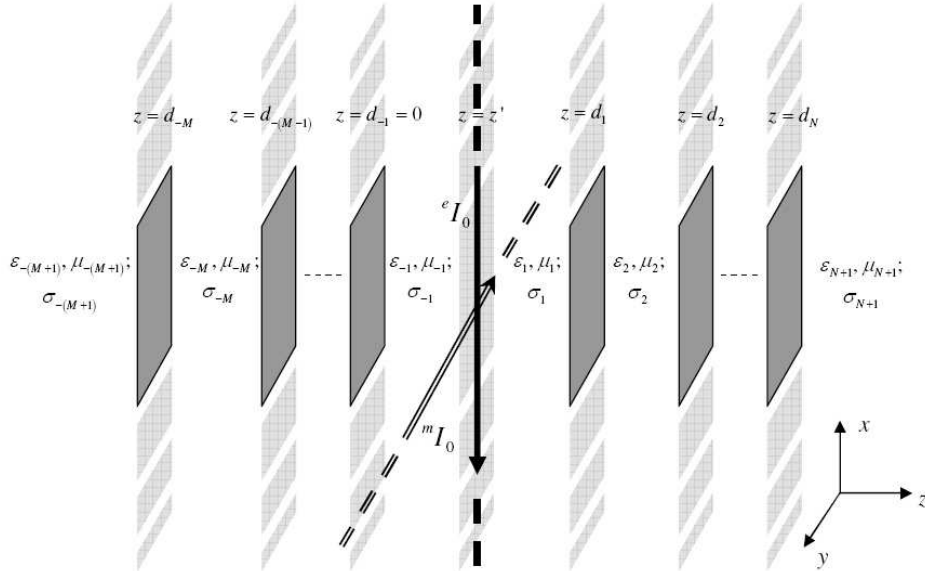

Figure 1: Two-dimensional configuration for the OLED model. The device consists of $M + N + 2$ layers, the interfaces of which are planes parallel to the $\widehat{xy}$ plane and the propagation direction is $\hat{z}$. The two-dimensional model excitation is a line source, sandwiched between layers (-1) and $(+1)$, which form together the active layer.

The transverse electromagnetic field components can be expressed [15] via the following spectral integrals

$$E_x(\vec{\rho}, \vec{\rho}') = \frac{1}{2\pi} \int_{-\infty}^{\infty} e_x(z, z') \cdot e^{j\vec{k}_t \cdot \vec{\rho}}, \qquad H_y(\vec{\rho}, \vec{\rho}') = \frac{1}{2\pi} \int_{-\infty}^{\infty} h_y(z, z') \cdot e^{j\vec{k}_t \cdot \vec{\rho}}, \qquad (2)$$

where $e_x(z, z')$, $h_y(z, z')$ and $\vec{k}_t = k_t \hat{t}$, denote the plane-wave spectrum and transverse wave vector, respectively. The transverse plan-wave spectral amplitudes can be readily expressed in terms of

the one-dimensional Green's function, $g(z, z')$,

$$e_x(z, z') = j\beta\widetilde{Z}J_s g(z, z') - M_s \frac{dg(z, z')}{dz}, \qquad h_y(z, z') = -J_s \frac{dg(z, z')}{dz} + j\beta\widetilde{Y}M_s g(z, z'), \qquad (3)$$

where

$$\begin{smallmatrix}e\\m\end{smallmatrix}\widetilde{Z} = 1/\begin{smallmatrix}e\\m\end{smallmatrix}\widetilde{Y} = Z(k/\beta)^{\pm 1} \quad \text{and} \qquad {}^e J_s = {}^e I_0, \quad {}^m M_s = {}^m I_0, \quad {}^m J_s = {}^e M_s = 0. \qquad (4)$$

The wave equation and associated constraints for $g(z, z')$ are outlined in [12, 14, 15]. Note that $g_n(z, z')$ specifies $g(z, z')$ at the n-th layer. Expanding the work presented in [14] we express the Green function in the various layers in terms of reflection, transmission and source coefficients, as demonstrated in Table 1. The recursive expressions for the total reflection and transmission coefficients are derived via the constraints listed in [15], to the right (forward) and to the left (reversed) of the active layer, for both electric and magnetic polarizations.

Table 1: Recursive relations of the one-dimensional Green function.

	Forward Direction $(z > z')$	Reversed Direction $(z < z')$
One-dimensional Green function	$g_n(z, z') = \dfrac{e^{j\beta_1 z'}}{2j\beta_n} \cdot \dfrac{1 - \widehat{R}_{-1}(k_t) e^{-2j\beta_1 z'}}{1 - R_1(k_t)\,\widehat{R}_{-1}(k_t)}\left[\displaystyle\prod_{p=2}^{n} T_p(k_t)\right]\left[e^{-j\beta_n z} - R_n(k_t)\,e^{j\beta_n z}\right]$	$g_n(z, z') = \dfrac{e^{-j\beta_1 z'}}{2j\beta_n} \cdot \dfrac{1 - R_1(k_t) e^{2j\beta_1 z'}}{1 - R_1(k_t)\,\widehat{R}_{-1}(k_t)}\left[\displaystyle\prod_{p=n}^{-2}\widehat{T}_p(k_t)\right]\left[e^{j\beta_n z} - \widehat{R}_n(k_t)\,e^{-j\beta_n z}\right]$
Total reflection coefficient	$R_n(k_t) = \left\{\Gamma_n(k_t) + \dfrac{\left[1 - \Gamma_n^2(k_t)\right]R_{n+1}(k_t)\,e^{2j\beta_{n+1}d_n}}{1 + \Gamma_n(k_t)R_{n+1}(k_t)\,e^{2j\beta_{n+1}d_n}}\right\}e^{-2j\beta_n d_n}$	$\widehat{R}_n(k_t) = \left\{\widehat{\Gamma}_n(k_t) + \dfrac{\left[1 - \widehat{\Gamma}_n^2(k_t)\right]\widehat{R}_{n-1}(k_t)\,e^{-2j\beta_{n-1}d_n}}{1 + \Gamma_n(k_t)\widehat{R}_{n-1}(k_t)\,e^{-2j\beta_{n-1}d_n}}\right\}e^{2j\beta_n d_n}$
Recursion base condition	$R_{N+1}(k_t) = 0$	$\widehat{R}_{-(M+1)}(k_t) = 0$
Total transmission coefficient	$T_n(k_t) = \dfrac{\left[1 + \Gamma_{n-1}(k_t)\right]e^{j\left(\beta_n - \beta_{n-1}\right)d_{n-1}}}{1 + \Gamma_{n-1}(k_t)R_n(k_t)\,e^{2j\beta_n d_{n-1}}}$	$\widehat{T}_n(k_t) = \dfrac{\left[1 + \widehat{\Gamma}_{n+1}(k_t)\right]e^{-j\left(\beta_n - \beta_{n+1}\right)d_{n+1}}}{1 + \widehat{\Gamma}_{n+1}(k_t)\widehat{R}_n(k_t)\,e^{-2j\beta_n d_{n+1}}}$
Local reflection coefficient	$\Gamma_n = \dfrac{1 - \gamma_n}{1 + \gamma_n}$	$\widehat{\Gamma}_n = -\Gamma_{n-1} = \dfrac{1 - \widehat{\gamma}_n}{1 + \widehat{\gamma}_n}$
Impedance ratio	$\begin{smallmatrix}e\\m\end{smallmatrix}\gamma_n = \left(\dfrac{\widetilde{Z}_{n+1}}{\widetilde{Z}_n}\right)^{\pm 1}$	$\begin{smallmatrix}e\\m\end{smallmatrix}\widehat{\gamma}_n = 1/\begin{smallmatrix}e\\m\end{smallmatrix}\gamma_{n-1} = \left(\dfrac{\widetilde{Z}_{n-1}}{\widetilde{Z}_n}\right)^{\pm 1}$

Power relations associated with our OLED layered model are crucially important for evaluating and optimizing the device performance. The most significant relations are the so-called radiation (emission) pattern, radiation power and absorption power, given via

$$S_\rho(\theta, \omega) = \left.\frac{1}{2}\left[\vec{E}\left(\vec{\rho}, \vec{\rho}'\right) \times \vec{H}^*\left(\vec{\rho}, \vec{\rho}'\right)\right] \cdot \hat{\rho}\right|_{\rho \to \infty}, \qquad (5)$$

$$S_r(z, \omega) = \int_{-\infty}^{\infty} dy \left\{\left.\frac{1}{2}\left[\vec{E}\left(\vec{\rho}, \vec{\rho}'\right) \times \vec{H}^*\left(\vec{\rho}, \vec{\rho}'\right)\right] \cdot \hat{z}\right|_{z > z'}\right\} \qquad (6)$$

and

$$S_a(z, \omega) = \int_{-\infty}^{\infty} dy \left\{\left.\frac{1}{2}\left[\vec{E}\left(\vec{\rho}, \vec{\rho}'\right) \times \vec{H}^*\left(\vec{\rho}, \vec{\rho}'\right)\right] \cdot (-\hat{z})\right|_{z < z'}\right\}, \qquad (7)$$

respectively. Substituting (2)–(3) in conjunction with Table 1 into (5)–(7) renders closed-form expressions for the power relations, as summarized in Table 2.

Apparently, coherence effects must be considered when analyzing OLED device, as may be concluded from [9, 10]. Since spatial coherence can, in general, be neglected when considering

OLEDs [16], we take into consideration only the temporal coherence effect on the power relation [13], i.e.,

$$\mathcal{S}(\vec{\rho}) = \int_0^\infty p(\omega) \, S(\vec{\rho}, \omega) \, d\omega \tag{8}$$

where $S(\vec{\rho}, \omega)$ denotes any of the power relations given in (5)–(7), and $p(\omega)$ is the coherence function of the device, i.e., the appropriately normalized free-space photoluminescence (PL) spectrum. The Gaussian distribution, being simple analytic and effectively bounded function, is selected as the coherence function for our simulations. The Gaussian width, $\Delta\omega$, is inversely proportional to the coherence length [13], L_c, given via, $L_c = 2\pi c/\Delta\omega = \lambda^2/\Delta\lambda$.

Table 2: Power relations for the OLED model. The line source power, P_n, denotes the radiation power of a line source in an unbounded homogeneous medium (the n-th layer).

Radiation pattern, Eq. (5)	$S_\rho(\theta, \omega) = \dfrac{P_{N+1}}{\pi\rho} \left\| \left\{ \dfrac{e^{j\beta_1 z'} - \widehat{R}_{-1}(k_t)\, e^{-j\beta_1 z'}}{1 - R_1(k_t)\,\widehat{R}_{-1}(k_t)} \left[\prod_{p=2}^{N+1} T_p(k_t) \right] \right\} \Big|_{k_t = k_{N+1}\sin\theta} \right\|^2$
Radiation power, Eq. (6)	$S_r^n(z, \omega) = \dfrac{P_n}{\pi} \int_{-\infty}^{\infty} dk_t \left\{ \dfrac{e^{2\Im\{\beta_n\}z}}{\beta_n} \left| \dfrac{e^{j\beta_1 z'} - \widehat{R}_{-1}(k_t)\, e^{-j\beta_1 z'}}{1 - R_1(k_t)\,\widehat{R}_{-1}(k_t)} \left[\prod_{p=2}^{n} T_p(k_t) \right] \right|^2 \left[1 - R_n(k_t)\,e^{2j\beta_n z} \right]\left[1 + R_n(k_t)\,e^{2j\beta_n z} \right]^* \right\}$
Absorption power, Eq. (7)	$S_a^n(z, \omega) = \dfrac{P_n}{\pi} \int_{-\infty}^{\infty} dk_t \left\{ \dfrac{e^{-2\Im\{\beta_n\}z}}{\beta_n} \left| \dfrac{e^{-j\beta_1 z'} - R_1(k_t)\, e^{j\beta_1 z'}}{1 - R_1(k_t)\,\widehat{R}_{-1}(k_t)} \left[\prod_{p=n}^{-2} \widehat{T}_p(k_t) \right] \right|^2 \left[1 + \widehat{R}_n(k_t)\,e^{-2j\beta_n z} \right]\left[1 - \widehat{R}_n(k_t)\,e^{-2j\beta_n z} \right]^* \right\}$
Line source power, P_n	$^e P_n = \dfrac{k_n Z_n |^e I_0|^2}{16}; \quad ^m P_n = \dfrac{k_n Y_n |^m I_0|^2}{16}$

3. RESULTS

The potential promise of our modeling approach, outlined in the previous section, is demonstrated through radiation pattern simulations, incorporating coherence effects, of a prototype device. A basic configuration of a five layer bottom-emitting (BE) OLED is selected [1]. The elementary device, specified in Table 3 and depicted in Fig. 2, corresponds to Fig. 1 (setting $M = 1$ and $N = 3$) with an electric line source excitation located at $z' = 20\,\text{nm}$, radiates typically at $\lambda \approx 600\,\text{nm}$ [17].

n	Layer Material	n_n	κ_n	$d_n[\text{nm}]$
-2	Silver	0.124	3.73	$-\infty$
-1	MEH-DOO-PPV	1.9	0.01	0
$+1$	MEH-DOO-PPV	1.9	0.01	200
$+2$	ITO	1.85	0.0065	300
$+3$	Glass	1.5	0	10^6
$+4$	Air	1	0	$+\infty$

Table 3: Geometrical and electrical properties of a prototype BE-OLED, corresponding to Fig. 1. Data is retrieved from [6, 17].
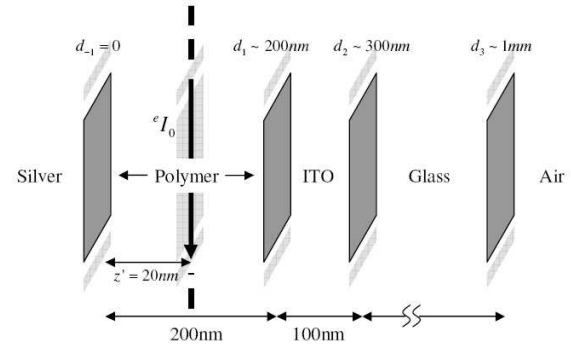
Figure 2: Physical configuration of the prototype BE-OLED specified in Table 3.

Radiation patterns of prototype BE-OLEDs with varying either glass layer thickness or coherence length are depicted in Fig. 3 and Fig. 4, respectively. While Fig. 3 patterns correspond to four glass layers with varying thickness, $\backsim 1\,\text{nm}$ to $\backsim 1\,\text{mm}$, and $L_c = 150\,\mu\text{m}$, Fig. 4 is associated with three coherence lengths, $5\,\mu\text{m}$ to $150\,\mu\text{m}$, whereas the glass layer thickness is set to $\backsim 1\,\mu\text{m}$.

It is readily observed that for device thickness (represented by the dominant glass thickness) much smaller than the coherence length, interference effects are noticeable whereas the reversed situation results in a quasi-Lambertian radiation pattern. Furthermore, these observations agree
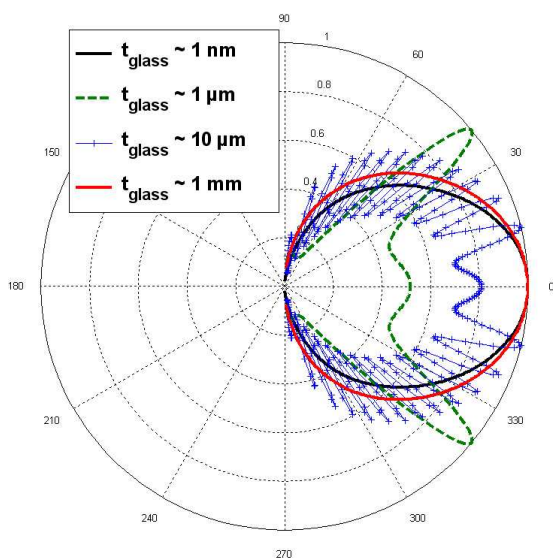
Figure 3: The effect of glass thickness on the OLED radiation pattern, for electric line source excitation, constant coherence length $L_c = 150\,\mu m$ and four glass thickness values: $\backsim 1\,nm$ (black solid line), $\backsim 1\,\mu m$ (green dashed line), $\backsim 10\,\mu m$ (blue line with plus marker) and $\backsim 1\,mm$ (red solid line).
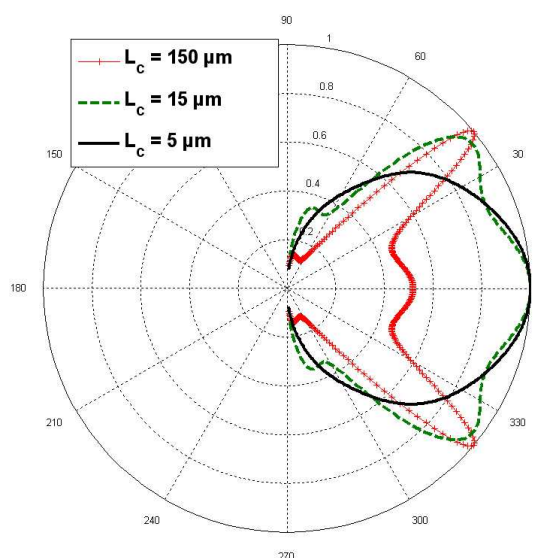
Figure 4: The effect of coherence length on OLED radiation pattern, for electric line source excitation, constant glass thickness of $\backsim 1\,\mu m$ and three cohrence length values: $5\,\mu m$ (black solid line), $15\,\mu m$ (green dashed line) and $150\,\mu m$ (red line with plus markers).

well with experimental measurements taken in the two limits, [9] and [6], respectively. Note that the two extremes correspond to radiation patterns which reflect either the device properties, i.e., thickness dependent interference, or the source characteristics, leading to spectral broadening dominated pattern.

4. CONCLUSION

A complete analysis for the electromagnetic radiation from two-dimensional OLEDs, incorporating temporal coherence effects, has been presented and verified through numerical simulations. The resultant analytical expressions preserve the physical intuition of the device optics and allow for an efficient implementation and design. Furthermore, as demonstrated for the prototype device, they establish a clear and simple relation between the device structure and the radiation pattern.

REFERENCES

1. Friend, R. H., R. W. Gymer, A. B. Holmes, J. H. Burroughes, R. N. Marks, C. Taliani, D. D. C. Bradley, D. A. Dos Santos, J. L. Brédas, M. Lögdlund, and W. R. Salaneck, "Electroluminescence in conjugated polymers," *Nature*, Vol. 397, 121–128, 1999.

2. Tessler, N., G. J. Denton, and R. H. Friend, "Lasing from conjugated-polymer microcavities," *Nature*, Vol. 382, No. 6593, 695–697, 1996.

3. Tessler, N., V. Medvedev, M. Kazes, S. H. Kan, and U. Banin, "Efficient near-infrared polymer nanocrystal light-emitting diodes," *Science*, Vol. 295, 1506–1508, 2002.

4. Forrest, S. R., "Active optoelectronics using thin-film organic semiconductors," *IEEE J. Sel. Top. Quantum Electron.*, Vol. 6, No. 6, 1072–1083, 2000.

5. Lu, M.-H. and J. C. Sturm, "Optimization of external coupling and light emission in organic light-emitting devices: Modeling and experiment," *J. Appl. Phys.*, Vol. 91, No. 2, 595–604, 2002.

6. Celebi, K., T. D. Heidel, and M. A. Baldo, "Simplified calculation of dipole energy transport in a multilayer stack using dyadic Green's functions," *Opt. Express*, Vol. 15, No. 4, 1762–1772, 2007.

7. Smith, L. H. and W. L. Barnes, "Using a low-index host layer to increase emission from organic light-emitting diode structures," *Org. Electron.*, Vol. 7, No. 6, 490–494, 2006.

8. Kim, J.-S., P. K. H. Ho, N. C. Greenham, and R. H. Friend, "Electroluminescence emission pattern of organic light-emitting diodes: Implications for device efficiency calculations," *J. Appl. Phys.*, Vol. 88, No. 2, 1073–1081, 2000.

9. Michelotti, F., G. Roma, A. Belardini, N. Danz, A. Pace, F. Sarto, and R. M. Montereali, "Micro-cavity organic light emitting diodes for biochip applications," *J. Non-Cryst. Solids*, Vol. 352, No. 23–25, 2476–2479, 2007.

10. Kahen, K. B., "Rigorous optical modeling of multilayer organic light-emitting diode devices," *Appl. Phys. Lett.*, Vol. 78, No. 12, 1649–1651, 2001.

11. Chance, R. R., A. Prock, and R. Silbey, "Molecular fluorescence and energy transfer near interfaces," *Adv. Chem. Phys.*, Vol. 37, 1–65, 1978.

12. Felsen, L. B. and N. Marcuvitz, *Radiation and Scattering of Waves*, 1st edn., Prentice-Hall, Englewood Cliffs, N.J., 1973.

13. Wolf, E., "Coherence and radiometry," *J. Opt. Soc. Am.*, Vol. 68, No. 1, 6–17, 1978.

14. Einziger, P. D., L. M. Livshitz, and J. Mizrahi, "Rigorous image-series expansions of quasistatic Green's functions for regions with planar stratification," *IEEE T. Antenn. Propag.*, Vol. 50, No. 12, 1813–1823, 2002.

15. Razansky, D., D. F. Soldea, and P. D. Einziger, "Bounds and estimates for power absorption in two-dimensional highly lossy configurations," *J. Appl. Phys.*, Vol. 95, No. 12, 8298–8308, 2004.

16. Saxena, K., D. S. Mehta, R. Srivastava, and M. N. Kamalasanan, "Spatial coherence properties of electroluminescence from Alq3-based organic light emitting diodes," *Appl. Phys. Lett.*, Vol. 89, No. 6, 061124, 2006.

17. Leger, J. M., S. A. Carter, B. Ruhstaller, H.-G. Nothofer, U. Scherf, H. Tillman, and H.-H. Hrhold, "Thickness-dependent changes in the optical properties of PPV- and PF-based polymer light emitting diodes," *Phys. Rev. B*, Vol. 68, No. 5, 054209, 2003.

Fourth Order Moment Statistical Characteristic of the Laser Pulse Scattering on One-dimensional Random Rough Surface

Ming-Jun Wang[1, 2], **Zhen-Sen Wu**[2], **Ying-Le Li**[1], **and Geng Zhang**[2]

[1]Department of Physics, Xian Yang Normal College, Xianyang, Shaanxi 712000, China
[2]School of Science, Xidian University, Xi'an, Shaanxi 710071, China

Abstract— The fourth order moment statistical characteristic of the pulse wave scattering on one-dimensional random rough surfaces is investigated numerically and presented based on Kirchhoff's approximation in detail. As a special application, the fourth order moment (FOM) is simplified to get the two positions and two frequencies mutual correlation function of the pulse wave scattering on random rough surfaces. Which the distributions of FOM change with the coherence bandwidth frequency difference and scattering angles in different incidence angles, the mean of fluctuating heights and coherent length on the random rough surface is computed. The numerical results show that the mean of fluctuating heights and coherent length of random rough surfaces have an important effect on the value of the FOM. Generally, the largest scattering value and smallest coherence bandwidth of the FOM occur at the specular directions. That is to say, the FOM will have a lower coherent component and its distributions of scattering angles will be broadened in relation to the rugged surface. At same time, the FOM decreased rapidly as there was a bigger coherence bandwidth frequency difference.

1. INTRODUCTION

There has been an interesting and need for understanding the statistical characteristics laser scattered from one dimensional random rough surface [1–3]. A.Ishimaru has used the two-frequency mutual coherent function to study pulse scattering from one dimensional perfectly conducting rough surface [4, 5]. Using physical optics approximation, Wu Z. S. has studied the scattering from rough objects with random shapes and given the theoretical formula of the mono-static and bi-static scattering cross section [6, 7]. Chen Hui has applied Kirchhoff and physical optics approximation to derive the two-frequency mutual coherence function of the pulse plane wave and pulse beam scattering from two dimensional randomly rough surface [8]. But all above only studied the second order moment of the scattering field. In order to study the fluctuation of the scattering intensity and improve the imaging quality, For fourth order statistical characteristics laser scattered field, people always study it on electromagnetic wave propagation through random media [1–3, 9].

This paper investigates the fourth order moment statistical characteristic of the plane scattering from two dimensional random rough surface. In this case, we present analytical expression for the fourth order moment for plane wave scattering based on the Kirchhoff approximation. The numeric results of the two-frequency and two-position mutual coherence function of the field are calculated at 1.06 μm. It is noted that the fourth moment of smooth surface has stronger enhanced backscattering effect, while that of rough surface is lower.

2. THE FOURTH ORDER MOMENT OF THE SCATTERING E-FIELD FROM RANDOM ROUGH SURFACES

2.1. Two Frequency Coherence Function of the Scattered Wave from Rough Surfaces

The Kirchhoff approximation for the scattered waves from rough surfaces has been extensively studied, and therefore we give only a brief summary necessary for our discussion. When the rms height δ is large compared with wavelength, the geometric optical approximation is applied, so the far scattered field is given by [1, 4, 5]

$$E_s = k \cos \theta_s \sqrt{\frac{2\pi}{kR}} e^{i(k_i R - \pi/4)} T\left(\vec{k}_i, \vec{k}_s\right) \tag{1}$$

where k is the wave-number, $k = 2\pi/\lambda$, θ_s is the scattering angle, R is the distance from the far field to the original point. And the transition matrix T can be written as [12]

$$T(k_i, k_s) = \frac{F_2}{2\pi} \int R_1 e^{i\left(\vec{k}_i - \vec{k}_s\right) \cdot \vec{r}_1} dx_1 \tag{2}$$

R_1 is the Fresnel reflection coefficient at r_1, $\vec{k}_i$ is the incidence wave vector and $\vec{k}_s$ is the scattering wave vector and F_1 which can be written as:

$$F_1 = \frac{1 + \cos\theta_i \cos\theta_s - \sin\theta_i \sin\theta_s \cos\phi_s}{\cos\theta_s(\cos\theta_i + \cos\theta_s)} \tag{3}$$

$$\vec{k}_i = k\sin\theta_i\hat{x} - k\cos\theta_i\hat{z}, \quad \vec{k}_s = k\sin\theta_s\cos\phi_s\hat{x} + k\sin\theta_s\sin\phi_s\hat{y} + k\cos\theta_s\hat{z} \tag{4}$$

Two-frequency scattering cross section per unit area of rough surface is

$$\sigma_0 = (R/L) < T_{sf}T'^*_{sf} > \tag{5}$$

$$T = < T_s > + T_{sf}, \quad T' = < T'_s > + T'_{sf} \tag{6}$$

$$\vec{k}'_i = k'\sin\theta'_i\hat{e}_x + k'\cos\theta'_i\hat{e}_y, \quad \vec{k}'_s = k'\sin\theta'_s\hat{e}_x + k'\cos\theta'_s\hat{e}_y \tag{7}$$

L is the illumination distance, $k = \omega/c$, $k' = \omega'/c$, $< T_s >$ is the coherent part and T_{sf} is the incoherent part.

Using the (1) and (2), we get

$$\sigma_0 = 2\pi\sqrt{kk'}\cos\theta_s\cos\theta'_s\left(< T_f(\omega)T'^*_f(\omega) > /L\right)e^{i\left(\vec{k} - \vec{k}'\right)\cdot R} \tag{8}$$

$< T_f(\omega)T'^*_f(\omega') >$ is two-frequency coherence function of the scattered wave from rough surfaces.

2.2. The Fourth Order Statistical Characteristics of the Scattering from Random Rough Surfaces
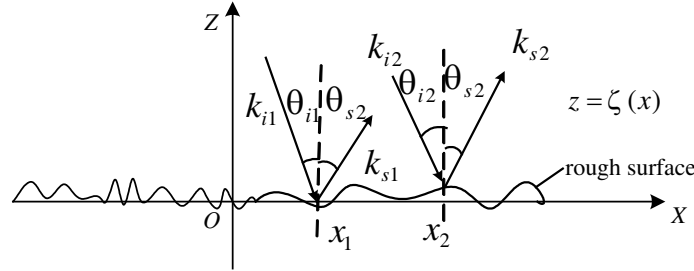
Figure 1: Geometry relations of rough surface two scattering points.

Two points scattering characteristics geometry relations from rough surface in Fig. 1. The scattering E-field from points x_1, x_2 are given by

$$\vec{E}_{s1} = \vec{E}_{s1}(k_{i1}, k_{s1}) = < \vec{E}_{s1} > + \vec{E}_{sf1}, \quad \vec{E}'_{s1} = \vec{E}'_{s1}(k'_{i1}, k'_{s1}) = < \vec{E}'_{s1} > + \vec{E}'_{sf1}$$
$$\vec{E}_{s2} = \vec{E}_{s2}(k_{i2}, k_{s2}) = < \vec{E}_{s2} > + \vec{E}_{sf2}, \quad \vec{E}'_{s2} = \vec{E}'_{s2}(k'_{i2}, k'_{s2}) = < \vec{E}'_{s2} > + \vec{E}'_{sf2} \tag{9}$$

According to the Eqs. (5) $\sim$ (7), the fourth order statistical characteristics of rough surface scattering is defined

$$\Psi = (R/L)^2 < E_{1sf}E'^*_{1sf}E_{2sf}E'^*_{2sf} > \tag{10}$$

So the fourth order moment of the scattering from random rough surface is given by

$$\Psi = 4\pi^2\sqrt{k_1k'_1}\sqrt{k_2k'_2}\cos\theta_{s1}\cos\theta'_{s1}\cos\theta_{s2}\cos\theta'_{s2}(< T_{1f}T'^*_{1f}T_{2f}T'^*_{2f} > /L^2)e^{i\left(\vec{k}_1 - \vec{k}'_1\right)\cdot R}e^{i(\vec{k}_2 - \vec{k}'_2)\cdot R'} \tag{11}$$

Assuming the rough surface is Gaussian rough surface, as the property of the Gaussian variables shows [1, 3], the fourth order moment is equal to a sum of products of all possible covariance, then, get the following

$$< T_{1f}T'^*_{1f}T_{2f}T'^*_{2f} > = < T_{1f}(k_{s1}, k_{i1})T'^*_{1f}(k'_{s1}, k'_{i1})T_{2f}(k_{s2}, k_{i2})T'^*_{2f}(k'_{s2}, k'_{i2}) >$$
$$= < T_{1f}(k_{s1}, k_{i1})T'^*_{1f}(k'_{s1}, k'_{i1}) > < T_{2f}(k_{s2}, k_{i2})T'^*_{2f}(k'_{s2}k'_{i2}) >$$
$$+ < T_{1f}(k_{s1}, k_{i1})T'^*_{2f}(k'_{s2}, k'_{i2}) > < T'^*_{1f}(k'_{s1}, k'_{i1})T_{2f}(k_{s2}, k_{i2}) >$$
$$+ < T_{1f}(k_{s1}, k_{i1})T_{2f}(k_{s2}, k_{i2}) > < T'^*_{1f}(k'_{s1}, k'_{i1})T'^*_{2f}(k'_{s2}, k'_{i2}) > \tag{12}$$

From Refs. [4, 5]

$$< T_f T_f'^* >= H_1' H_1 \int dx_1 dy_1 \int dx_1' dy_1' \left[< \exp\left(-i\vec{v}\cdot\vec{r}_1 + i\vec{v}'\cdot\vec{r}_1'\right) > \right] \quad (13)$$

Let $\vec{v} = \vec{v}\hat{x} + v_z\hat{z}$, $\vec{r}_1 = x_1\hat{x} + \xi\hat{z}$. As $\vec{r}_1 = \vec{r}_2$, $\vec{r}_1' = \vec{r}_2'$, $v_1 = v_2$, $v_1' = v_2'$, and assuming a Gaussian distribution for $\xi(x)$. Using (13), the first item of (12) for an example to derive the expression

$$< T_{1f} T_{1f}'^* >< T_{2f} T_{2f}'^* >= H_1 H_1' H_2 H_2' \Phi_{s11'} \Phi_{11'} \Phi_{s22'} \Phi_{22'}$$

$$\Phi_{11'} = \exp\left[\left(-v_{z1}^2 + v_{z1'}^2\right)\sigma^2/2\right] \sqrt{\frac{\pi}{v_{z1}v_{z1'}}} \left(\frac{l}{\sigma}\right) \exp\left(-v_{c11'}^2 l^2/4 v_{z1} v_{z1'}\sigma^2\right) \quad (14)$$

$$\Phi_{s11'} = \frac{1}{L}\int_{-\infty}^{\infty} \exp\left(-\pi x_{c11'}^2/L^2\right)\exp\left(-iv_{d11'}x_{c11'}\right)dx_{c11'} = \exp\left(-\pi v_{d11'}^2 L^2/4\pi\right) \quad (15)$$

$$\Phi_{22'} = \exp\left[\left(-v_{z2}^2 + v_{z2'}^2\right)\sigma^2/2\right] \sqrt{\frac{\pi}{v_{z2}v_{z2'}}} \left(\frac{l}{\sigma}\right) \exp\left(-v_{c22'}^2 l^2/4 v_{z2} v_{z2'}\sigma^2\right) \quad (16)$$

$$\Phi_{s22'} = \frac{1}{L}\int_{-\infty}^{\infty} \exp\left(-\pi x_{c22'}^2/L^2\right)\exp\left(-iv_{d22'}x_{c22'}\right)dx_{c22'} = \exp\left(-\pi v_{d22'}^2 L^2/4\pi\right) \quad (17)$$

The functions of subscript $11'$ are the first factor $< T_{1f} T_{1f}'^* >$, and the functions of subscript $22'$ are the second factor $< T_{2f} T_{2f}'^* >$ in the first items Eq. (2). And using the same method we also get fourth order moment

$$\Psi = 4\pi^2 \sqrt{k_1 k_1'} \sqrt{k_2 k_2'} \cos\theta_{s1} \cos\theta_{s1}' \cos\theta_{s2} \cos\theta_{s2}' e^{i\left(\vec{k}_1 - \vec{k}_1'\right)\cdot R} e^{i\left(\vec{k}_2 - \vec{k}_2'\right)\cdot R'}$$

$$\left[H_1 H_1' \Phi_{s11'} \Phi_{11'} H_2 H_2' \Phi_{s22'} \Phi_{22'} + H_1 H_2' \Phi_{s12'} \Phi_{12'} H_1' H_2 \Phi_{s1'2} \Phi_{1'2} + H_1 H_2 H_1' H_2' \sigma^2 L^4\right] \quad (18)$$
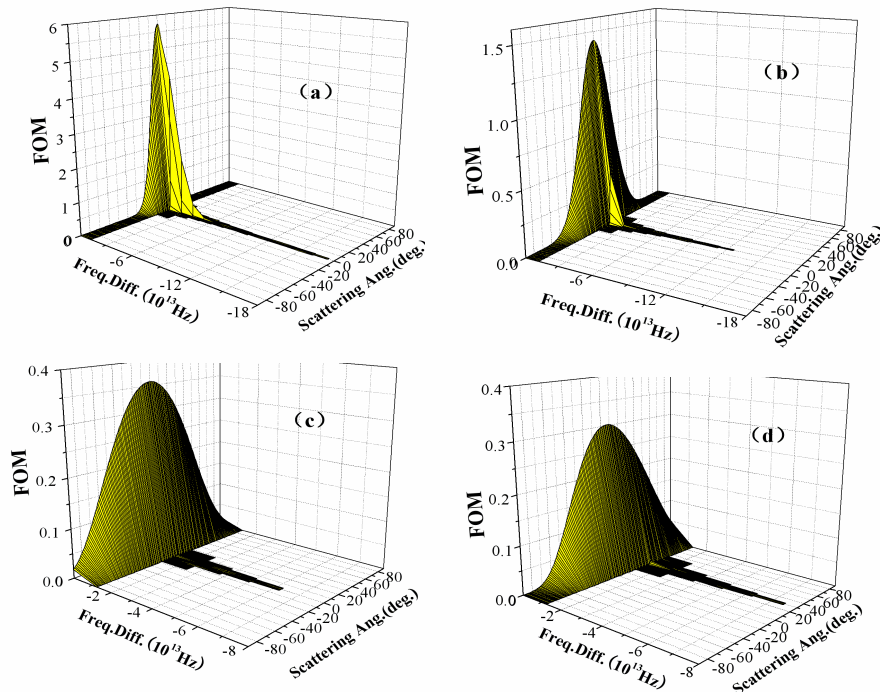

Figure 2: Fourth order moment change with scattering angle and frequency difference.

3. THE NUMERICAL RESULTS AND DISCUSSION

As an example, we utilize the Eq. (18) to calculate the scattering fourth order moment for one dimensional gaussian random rough surface in Figs. 2(a)–(b). The incident wave length λ is $1.06\,\mu$m, and the surface illumination length L is 100λ. The incident angle is $0°$, $30°$, coherent length $l = 12\lambda$ and the mean of fluctuating heights (rms) $\sigma = 1\lambda$, 2λ, 4λ, respectively in Figs. 2(a)–(c) and (d). The Numerical results are shown that as the frequency difference is 0, and in the specular direction, there are the biggest values while the values in frequency difference greater than 0 and other direction can be neglected; With the increase of rms height, when the frequency difference is 0, there is also the biggest values which is greatly smaller than that of the smaller rms height, and the values decreases in specular direction but increase in other direction. In addition, we can see that the increase of rms height results in the correlation bandwidth narrowing.

4. CONCLUSION

This paper presents a study of the fourth order moment of the scattering from one dimensional randomly rough surface. Based on the Kirchhoff approximation and the geometric optical approximation, analytical expressions are derived and the numerical results are given. It is shown that only in the specular direction and 0 frequency difference, the fourth order moment can be considered when the surface is fairly smooth. However, this paper only discusses the case when the incident wave is plane wave which is the simplest case. In fact, discussing the scattering intensity fluctuation and speckle effect when the incident wave is pulse waves or laser pulse beams has more important meaning, Next step we will work for this subjects.

REFERENCES

1. Ishimaru, A., *Wave Propagation and Scattering in Random Media*, Academic press, New York, 1978.
2. Ogilvy, J. A., *Theory of Wave Scattering from Random Rough Surfaces*, Adam Hilger, Bristol, 1991.
3. Tsang, L., J. A. Kong, and K. H. Ding, *Scattering of Electromagnetic Waves*, Vol. 1, Theory and Applications, Wiley Interscience, New York, 2000.
4. Ishimaru, A., L. Ailes-Sengers, P. Phu, and D. Winebrenner, "Pulse broadening of enhanced backscattering from rough surfaces," *Waves Random Media*, Vol. 4, No. 4, 453–465, 1994.
5. Ishimaru, A., L. Ailes-Sengers, P. Phu, and D. Winebrenner, "Pulse broadening and two-frequency mutual coherence function of the scattered wave from rough surfaces," *Waves Random Media*, Vol. 4, No. 2, 139–148, 1994.
6. Wu, Z., "IR laser backscattering by arbitrarily shaped dielectric object with rough surface," *SPIE's International Symposium on Optical Science and Engineering*, 21–26, San Diego, USA, July 1991.
7. Wu, Z. and S. Cui, "Bistatic scattering by arbitrarily shaped objects with rough surface at optical and infrared frequencies," *Inter. J. of Infrared and Mill. Waves*, Vol. 13, No. 4, 537–549, 1992.
8. Chen, H. and Z. S. Wu, " Infrared laser pulse scattering from randomly rough surface," *Inter. J. of Infrared and Mill. Waves*, Vol. 25, No. 4, 1211–1219, 2004.
9. Yang, R. K., Z. S. Wu, and P. R. Zhang, "Study of scintillation for infrared laser beam propagating in atmospheric turbulence on Earth-space paths," *Int. J. IR and Mill. Waves*, Vol. 125, No. 66, 1001–1003, 2004.

Investigation of Novel Ultrasonic Positioning Method Installed in Sensor Network

Mitsutaka Hikita, **Yasushi Hiraizumi**, **Hiroaki Aoki**,
Junji Matsuda, **and Tomoaki Watanabe**

Faculty of Global Engineering, Kogakuin University, Shinjuku-ku, Tokyo 163-8677, Japan

Abstract— A new concept called "Sensor Network" has been proposed with the development of mobile communications system such as cellular phone, radio LAN and Bluetooth. Signals from many sensor nodes spread in a wide area are gathered to a center node by technology similar to that used in mobile communications. Sensor network enables home/office circumstance control, environment monitoring and protection based on the collected data. We proposed a novel ultrasonic positioning method which can be installed in sensor network. This method has also been investigated to be combined with 2.4-GHz ZigBee, which has been regulated by IEEE 802.15.4 as wireless-communications medium for sensor network.

1. INTRODUCTION

Cellular-phone systems have spread all over the world and their technologies continue to be developed during the first half of the twenty-first century. On the other hand outstanding technical innovations have been observed in a sensor area. A new concept called "Sensor Network" has been proposed recently by combining sensor technologies with such mobile communications system as cellular-phone, radio LAN and Bluetooth. In this network system, a lot of sensors distributed in a certain area such as home, office and public places are connected via rather simple private radio communications network. It is thought to have a big impact on our lives and to grow to be a giant industry like the cellular-phone systems. It will also contribute to future ecology, i.e., energy saving and environmental preservation. In our laboratory, we have been studying the sensor network to achieve comfortable living circumstances by home/office sensing and control [1].

In this paper, we proposed a new ultrasonic positioning method which can be installed in the sensor network. Sensor nodes require long-term operation with extremely low-power supply, such as several-year operation with a single battery. However, such conventional ultrasonic positioning methods as a pulse-echo method and a code-division [2] method require high voltage handling capability for the transmitter and complicated signal processing procedures for receiver, respectively. Therefore, their devices can not be included in sensor nodes. Our proposed method can provide not only the low-power consumption but also removal of the complicated signal processing procedures, which overcomes the limitation for use as sensor-node devices [3]. Our method has also been invented based on ZigBee-based sensor network. ZigBee has been regulated by IEEE802.15.4 [4] and ZigBee Alliance as a wireless-communication medium used in low-power and long-term radio systems such as sensor network. We also showed the feasibility of the new method by fundamental experiments compared with those of conventional method. The sensor network including our positioning devices will possibly be applied not only to home/office monitoring but also to care for old people, prevention of crime and watch in hospitals.

2. SENSOR NETWORK BASED ON ZIGBEE

Texas Instrument, one of the most leading semiconductor-IC manufacturing companies, has a vision as shown in Fig. 1 for ZigBee-based sensor network. Almost all office/home controls, i.e., light, temperature/humidity, security/alarm, smoke-detection/alarm and other notifications, are conducted via sensor network. T. I. has a plan to provide semiconductor ICs used in such network systems. Our proposed ultrasonic positioning devices will be installed in sensor nodes together with these developed ICs.

ZigBee has been regulated by IEEE802.15.4 [4] and ZigBee Alliance as a wireless-communication medium for the sensor network. Wireless specifications for ZigBee are shown in Table 1 compared with other recent radio communications media, i.e., Bluetooth and UWB 2.4 GHz, same frequency as Bluetooth, is used, but data rate is 250 Kbps which is about one third of Bluetooth. The number of connectable nodes per one network is 65,535, which is completely different from Bluetooth's 7 slaves per one master pico-network. Because of these characteristics, sensor-network is the most adequate application for ZigBee.

The protocol structure for ZigBee is shown in Fig. 2, which has almost same layer configuration as Bluetooth or other radio communications systems. The network topology is shown in Fig. 3. Not only conventional star-link type topology for Bluetooth and radio LAN but also new mesh-link type and tree-type topology can be possible for ZigBee, which can achieve 65,535 nodes connected per one network. Other one of the most important features required for ZigBee is the extremely low-power consumption, which provide several-year operation with a single battery.
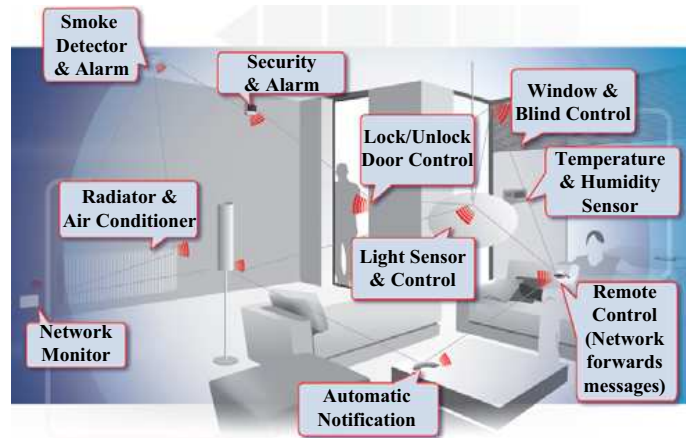
Figure 1: Concept of ZigBee-based sensor network (slightly modified from originals in H. P. of T. I.).

Table 1: Wireless specifications for ZigBee, Bluetooth and UWB.

System	ZigBee (IEEE802.15.4)	Bluetooth (IEEE802.15.1)	UWB (IEEE802.15.13)
Frequency Modulation	2.4 GHz OQPSK/DSSS	2.4 GHz GFSK/FHSS	3.1-10.6 GHz MCOFDM
Output Data Rate	0 dBm 250 Kbps	0 dBm 721 Kbps	-41.3 dBm/MHz 10^2~10^3 Mbps
Number of Nodes	65,535 Nodes /1 Network	7 Slaves /1 Master pico-net.	
Application	Sensor network Home network Temperature /Light/Air con. Security	PC, Printer Key board Audio sets Hands free	
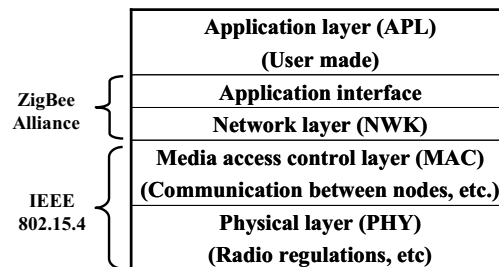
Figure 2: Protocol structure for ZigBee.

3. CONVENTIONAL ULTRASONIC MEASURING METHODS

One of the images for sensor network installed in homes and offices is shown in Fig. 4. Many sensor nodes with various sensors will be arranged, where nodes can communicate one another via ZigBee. Sensed signals from all nodes are gathered to a center node. The center node not only supervises sensor nodes and processes collected data but also sends control signals to other

installations. Ultrasonic devices within sensor nodes transmit and receive ultrasonic waves to measure the distances between sensor nodes and reflected objects.

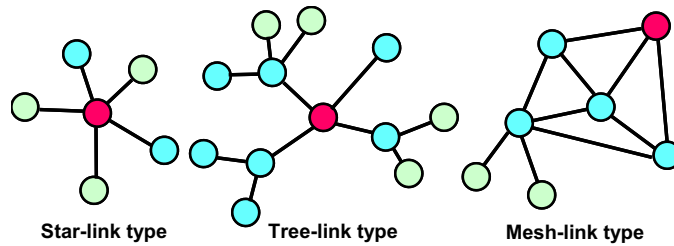Star-link type   Tree-link type   Mesh-link type

Figure 3: Network topology for ZigBee. Only star-link type is accepted for Bluetooth and radio LAN. All three link types are possible for ZigBee.
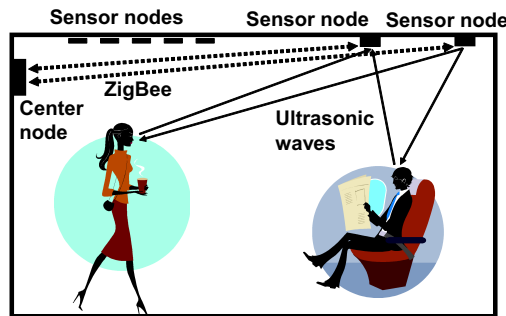

Figure 4: Image example for sensor network with ultrasonic positioning method.

Two major measuring methods have been used in conventional positioning systems. One is a pulse-echo method, and the other is a code-division method [2]. Assuming two objects as shown in Fig. 5, the pulse-echo method measures time delays between the transmitted and received burst-type ultrasonic pulses as shown in Fig. 6. In this case, the width of received pulses is spread comparatively compared with that of the transmitted pulses, which requires very narrow transmitted pulse to achieve enough space resolution. However, if we keep the same transmitted energy the amplitude of transmitted pulses must be increased with narrowing their width. Therefore, high voltage handling capability is needed for a final stage amplifier in the transmitter, which reveals that adopting this method to sensor nodes is very difficult.
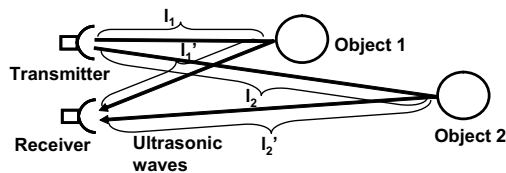

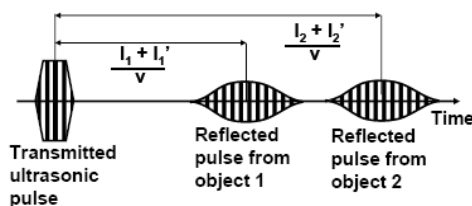Figure 5: Measurement model with two objects as well as transmitter and receiver.

Figure 6: Transmitted and received ultrasonic pulses for pulse-echo method.

The other code-division method uses Binary-phase-shift-keying (BPSK) signals as transmitted ultrasonic waves shown in Fig. 7's left figures. The $0/\pi$-phase combination is determined based on specific codes, such as Barker, M-sequence, and Gold codes. Output signals from a sliding correlator within the receiver have sharp peaks which correspond to auto-correlation function between the coded signal in received ultrasonic waves and a replica signal, i.e., reverse-order-coded signal, from the receiver. This method provides high space resolution as shown in Fig. 7. However, very high speed digital-signal-processing (DSP) ICs and large memories are necessary to achieve function of the sliding correlator, which also reveals it difficult to adopt this method to sensor nodes.
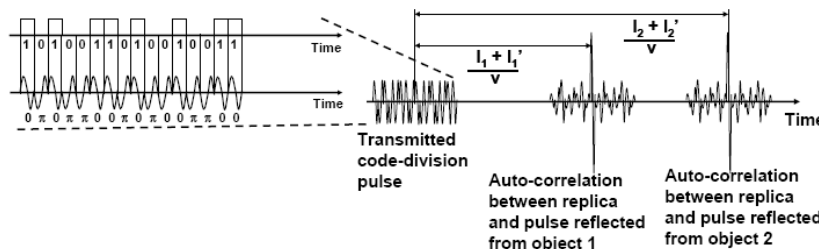

Figure 7: Transmitted and received ultrasonic pulses for code-division method. Left figures show specific code and BPSK signals.

4. NOVEL ULTRASONIC POSITIONING METHOD FOR SENSOR NETWORK

ZigBee which uses 2.4 GHz has been regulated by IEEE802.15.4 as one of radio standards for low data-rate communications media such as a sensor network [4]. As a block diagram is shown in Fig. 8(a), the sensor node used in conventional ZigBee system consists of 2.4-GHz VCO (Voltage-Controlled Oscillator) locked to TCXO (Temperature-Compensated Xtal Oscillator, i.e., 13 MHz) and other circuit components. We have proposed a new sensor node structure which is constructed with not only above circuits but also new ultrasonic-positioning devices as shown in Fig. 8(b).

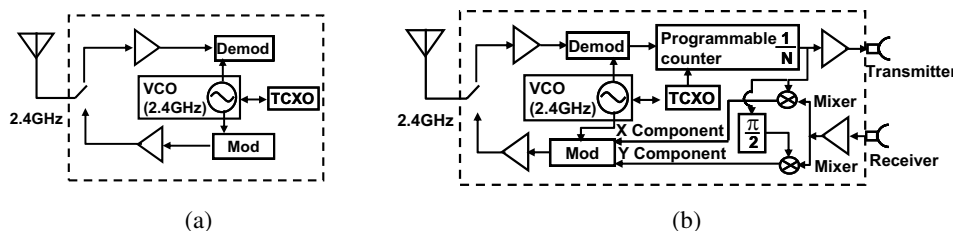(a)   (b)

Figure 8: Block diagram for ZigBee-based sensor node. (a) Conventional node. (b) New node with ultrasonic-positioning devices.

A same model as shown in Fig. 5 which consists of an ultrasonic transmitter, an ultrasonic receiver and two reflecting objects is considered. In general, an ultrasonic transducer has frequency characteristics. Assuming that the transmitted waves from the transducer have a frequency band-width from f_1 to f_2, and have single-peaked characteristics for amplitude and flat characteristics for phase respectively within the bandwidth as shown in Fig. 9(a), the received waves have reduced-amplitude and increased-phase characteristics due to the propagation and the reflection as shown in Fig. 9(b). A transmission function, $G(\omega)$, between the transmitter and the receiver can be obtained by a fraction between the characteristics in Figs. 9(a) and (b). Therefore, the magnitude of $G(\omega)$, $|G(\omega)|$, is calculated as the relative amplitude between Figs. 9(a) and (b). The argument of $G(\omega)$, $\angle G(\omega)$, is given by that in Fig. 9(b) because the phase of the transmitted waves is assumed to be zero. An Inverse Fourier Transform of $G(\omega)$ leads to the impulse response between the transmitter and the receiver. This impulse response includes signals of sin(x)/x-shaped function, as shown in Fig. 10. Center of the each signal of sin(x)/x-shaped function reveals a time delay of traveling ultrasonic waves between the transmitter and the receiver way of the corresponding reflecting object. Thus we can obtain the distance from the transmitter/receiver to the object by multiplying the time delay and wave velocity, i.e., about 340 m/s.
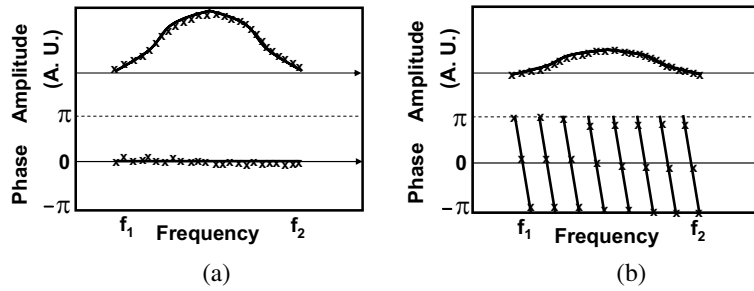
Figure 9: Amplitude and phase characteristics of ultrasonic continuous waves. Actual data at discrete frequencies have values as indicated by crosses. (a) Transmitted waves. (b) Received waves.
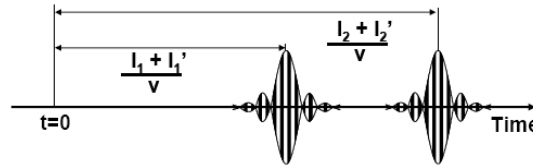

Figure 10: Impulse response, Inverse Fourier Transform of $G(\omega)$, between transmitter and receiver.

In an actual model, the transmitter/receiver devices are included in a sensor node as shown in Fig. 8(b). We divide the frequency from f_1 to f_2 into $n+1$ of discrete frequencies, i.e., $f(1) = f_1$, $f(2) = f_1 + (f_2 - f_1)/n$, ..., $f(n+1) = f_2$. These frequencies are also determined same as discrete frequencies used in IFFT (Inverse Fast Fourier Transform) procedure, and can be produced by counting signal, i.e., 13-MHz signal for example, from TCXO using a programmable counter as shown in Fig. 8(b). Arbitrary counting numbers, Ns, are designated from the center node via ZigBee, which provides signals at several tens KHz to the transmitter.

Examples of transmitted waves at the above $n+1$ of discrete frequencies are shown in Fig. 11(a), while those of the received waves are shown in Fig. 11(b). One transmitter (sensor node) transmits Fig. 10(a)'s waves, and several receivers (sensor nodes) around the transmitter receive Fig. 10(b)'s waves. In each receiver, only amplitudes and phases of the received waves are measured, i.e., the crosses shown in Fig. 9(b), and they are sent to the center node via ZigBee. Therefore, sensor nodes can achieve very low-power consumption because they only transmit and receive ultrasonic continuous waves with small amplitude. In the center node, IFFT procedure is executed on the basis of data for amplitudes and phases from sensor nodes. So, delay times, i.e., distances, from sensor nodes to objects are obtained as impulse response between the transmitter and receivers. Tangible advantages of the proposed system are that distance information can be automatically gathered in the center node, and all digital signal processing procedures such as IFFT are executed only in the center node with another power supply.
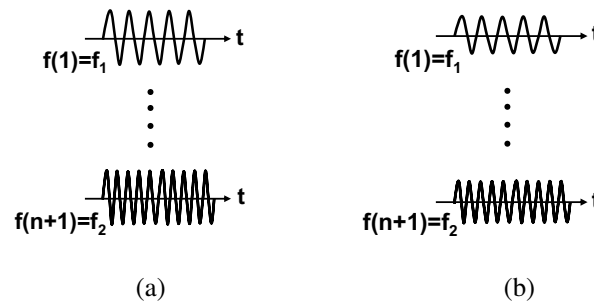

Figure 11: Continuous waves at $n+1$ of discrete frequencies from $f(1) = f_1$ to $f(n+1) = f_2$. (a) Transmitted waves. (b) Received waves.

5. EXPERIMENTAL RESULTS

In order to confirm our proposal, we experimentally investigated the method using a simple model shown in Fig. 12. The transmitter and the receiver are arranged about 50 cm apart to each other.

As a comparison, we first measured the same distance by a conventional pulse-echo method with burst type of 40-KHz ultrasonic pulses. Experimental results of the conventional method are shown in Figs. 13(a) and (b), where (a) and (b) show the transmitted and received pulses, respectively. The received pulse is delayed by about 1.5 msec which corresponds to about 50 cm.
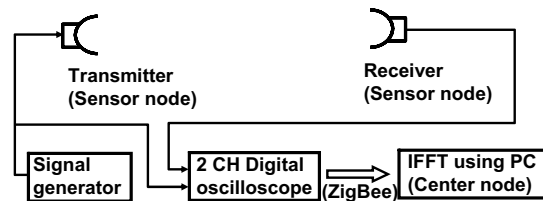

Figure 12: Experimental set-up with transmitter and receiver about 50 cm apart to each other.
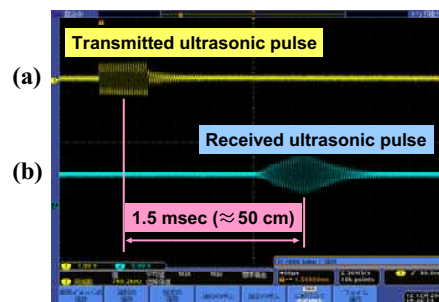

Figure 13: Experimental results for conventional pulse-echo method. (a) Transmitted pulse. (b) Received pulse.

In the new method, the transmitted and received continuous waves are displayed in an oscilloscope, and measured data of relative amplitudes and phases between them are sent to a PC as shown in Fig. 12. Relative amplitudes and phases are shown in Figs. 14(a) and (b), respectively. From the figures, 3-dB bandwidth of the transducer is about 1.9 KHz, and almost no waves can be transmitted at the outside frequency range of 37.4–43.5 KHz. Impulse response characteristics between the transmitter and the receiver obtained by IFFT procedure on the basis of data in Figs. 14(a) and (b) are shown in Fig. 15. The distance of about 50 cm can be clearly recognized, and the threshold characteristics are better than those obtained by the conventional method. These results confirmed validity of our proposed novel method.
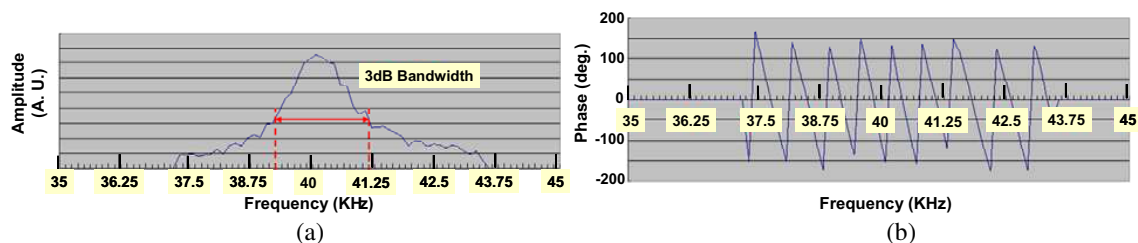

Figure 14: Measured data for proposed method. (a) Amplitude characteristics. (b) Phase characteristics.
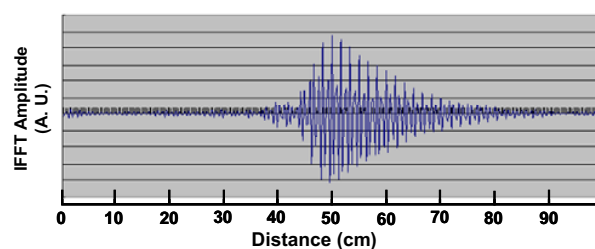

Figure 15: Measured distance with proposed method, i.e., IFFT results using Fig. 14's data.

6. CONCLUSIONS

Sensor network which consists of small-sized radio-communication infrastructures has been investigated as most promising means to monitor and control home/office circumstances. We proposed a novel ultrasonic distance measuring method which can be used as ultrasonic positioning installed in sensor network. We also devised a new sensor node structure which included not only network functions but also ultrasonic sensing devices. The fundamental experiments with the transmitter and the receiver facing to each other showed validity of the proposed method as well as possibility of new network systems including the positioning method.

REFERENCES

1. Home Page of Kogakuin University (http://www.ns.kogakuin.ac.jp/~wwa1022/).
2. Dixon, R. C., *Spread Spectrum Systems*, John Wiley & Sons, 1976.
3. Japanese patent pending.
4. IEEE Standard 802.15.4: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Low Rate Wireless Personal Area Networks (LR-WPANs), 2003.

Wireless Sensor Networks in Agriculture: Cattle Monitoring for Farming Industries

Kae Hsiang Kwong, **Tsung Ta Wu**, **Hock Guan Goh**, **Bruce Stephen**,
Michael Gilroy, **Craig Michie**, **and Ivan Andonovic**

Centre for Intelligent Dynamic Communications, Department of Electronic and Electrical Engineering
University of Strathclyde, Glasgow, United Kingdom

Abstract— This paper investigates an adaptation of Wireless Sensor Networks (WSNs) to cattle monitoring application. The proposed solution facilitates a desired requirement of continuously assessing the condition of individual animal, aggregating and reporting these data to the farm manager. There are several existing approaches to animal monitoring, from using a store and forward mechanisms to employing a GSM technique. These approaches for monitoring livestock health can only provide sporadic information and introduce a considerable cost in staffing and physical hardware. The core of this study is to overcome the aforementioned drawbacks by using alternative low cost, low power consumption sensor nodes, which are capable of providing real-time communications at a reasonable hardware cost. In this paper, the hardware and software have been carefully designed to provide early indication of possible outbreaks while conforming to WSNs' stringent limitations.

1. INTRODUCTION

The farming industry is an important sector of the UK economy with an annual turnover of around £ 5.8 billion. Close to three quarters of the country's land are used for farming and which employs over half a million people [1]. In recent history, there have been a number of major animal disease outbreaks in the UK. The two most significant incidents were the bovine spongiform encephalopathy (BSE) outbreak and foot and mouth disease (FMD) outbreak, which took place in 1986 and 2001 respectively. 4.5 million cows were incinerated after the identification of BSE and over 4 million livestock were slaughtered to stop the spread of foot and mouth disease. In total, these diseases are estimated to have cost the UK economy £ 13 billion [2].

It is always desirable to implement a health monitoring programme to observe the condition of animal as well as to prevent the outbreaks of animal diseases. Various researchers [3–5] have used wireless networks with a mounted sensory device (i.e., collar) to track individual animal activity and to monitor the health condition. One noteworthy example of this is 'ZebraNet' [3]. Devices mounted on the zebra routinely exchange data such as GPS position with all other devices that fall within their transmission range. If sufficient memory space is available, a user could then download historical position data of multiple animals by approaching a single zebra. Sikka et al. [4] deployed a mixture of static and mobile sensors to monitor soil moisture and cow movement. The static nodes measure soil moisture profiles at varying depths while the mobile nodes are carried by the livestock to investigate animal behaviours. Both schemes [3, 4] are based on a store and forward approach which may not be applicable in WSNs since the memory space of sensor node is scarce. Moreover, the maintenance and labour cost is high since the retrieval of the aggregated measurement from animals requires physical human access. The approach employed by [3, 4] may also suffer a considerable delay in packet delivery if the stored data is not frequently updated with the newly measured data. Mayer et al. [4] retrieve data from animal mounted device by utilising GSM infrastructure to facilitate real-time communication. However, this becomes prohibitively expensive in the case of monitoring large number of animals (i.e., a collar can typically cost approximately 1700 Euros) [5].

Having discussed the basic problems of present schemes [3–5], the ultimate goal of this article is to design a network system that uses alternative low-cost, low power consumption sensor nodes to facilitate a real-time health monitoring application. To achieve this goal, firstly, an antenna diversity collar is designed to improve the performance of radio coverage in farm field. Secondly, in contrast to traditional store and forward mechanism, a particular routing protocol is presented to facilitate multi-hop connectivity. In this protocol, each node determines its next nearby hop based on current network connectivity and traverse hop distance, that falls between destination point and present position. Utilising the above information, the packet can be forwarded straightforwardly by comparing the traverse hop distance between the sender and the most current preferred receiver. The packet will move one hop closer to the destination at each forwarding stage. Eventually,

the forwarded packet will arrive at the base station with an implicit routing path. The designed protocol skips the time spent in creating and maintaining an explicit routing path that results in shorter packet delay. To our best knowledge, no routing schemes are currently available to support animal monitoring.

2. CHALLENGES

Sensor nodes are known for their constrained capacities in terms of energy storage, limited computation power and memory capability, etc. For example, a MICAz node [6] is powered by two alkaline AA batteries and has only one 4 MHz processor with 128 KB of memory and 4 KB of RAM. Giving these inherent limited capacities, implementing communication devices for cattle monitoring application raises specific and severe challenges. The major challenges include:

- **Radio attenuation caused by animal body**. The radio performance needs to be improved from the attenuation caused by the animal body. In cattle health monitoring applications, the cattle are generally fed in herd and this massively increases the surface area, which will seriously affect the radio performance. The reason is due to radio absorption taking place at the animal's body [7]. Ideally, communications hardware should be capable of minimising the impact of through body attenuation. This will be addressed in Section 3.

- **Memory is scarce in sensor nodes.** As a result of the severe memory constraints inherent within sensor networks, the traditional store and forward approach is not feasible. Therefore a robust routing protocol is needed to instantly forward the measured data back to base station through multi-hop.

- **Mobility**. Presently, in most WSN monitoring applications, e.g., [9], the sensor nodes are assumed to be static. One of the major issues raised when the sensor nodes are mounted on the cattle the location of sensor nodes changes frequently. The network topology and routing path configurations should be dynamic in responding to frequent movement of animal and embrace the impact occurred in the performance of packet delivery.

3. WIRELESS COMMUNICATIONS

In the cattle monitoring system, the animal is free to roam around and wireless methods are considered the only feasible option to establish and maintain communications between a base station and collar attached on cattle. Access to the range of radio frequency bands is regulated by various standards bodies such as International Telecommunication Union (ITU). WSNs use unlicensed bands, which are the ISM — industrial, scientific and medical radio bands, were originally reserved internationally for non-commercial used of radio frequency (RF) electromagnetic fields. In this paper, most of the trials are conducted using 2.4 GHz ISM band, as this available globally.

An estimation of signal penetration rate through animal's body can be made using data obtained from a detailed study of dielectric properties of body tissues [7]. While this is not completely accurate, the electric properties of mammalian tissue are expected to be reasonably similar between species. Table 1 summarises penetration depth of common ISM band frequencies. Table 1 shows that the penetration rate is less than 2.5 cm in fleshy tissues (skin/muscle) when transmitting at 2.4 GHz. The lower frequency band of 315 MHz has a better penetration depth but the width of cow

Table 1: Frequencies and tissue penetration depth [7].

	315 MHz			2.4 GHz		
	Conductivity (S/m)	Relative permittivity	Penetration depth (m)	Conductivity (S/m)	Relative permittivity	Penetration depth (m)
Blood	1.3212	65.375	0.03651	2.5024	58.347	0.016407
Bone cortical	0.083944	13.386	0.23495	0.38459	11.41	0.046992
Fat	0.039795	5.6239	0.3225	0.10235	5.2853	0.11956
Muscle	0.77442	58.001	0.055463	1.705	52.791	0.022785
Skin (dry)	0.64898	49.249	0.060904	1.4407	38.063	0.022956

neck is approximately 0.25 m, therefore it is not possible for any radio signal to reliably penetrate through the animal's body and establish network connectivity.

An antenna diversity scheme was examined, in which two antennae (placed at top left and top right of the collar) can be used to optimise the collar radio coverage. The locations of the antennae are carefully considered so that signal propagates efficiently outward with fewer impediments from the other cattle in the proximity. Figure 1(c) illustrates positions of the antennae and Figure 1(b) shows how the signal propagates outward in a desired fashion. Farm trials of antenna diversity enabled collars were carried out at a research farm. A single antenna diversity collar was mounted on a dairy cow placed within a small herd of 14 cows in a 12 m × 20 m collecting yard. A set of RSSI was collected for each collar during a pre-set time interval. The received RSSI is consistently 10 to 15 dB higher for one antenna (Figure 2). It may be deduced that antenna 1 clearly is oriented with the antenna in the line of sight of the base station whereas antenna 2 is oriented away from the base station. As a consequence, antenna 1 records a consistently higher received signal strength and a greater number of packets are received during the same time interval.
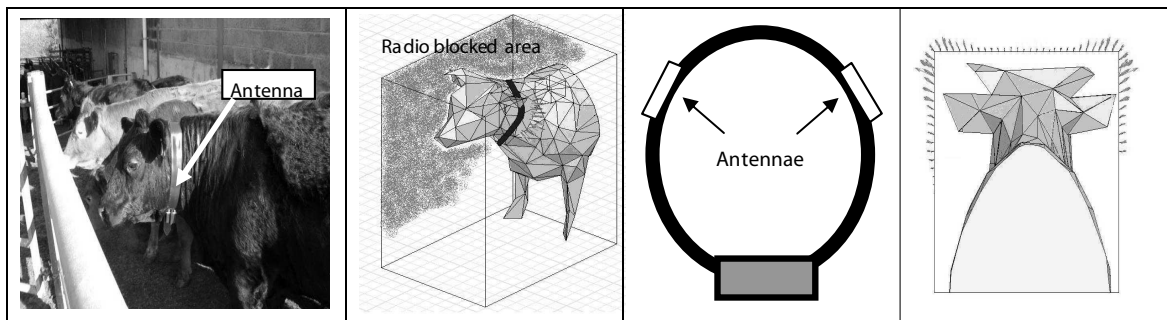

Figure 1: (a) Common single antenna placement on collar; (b) Radio signal can only propagate through one direction; (c) Enhanced antennae placement on a collar; (d) RF propagation of redesigned collar.
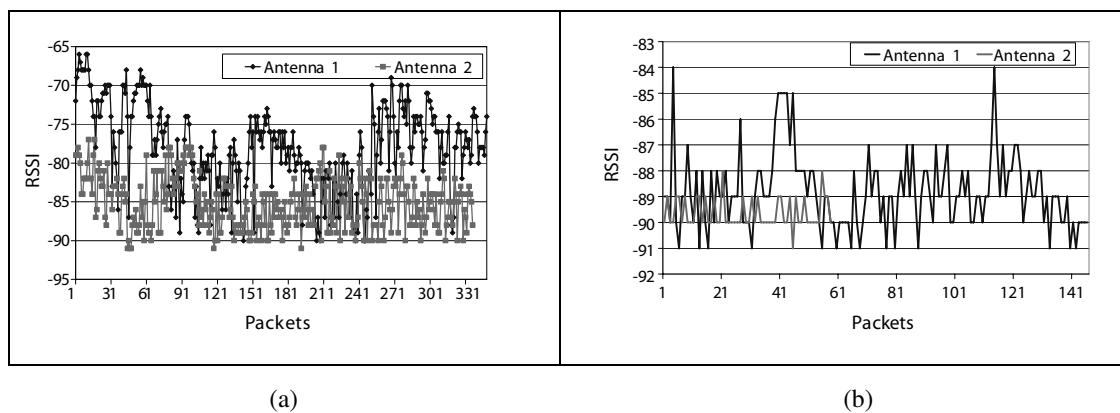

Figure 2: (a) RSSI from the both antennae with −10 dBm power transmission; (b) RSSI and received packet count for a cow at the edge of the network.

4. THE DESIGN OF IMPLICIT ROUTING PROTOCOL

In the cattle monitoring system, the connectivity between each collar is said to be sporadic leading to an unstable routing path and resulting in increased packet delay. To diminish the impact of mobility, an Implicit Routing Protocol (IRP) is designed particularly for the cattle monitoring systems. The proposed IRP works in the two phases: configuration phase and data forwarding phase. During the configuration phase, the BS periodically floods a TIER message throughout entire network. This TIER message contains a BS's ID field, and a hop count field. The hop count field is used to track the number of hops the TIER message has traveled from the base station. The tiers are numbered starting from the base station. A collar in a given tier, n, which represents it is n-th tier away from the BS. This critical information is defined as *TIER ID*. As animals are free to move around, the base station is required to send a TIER message periodically at intervals

of T_s to maintain this configuration correctly. At the data forwarding phase, if the collar is desired to report its measured data back to the base station, it will form a packet containing its current *TIER ID* and measurement data. This packet is then broadcasted data to its vicinity. Only collars which have a smaller *TIER ID* will respond with an acknowledgment (ACK). This collar after acknowledging to the source collar will broadcast the received packet. Collars that have an equal or larger *TIER ID* will discard the received data immediately. This forwarding rule will then repeat until the data arrives at the BS. The idea behind this forwarding rule is that the measured data will move one hop closer to the BS at each forwarding stage.

This section further investigates the performance of proposed routing scheme and verifies the effectiveness through empirical experiments. The IRP is implemented on the MICAz node using TinyOS [8] sensor network operation system. The testbed is configured with a 3-hop network with one source node, one base station (BS) and N pairs of intermediate relay nodes. Figure 3 illustrates the test bed configuration $N = 4$, and in each tier there are 4 relay nodes. During each experiment, the source node generates 10,000 packets in which each packet has 85 bytes in payload with an interval of 250 ms. In order to simulate the phenomena of moving cows leading to a sporadic link between any sensor nodes, an asynchronous random "on/off" mechanism is implemented. When a sensor node stays in "off" mode, it represents the cow has moved out of the communication range and when a sensor node is switched back to "on" mode it represents previous cow (or a new cow) has enter the communication range. This "on/off mechanism" is characterized by an "off" probability P_{off} which determines the probability the sensor node's radio stay at "off" mode. Figures 3 and 4 show the network performances that are impacted severely as P_{off} increases. This is due to the fact that the amount of time the sensor node in "off" state is prolonged. However, the performance is improved when the number of sensor nodes in each tier increases.
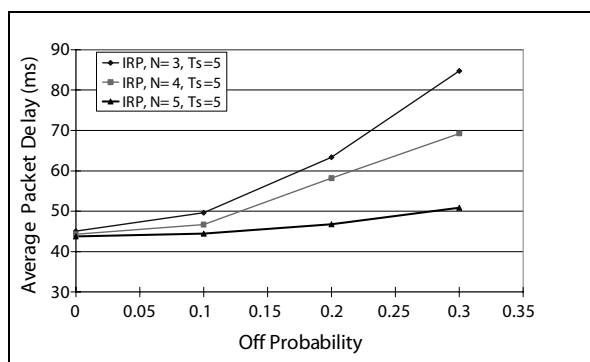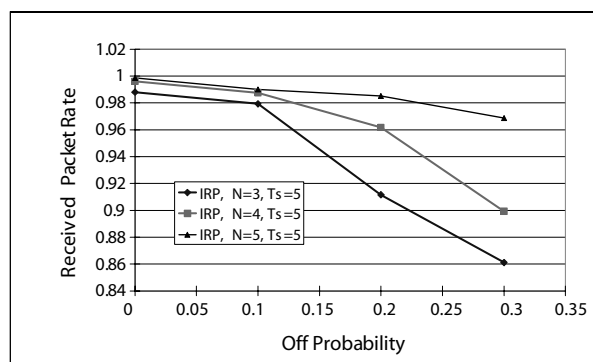

Figure 3: Average packet delay.

Figure 4: Received packet rate.

5. CONCLUSIONS

An overview of a couple of issues relating to cattle monitoring have been reported. This paper provides an alterative solution for animal monitoring system by using low-cost, low power consumption sensor nodes. To facilitate real-time reporting while overcoming mobility caused by animal movement, an Implicit Routing Protocol (IRP) is particularly designed. The experimental results indicated that the proposed IRP can significantly diminish the impact of mobility under varying "off" probability, different quantity of sensor node and the occurrence of network reconfiguration. In the near future, the designed routing protocol is expected to be used in the farm trial in order to study its operation and implication in the field.

REFERENCES

1. Agriculture in the United Kindom, https://statistics.defra.gov.uk/esg/publications/auk/2007/default.asp.

2. Matthews, K. H. and J. Buzby, "Dissecting the challenges of mad cow and foot-and-mouth disease," *Agricultural Outlook*, 4–6, August 2001.

3. Zhang, P., C. M. Sadler, S. A. Lyon, and M. Martonosi, "Hardware design experiences in ZebraNet," *Processsdings of SenSys*, Baltimore, Maryland, USA, November 2004.

4. Sikka, P., P. Corke, P. Valencia, C. Crossman, D. Swain, and G. Bishop-Hurley, "Wireless adhoc sensor and actuator networks on the farm," *Processdings of IPSN'06*, 492–499, April 2006.
5. Environmental Studies, http://www.environmental-studies.de.
6. Crossbow Technology Inc., http://www.xbow.com.
7. Gabriely, S., R. W. Lau, and C. Gabriel, "The dielectric properties of biological tissues: III. Parametric models for the dielectric spectrum of tissues," *Physics in Medicine and Biology*, Vol. 41, 2271–2293, 1996.
8. TinyOS website, http://www.tinyos.net.
9. Szewczyk, R., J. Polastre, A. Mainwaring, J. Anderson, and D. Culler, "An analysis of a large scale habitat monitoring application," *Processdings of SenSys*, Baltimore, Maryland, USA, November 2004.

Time Table Transfer Time Synchronization in Mobile Wireless Sensor Networks

Reza Khoshdelniat[1], **Moh Lim Sim**[1], **Hong Tat Ewe**[2], **and Tan Su Wei**[1]

[1]Faculty of Engineering, Multimedia University, Cyberjaya, Malaysia
[2]Faculty of Information and Communication Technology
Universiti Tunku Abdul Rahman, Malaysia

Abstract— Wireless sensor networks have the abilities to be applicable in most situations and conditions to monitor our environment and have a better control of our surrounding. In most applications having accurate time synchronization among the nodes in the network would be a crucial issue. Since all the clocks at the nodes do not work the same way, there are time offset and clock drift among the nodes. In mobile sensor networks the condition and requirements for time synchronization are different from static networks. Most of the time synchronization protocols designed for static networks will not be applicable to mobile sensor networks. We propose a new protocol for time synchronization of mobile wireless sensor networks which has a much lower convergence time with good precision and good abilities to be attack resilient. We describe how it works and its performance and also discuss the simulation results and compare its performance with other protocols.

1. INTRODUCTION

The wide application of wireless sensor networks has driven the attention of researchers and there have been lots of researches done on different protocol layers of sensor networks. Wireless sensor nodes have limitations such as small coverage range, limited power source, limited memory and computation power. Because the clocks on the nodes are imperfect, they would be relatively different from each other. In most applications such as military or health monitoring, its necessary that the nodes have accurate timing information about the other nodes in the network or all their clocks be synchronized with a specific clock. Time stamping measurements, network signal processing, localization, cooperative communication, medium access, sleeping schedules, and coordinated actuations are some situations that need synchronization in our wireless sensor network.

Some synchronization protocols such as [2, 4, 5, 7, 10] are designed and proposed for wireless sensor networks [1] which try to identify the clock offset and clock drift among the nodes. These protocols could be classified based on the way they work. One classification could be based on the way the nodes apply the synchronization correction [1], either the protocol corrects the nodes clock [3, 11] or the nodes keep the relative clock offset between themselves and their neighbors and communicate with others in the correct time [4, 10]. Some protocols such as [8, 9] have been designed with consideration of mobility among the nodes. A suitable time synchronization protocol should consider the limitations that sensor networks are facing. Convergence time which is the time it takes to synchronize the network is an important issue; it should be as small as possible. Low message overhead, low computation overhead, energy efficiency and accuracy are important requirements of a time synchronization protocol.

2. EXISTING PROTOCOLS AND CHALLENGES IN MOBILE NETWORKS

Here we have classified the time synchronization protocols into four categories as below. (i) Round trip time (rtt) based synchronizations that use the message exchange between two normal nodes in the network. The time it takes to send and receive the reply message is used to calculate the propagation delay, relative clock offset or time boundaries for their synchronization. Examples of these protocols are TPSN [4, 5, 7]. Using the round trip time is a suitable way to calculate the clock offsets in a mobile network. Although the nodes are moving but the message exchange would be done fast and before the nodes get out of each other's coverage range. (ii) Reference broadcasting based synchronizations; this category of protocols such as RBS [10] depends on the existence of several normal nodes in the coverage range of a reference node. As shown in the figure on the right in Fig. 1 right, the reference node will broadcast a synchronization message, the nodes in the coverage range will record the time of receiving the message and then exchange their recorded times to calculate the relative clock offset. The problem with this method is that

it requires the existence of at least three nodes in each other's coverage area. In mobile networks the nodes would be moving randomly in different directions with different speeds, this makes it hard to find three nodes in the same coverage area to implement the synchronization. (iii) Level based synchronizations, it has two phases. First phase is the level discovery which creates a level hierarchy in the network. Second phase is the synchronization of the nodes in the network based on the level hierarchy created. Examples of such protocols are [6–8]. The main problem with these protocols in mobile sensor networks is that the level hierarchies would not remain for long since the nodes are moving. (iv) Route based synchronization that uses a defined path among the network. Depending on the protocol, the path might pass all the nodes in the network or just go through some of the nodes. A timing message will be sent through the nodes along the path, which will record the initial time and the end time of the message. By using the message traveling time, the nodes average the time to different segments of the loop and smooth over the error of the clocks. An example of this type of protocol is [9]. Creating a path in mobile sensor networks is not easy, especially if we want the path pass through all the nodes in the network.

Figure 1: (left) The message exchange between two nodes to calculate the round trip time, (right) RBS, (1) refrence node broadcasting the sync message, (2) receiver nodes exchanging the arrival time of the sync message.

3. TIME TABLE TRANSFER (TTT) TIME SYNCHRONIZATION

The time synchronization protocol that we propose uses the round trip time (rtt) to calculate the clock offset and the propagation delay between two nodes. As shown in Fig. 1, node A starts the message transmission at time T_1. Node B receives the message and timestamps it as T_2, and sends back a reply message to node A at time T_3. The message sent by node B includes T_2 and T_3. Node A receives the message at T_4 and timestamps it upon receiving it. Since there are clock offset (Δ) between the local clocks of the nodes and also the propagation delay (d), we can have the following relations:

$$T2 = T1 + d + \Delta, \quad T4 = T3 + d - \Delta \tag{1}$$

From the above equations we could calculate the clock offset and the propagation delay as follows:

$$\Delta = \frac{(T2 - T1) - (T4 - T3)}{2}, \quad d = \frac{(T2 - T1) + (T4 - T3)}{2} \tag{2}$$

In a mobile network, the mobility of the nodes can help to move the timing data in the network and then transfer them to other nodes. We have tried to make use of the mobility of the nodes instead of facing it as a problem. Each node is free to communicate with its neighbors, but the nodes are not allowed to have repeated communications, which means a pair of nodes would be able to communicate only once during the synchronization phase and so the message overhead in the network will decrease. The clock offsets associated with each node will be kept in the memory of the node in a time table, and upon communicating with a new node the time table would be transferred to the other node. The new node, using the offset time table would be able to calculate its relative clock offset with the nodes in the time table without even having any message exchange with them which is a great advantage. The message includes T_{1s}, T_{2r}, T_{3r}, clock offset, propagation delay, source address, destination address and the clock offset time table. T_{1s} is the time that the sender node sends the message as shown in Fig. 1 left. T_{2r} is the time the receiver node time stamps upon receiving the message, and T_{3r} is the time the receiver replies back to the sender node.

For example we describe how the protocol works in Fig. 2. We assume 9 nodes are moving randomly in a network after some message exchanges the nodes have the relative offset between

themselves and the nodes that their number is in their time table. In the second stage, nodes 5 and 3 communicate and since they have not exchanged time table before, they would exchange time table and also calculate their relative clock offset. The relative clock offset between nodes A and B is shown as $\Delta(A,B)$. So node 5 will add $\Delta(3,5)$ to its table and node 3 would add $\Delta(5,3)$ to its table. Node 3 also communicates with node 2, so they exchange their time tables and also calculate their relative clock offsets. Node 2 has $\Delta(4,2)$ and $\Delta(9,2)$ in its offset time table which is sent to node 3, so node 3 will calculate its relative clock offset with node 2 which is $\Delta(2,3)$ and then will be able to calculate its relative clock offset with nodes 4 and 9 from the received offset time table. Node 3 will add $\Delta(4,3)$ and $\Delta(9,3)$ to its offset time table as well with out communicating with nodes 4 and 9. Node 3 will calculate the offset as following: $\Delta(4,3) = \Delta(4,2) + \Delta(2,3)$, $\Delta(9,3) = \Delta(9,2) + \Delta(2,3)$.
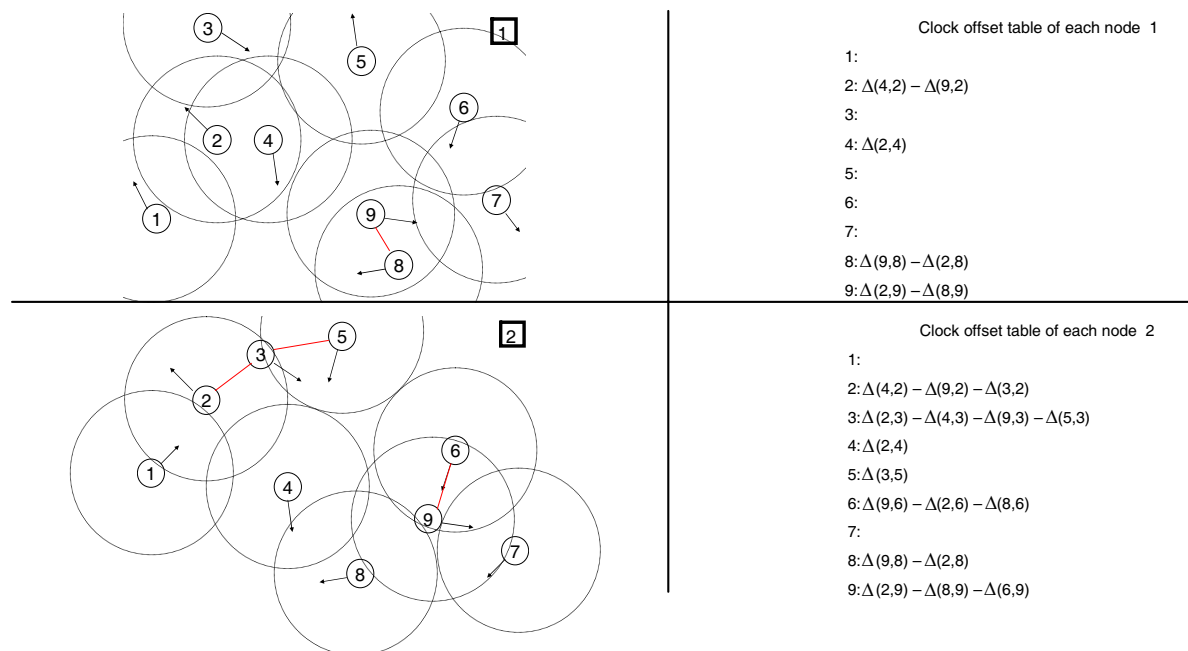

Figure 2: Time Table Transfer Time Synchronization. Nodes in the network in position 1, nodes 9 and 8 share their time tables (up), nodes after moving and getting in position 2, nodes 5, 3, 2 and also nodes 9 and 6 share their time table (down).

When the number of the nodes grows the size of the table and the size of the message will grow as well. In order to reduce the size of the message and the message overhead in the network, when a node sends the request message it will include the ID of the nodes that it does not have their clock offsets in its table. In this way when another node receives the request message, it will reply back and only include the offsets that the requesting node needs. The reason is that during our studies we realized a big part of the clock offset time table sent by the nodes would not be useful because the requesting node already has those clock offsets. So by deleting extra data from the message we would be able to significantly reduce the size of the message overhead in the network. We call this the incremental scheme for Time Table Transfer protocol.

The algorithm starts with each node sending a $REQUEST$ message after waiting for a random period of time. For example node A sends a $REQUEST$ message which includes the time the message is sent (T_{1s}) and the ID of the nodes that it does not have their clock offsets in its time table. Node A will wait for a period of time to receive any possible message. If any message was received, it will time stamp it upon receiving and then will check the type of the message. If the received message was a $REQUEST$ message from another node like node B, node A will name the time stamp of receiving the message as T_{2r} and send a $REPLY$ message to node B including the clock offsets than node B has requested, if it has them, and T_{1s}, T_{2r} and T_{3r}. If the received message was a $REPLY$ message from node C, node A will first check whether the message is destined for it or not and if it is, it will check the T_{1s} to make sure it is a $REPLY$ message to its latest $REQUEST$ message. Then node A will calculate the clock offset between itself and node C and also the propagation delay using T_{1s}, T_{2r}, T_{3r} and T_{4s}. Then node A will get the clock

offset time table in the message and by considering the relative clock offset $\Delta(A,C)$ it will adapt the clock offset to its own clock offset time table. Then node A would send an *ACKNOWLEDGE* message to node C including T_{1s}, T_{2r} and T_{3r}, the relative clock offset $\Delta(A,C)$, propagation delay, and also the clock offset of the time tables that have not been sent by node C, which means node C doesn't have them. In case the received message was a *ACKNOWLEDGE* message, node A will again check the destination address of the message, T_{1s}, T_{2r} and T_{3r} to make sure the message is the correct one. Then node A will add the clock offset in the message to its table and use it to adapt the clock offsets sent in the message to its own clock offset table. And then it will broadcast a *REQUEST* message and the algorithm continues.

In TTT (Time Table Transfer) time synchronization we have incremental time synchronization in the network, which means each node will be synchronized with part of a network before the whole network is synchronized completely. This would be a good advantage for different applications, because we don't have to wait until the network gets synchronized completely to perform a task. The nodes would be able to perform some tasks after a few rounds of message exchanges since they would be synchronized with a big part of the network. As the time passes the nodes would gain more timing information about the nodes in their network until each node has the clock offsets of all the other nodes in the network. TTT has a high accuracy as we use the round trip time to calculate the clock offset same as TPSN [5] with the average error of 16.9 μs.

4. SIMULATION

We have simulated the proposed protocol using Omnet++ simulator using its mobile framework. We simulated the protocol in different scenarios by changing the important parameters to evaluate the performance of the protocol. The parameters that have been changed are field size (m^2), speed (m/s) and different number of nodes. Fig. 3 shows the results of some of our simulations. Parameters such as convergence time (sec) and message overhead have been used to evaluate the performance. The nodes are initially placed in random positions and the direction of movement is random as well.
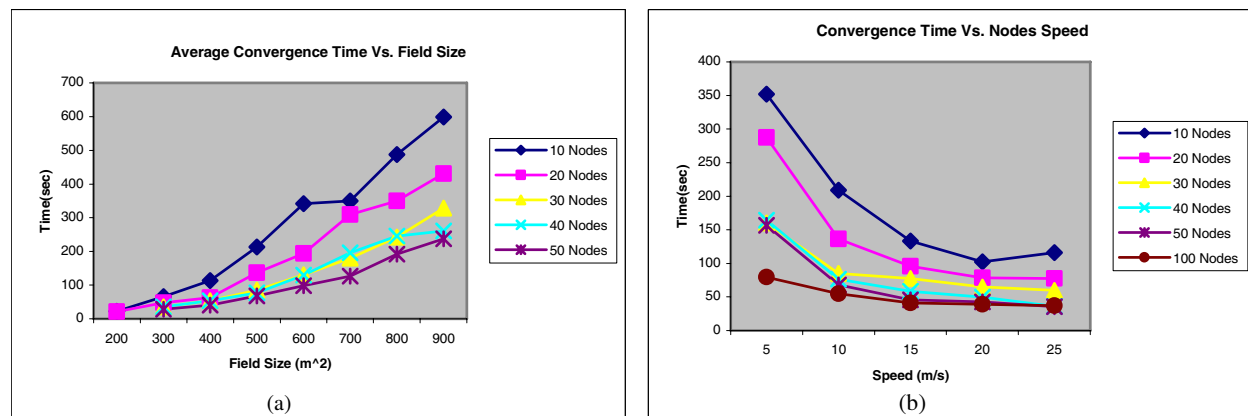

Figure 3: Simulation results on TTT.

As shown in Fig. 3(a), as the field size grows, the convergence time grows as well. The reason is the bigger distance that would appear between the nodes and it would take a longer time for them to get into each other's coverage area to communicate. It also shows that the higher the density of the nodes in the filed, the smaller the convergence time for the protocol. As it can be observed for the same field size, the higher number of nodes has less convergence time. In Fig. 3(b) and Fig. 4(a), we can see the effect of node speed on the convergence time and message overhead, with growth of the speed the nodes would be able to get in each other's coverage area faster and so exchange the messages which will result in smaller convergence time and message overhead. But when the speed increases, it will give the nodes less time to exchange messages, that's why we can see that as the speed increases, the convergence time and message overhead start to reduce as well, but as the speed grows more the convergence time and message overhead stop reducing and for higher speeds they would even start to grow. In Fig. 4(b), we have shown the results of comparison between TTT and RBS. The situations of the simulation for both protocols were the same to have a better and clear comparison. As we see the convergence time for TTT is much less than RBS,

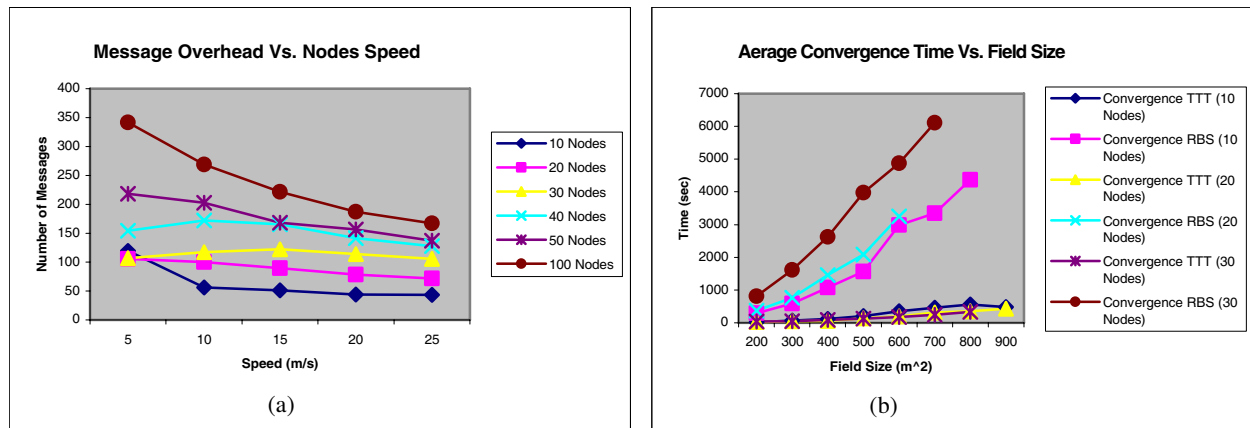same results appear for the comparison of the message overhead.

Figure 4: (a) Simulation results on TTT, (b) simulation results on comparing TTT and RBS.

5. CONCLUSION

Time Table Transfer (TTT) time synchronization protocol has been proposed for mobile networks since we use the mobility of the nodes as an advantage and tool to share the time table among the nodes. We have shown how the TTT works and studied the existing protocols to show the necessity of a good time synchronization protocol for mobile wireless sensor networks. Accuracy and desirable convergence time are features of the proposed protocol.

REFERENCES

1. Sundararaman, B., U. Buy, and A. D. Kshemkalyani, "Clock synchronization for wireless sensor networks: A survey," University of Illinois at Chicago, Mar. 2005.
2. Maróti, M., B. Kusy, G. Simon, and Á. Lédeczi, "The flooding time synchronization protocol," *Proc. 2nd Int. Conf. on Embedded Networks Sensor Systmes*, 39–49, ACM Press, Baltimore, MD, Nov. 2004.
3. Mock, M., R. Frings, E. Nett, and S. Trikaliotis, "Continuous clock synchronization in wireless real-time applications," *Proc. 19th IEEE Symposium on Reliable Distributed Systems*, 125–133, Oct. 2000.
4. Römer, K., "Time synchronization in Ad Hoc networks," *ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc) 2001*, 173–182, Long Beach, USA, Oct. 2001.
5. Ganeriwal, S., R. Kumar, and M. B. Srivastava, "Timing-sync protocol for sensor networks," *Proc. First ACM Int. Conf. on Embedded Networked Sensor Systems*, 138–149, Colifornia, L.A., Nov. 2003.
6. Su, W. and I. F. Akyildiz, "Time-diffusion synchronization protocols for sensor networks," *IEEE/ACM Transactions on Networking*, Vol. 13, No. 2, 384–397, 2005.
7. Sichitiu, M. L. and C. Veerarittiphan, "Simple, accurate time synchronization for wireless sensor networks," *Proc. IEEE WCNC 2003*, Vol. 2, 16–20, IEEE, New Orleans, LA, Mar. 2003.
8. Sun, K., P. Ning, and C. Wang, "Secure and resilient clock synchronization in wireless sensor networks," *IEEE Journal on Selected Areas in Communication*, Vol. 24, No. 2, 442–454, Feb. 2006.
9. Li, Q. and D. Rus, "Global clock synchronization in sensor networks," *Proc. IEEE INFOCOM 2004*, Vol. 1, 214–226, Mar. 2004.
10. Elson, J., L. Girod, and D. Estrin, "Fine-grained network time synchronization using reference broadcasts," *Proceedings of the Fifth Symp. Operating Systems Design and Implementation (OSDI 2002)*, 147–163, Boston, MA, Dec. 2002.
11. Ping, S., "Delay measurement time synchronization for wireless sensor networks," Intel Research, IRB-TR-03-013, June 2003.

Performance Analysis of Unsaturated Slotted IEEE 802.15.4 with Downlink and Uplink Traffic

Wei Wang[1], Lu Rong[2], and Yang Du[1]

[1]The Electromagnetics Academy at Zhejiang University
Zhejiang University, Hangzhou 310058, China
[2]Shanghai Research Center for Wireless Communications, Shanghai 200050, China

Abstract— In this work, we analyze the performance of unsaturated slotted IEEE 802.15.4 with downlink and uplink traffic in the beacon enabled mode. We assume that the coordinator has Poisson flow to every node through indirect transmission and each node has Poisson flow to coordinator through direct transmission. Based on the assumption, we study the normalized throughput of the network with different Beacon Order (BO) and traffic load, and investigate how the factors affect the performance by comparing the simulation results between downlink and uplink traffic, where the Network Simulator (NS-2) is used.

1. INTRODUCTION

IEEE 802.15TM working group designed IEEE 802.15.4 to meet the low-rate wireless personal area network (LRWPAN) which is characterized by its low complexity, low cost and low energy consumption. Several main factors have great impacts on the performance. Beacon order (BO) determines the duration of a superframe, with a range from 0 to 15. Maximum number of backoffs (macMaxCSMABackoffs) presents the maximum number of backoffs during the CSMA-CA algorithm when a note is about to perform Clear Channel Assessment (CCA). Other parameters include backoff exponent (BE), contention window (CW), and existence of ACK mechanism [1].

IEEE 802.15.4 standards define two types of transmission, direct transmission and indirect transmission. In the direct transmission, each node sends packets to the coordinator independently and access contention arises. In contention access period (CAP), CSMA/CA mechanism is used. Indirect mode is usually adopted when coordinator sends data to device. In this mode, coordinator notices the device in the beacon of next superframe and transmits the data packet after the device requests [1].

In this paper, we shall carry out a simulation based study to examine how different factors affect the network with downlink and uplink traffic. In Section 2, we will introduce our simulation environment and scenarios. Section 3 corresponds to performance of different parameters. Then we give our conclusions in Section 4.

2. SIMULATION PLATFORM

Major parameters pertinent to the simulation are shown in Table 1.

Table 1: Simulation parameters.

Effective packet length	90 bytes
Beacon order	4, 8
macMaxCSMABackoffs	4, 8
aMaxFrameRetries	3
Number of nodes	6
Traffic load	From 0.05 to 0.5 stepped by 0.05

The network operates at the 2.4 G band, with beacon enabled mode and star topology.

3. SIMULATION RESULT

3.1. Probability to Sense Channel

During the CSMA/CA mechanism, each node should sense channel after the backoff countdown. If the channel is found idle, the node starts to transmit the packet. On the other hand, if the channel is found busy at either CCA slot, the node performs the backoff stage again. Let us define τ to

be the probability to sense channel within an arbitrary slot. D_Backoffs and U_Backoffs stand for macMaxCSMABackoffs of downlink and uplink traffic.

From Figure 1, we can know that BO and macMaxCAMAbackoffs have great impacts on probability to sense channel. Node will enter idle state after a channel access failure when it is about to transmit request command to coordinator to ask for its packet and will transmit the command at next superframe. BO = 4 means more opportunity for node to sense channel than BO = 8. Bigger macMaxCSMABackoffs results in larger probability to sense channel because macMaxCSMABackoffs determines the maximum number of node to sense channel again after CCA failure.

τ of downlink traffic does not increase with the more traffic load for BO = 8, but present as a straight line, the reason of which is that the network coordinator can handle only a small amount of downlink traffic [3]. The same situation occurs in the network for BO = 4, but the coordinator can handle larger downlink traffic than BO = 8.

3.2. Throughput

The throughput is major performance metrics for 802.15.4 [3]. From Figure 2, we can find that τ has the most important impact on throughput, which means BO and macMaxCSMABackoffs decide the throughput of the network with downlink traffic.

3.3. Average Transmission Delay

Figure 3 shows the impacts on average transmission delay with different BO. It is clearly to us that coordinator can handle more downlink traffic when macMaxCSMABackoffs increases. The network will enter saturated state as downlink traffic load increases, which results in more and more downlink packets remained in the coordinator buffer and enlarge the average transmission delay.
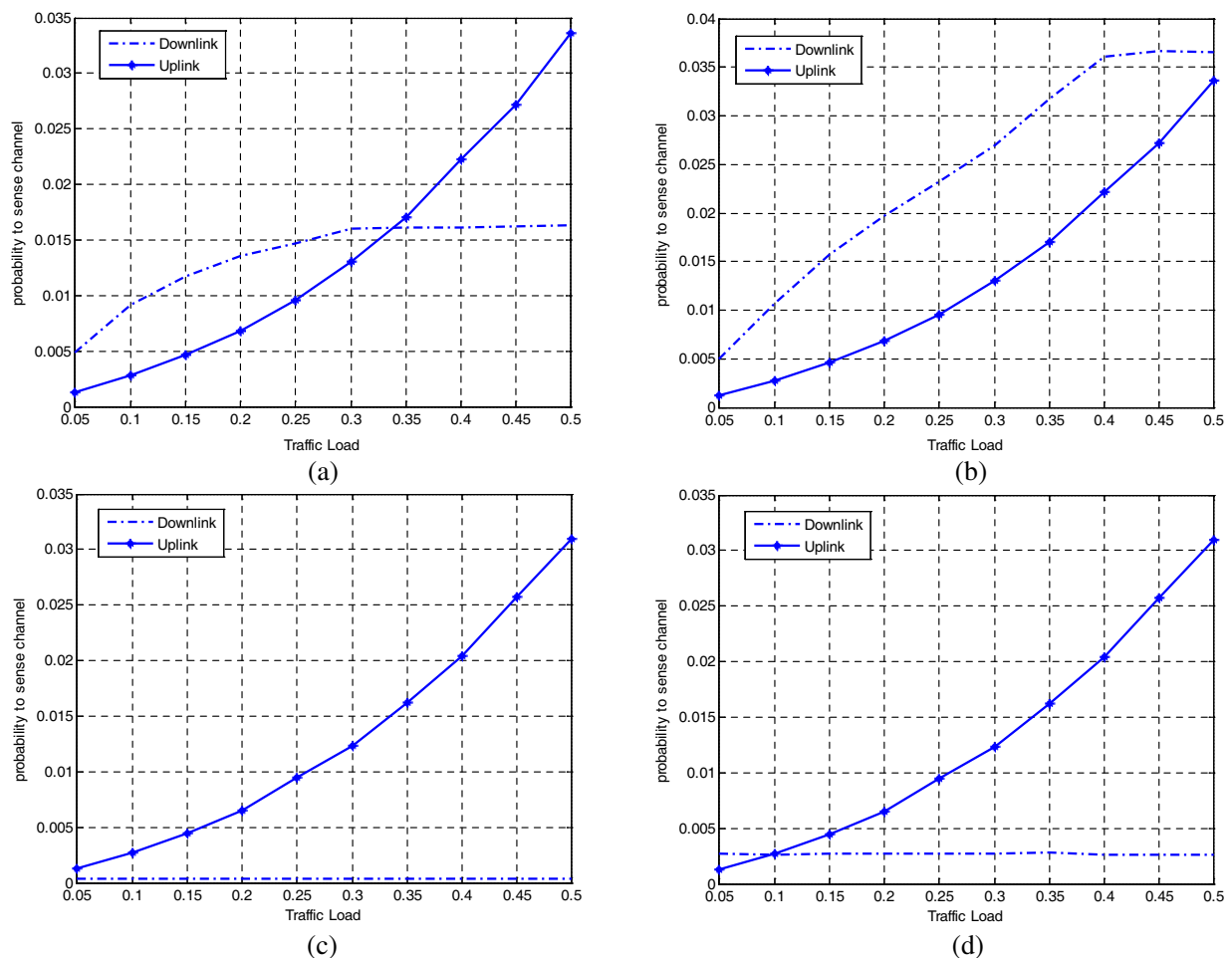

Figure 1: Probability to sense channel. (a) D_Backoffs = 4 U_Backoffs = 4 BO = 4, (b) D_Backoffs = 8 U_Backoffs = 4 BO = 4, (c) D_Backoffs = 4 U_Backoffs = 4 BO = 8, (d) D_Backoffs = 8 U_Backoffs = 4 BO = 8.
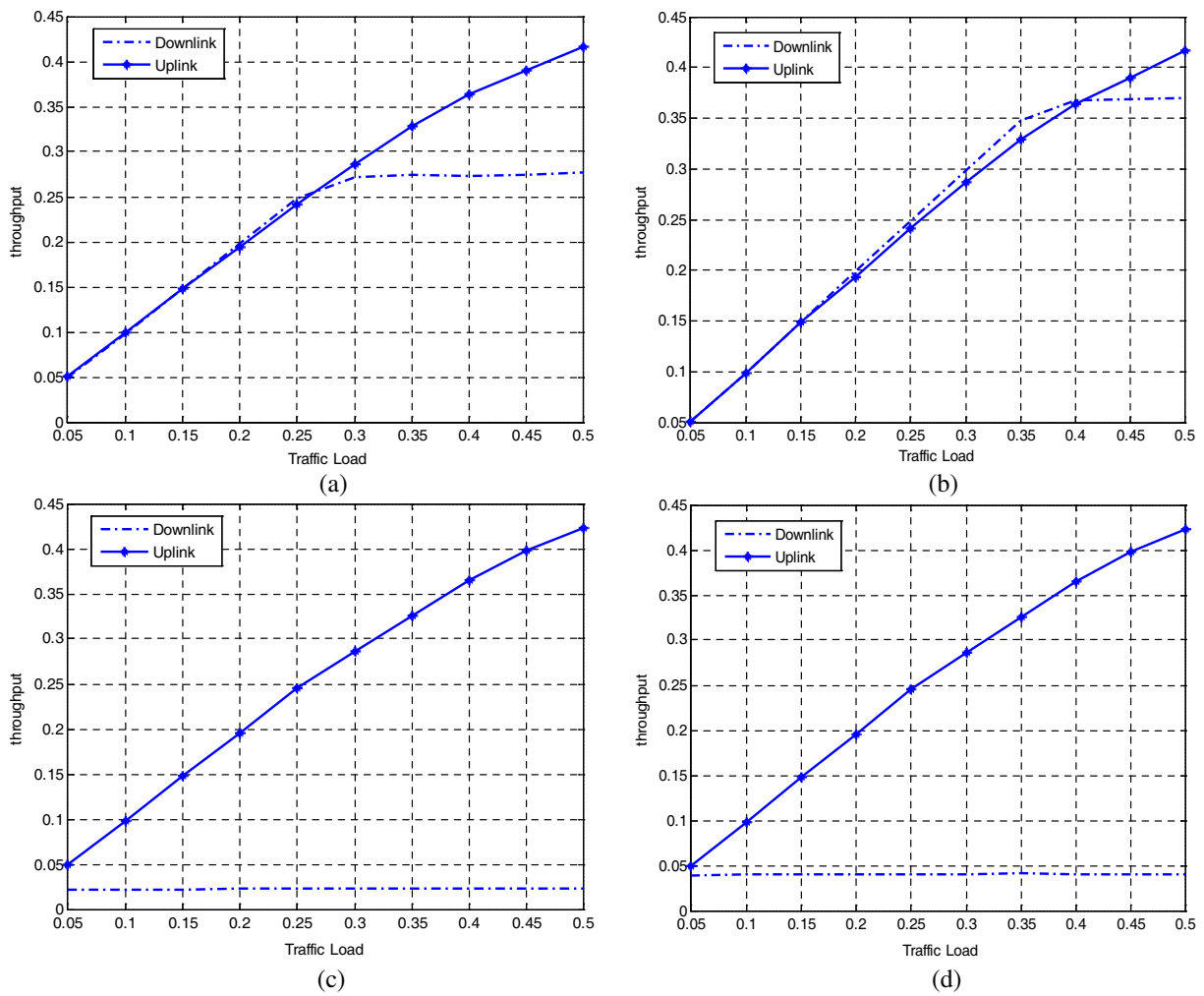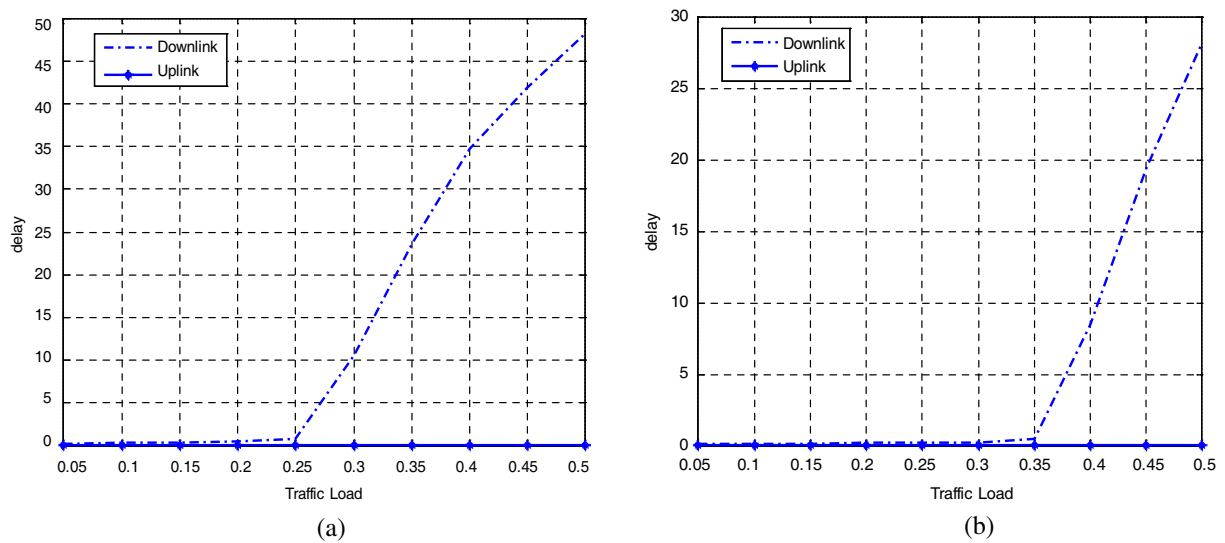
Figure 2: Throughput. (a) D_Backoffs = 4 U_Backoffs = 4 BO = 4, (b) D_Backoffs = 8 U_Backoffs = 4 BO = 4, (c) D_Backoffs = 4 U_Backoffs = 4 BO = 8, (d) D_Backoffs = 8 U_Backoffs = 4 BO = 8.

Figure 3: Transmission Delay. (a) D_Backoffs = 4 U_Backoffs = 4 BO = 4, (b) D_Backoffs = 8 U_Backoffs = 4 BO = 4.

4. CONCLUSIONS

In our work, we have analyzed the performance of Wireless Sensor Network with uplink and down-link traffic. Through the results of our simulation, we find BO and macMaxCSMABackoffs have great impacts on indirect transmission, and coordinator can handle only a small amount of down-link traffic. Shorter superframe and more opportunity for node to perform CCA will help to get better network throughput when operating with downlink traffic, but does not have much effect on uplink traffic.

REFERENCES

1. Standard for Part 15.4: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Low Rate Wireless Personal Area Networks (WPAN), IEEE Std. 802.15.4, IEEE, New York, 2003.
2. Pollin, S., M. Ergen, S. C. Ergen, B. Bougard, L. Van der Perre, F. Catthoor, I. Moerman, A. Bahai, and P. Varaiya, "Performance analysis of slotted IEEE 802.15.4 medium access layer," *Tech. Rep., DAWN Project*, 2005.
3. Park, T. R., T. H. Kim, J. Y. Choi, S. Choi, and W. H. Kwon, "Throughput and energy consumption analysis of IEEE 802.15.4 slotted CSMA/CA," *Electr. Lett.*, Vol. 17, No. 4, 361–376, Apr. 2006.
4. Misic, J., S. Shafi, and V. B. Misic, "Performance of a beacon enabled IEEE 802.15.4 cluster with downlink and uplink traffic," *IEEE Trans. Paral. Distr. Syst.*, Vol. 17, No. 4, 361–376, Apr. 2006.

Separation of Detection Authorities (SDA) Approach for Misbehavior Detection in Wireless Ad Hoc Network

Zan Kai Chong[1], **Moh Lim Sim**[1], **Hong Tat Ewe**[2], **and Su Wei Tan**[1]

[1]Multimedia University, Malaysia
[2]Tunku Abdul Rahman University, Malaysia

Abstract— Wireless ad hoc network is a distributed wireless network without the presence of a physical network infrastructure. Having a limited wireless transmission range, all the nodes need to cooperate with each other during packet forwarding in order to maintain the connectivity. In addition, the lack of energy resources may cause a selfish node to drop legitimate packets which eventually disrupt the network performance. In fact, many detection protocols have been proposed to discourage such behaviour. However, the implemented penalization is ineffective as not all the surrounding nodes execute in the same manner. Moreover, most of them have neglected the existence of the malicious node that could inflict them directly and indirectly. In this paper, we present the separation of detection authorities (SDA) approach to detect and exclude the misbehaved nodes in the network. The SDA distributes the detection authorities to the reporter, investigation agents, and the central authority where each of them takes charge of different detection process. To ensure accuracy, only the central authority will report the result. Our evaluation demonstrated that SDA can detect misbehaved nodes in the network effectively.

1. INTRODUCTION

In a wireless ad hoc network, all individual nodes has to cooperate with each other during packet forwarding primarily due to their limited transmission range and the lack of physical network infrastructure. Wireless ad hoc network provides the flexibility and scalability where the nodes are not confined to geographical restriction and are able to join or leave the network freely and randomly. Therefore, wireless ad hoc network has been widely deployed in military, scientific research, mission-critical operation and civilian application.

Despite having such benefits, wireless ad hoc network is notorious for poor administration as the wireless transmission is vulnerable to security attack. Contrary to the conventional wireless network, wireless ad hoc network does not have an absolute control over the nodes behaviour as they are owned by multiple authorities. As a result, legitimate packets may be dropped purposely by misbehaved nodes and might disrupt the network if not taken seriously.

Many research papers have discussed about the solutions to discourage erratic nodes in the network. Most emphasized on the detection scheme but neglected the importance of exclusion mechanism (penalization). For example, a simple hierarchical wireless ad hoc network has all the nodes sending data upward to the base station (root of the network). Assuming that misbehaved nodes are scattering around in the network and a downstream node wishes to penalize its upstream nodes erratic behaviour by refusing to forward packet for it. Due to the fact that the upstream node's data packet always flows upward to the base station only, the downstream penalization has no effect on the upstream node performance. Some protocols actually allow ambiguous accusation where the accuracy of the accusation is susceptible to the erratic manipulation. We proposed a detection scheme based on the separation of detection authorities (SDA) approach for misbehaviour detection in wireless ad hoc network to tackle imperfect localized penalization and ambiguous accusation.

An analogy to our SDA scheme is the operation of a democratic country where power is separated among several independent divisions for a fair execution of administrative power. In SDA, the detection authorities are distributed to the *reporter*, *investigation agents* and the *central authority* where each of them is responsible for certain part of the detection process. The victim of the misbehaved node will become the reporter of this deviation act by sending a secret accusation to the central authority. The central authority is the centrepiece that processes the accusation and issue a trustworthy conviction to the misbehaved nodes. Neighbours of the suspicious node will be appointed as the investigation agent by the central authority to verify the reliability of the accusation. Identity of the misbehaved node will be published by the central authority to all the nodes in the network, such penalization is deemed network-wide. Normally, the packet drop attack

is initiated by a single misbehaved node only. However, collaboration among several misbehaved nodes is considered in our simulation as well.

The rest of the paper is organized as follows: Section 2 reviews other related literature. Section 3 will be our system model. Next, Section 4 describes the design of SDA scheme and its operation. Section 5 is the evaluation of the SDA performance. The last section is the conclusion of this paper.

2. RELATED WORK

Felegyhazi et al. [1] presents a game theoretic model to analyse the cooperation in both dynamic and static scenarios. The simulation result shows that cooperation solely based on the self-interests of the nodes couldn't be realized in practical and an incentive mechanism is needed.

In SORI [2] all nodes maintain a confidence level table for them to exchange information with each other and penalize the bad reputation selfish node. They use one-way hashing to ensure the selfish node couldn't impersonate other nodes in improving its own reputation. However, a malicious node can always fake the information and keep condemning other innocent nodes and eventually causing a chaos in the network.

[10] is a reward-based scheme that relies on the secured module where it must be tamper resistance and protected from illegal manipulation. The secured module only feasible under a controlled environment. Thus, it is not suitable in real world implementation as the availability and the robustness of the modules are not guaranteed.

SMDP [4] is a session-based detection protocol and it use the principle of data flow conversation where the data flow in and flow out from a node should always be equal. At the end of each data session, all the nodes along the path will send the total packet they received to the previous hop and the total packet they transmitted to the next hop. After gathering all these transmission reports, all the nodes will rebroadcast the sum of the packets to the surrounding nodes. A node will be suspected if the total transmission is much different from the total reception. Digital signature has been used to ensure no one can fake the integrity of the report. However, the source can defame the next forwarder by reporting an incorrect number of total transmitted packet.

URSA [11] is a robust network access control that based on the ticket certification service through multiple node consensus and fully localized instantiation. It is a protocol that relies on the multi-signature (Threshold cryptography) to achieve the group trust model where each of the legitimate node holds a portion of the secret key (SK) and k portion of SKs are needed to renew the signature of the ticket. Our proposed SDA approach outperforms the URSA as URSA is affected by the connectivity of the network. In other words, insufficient k nodes in an area of network will cause an innocent nodes be excluded from the network. Besides, the robustness of URSA relies on the strength of the threshold cryptography and it needs to periodically refresh the network secret key to avoid the malicious node from obtaining k secret key illegally by joining the network repeatedly. SDA only requires the majority of the neighbouring nodes to be helpful to strengthen the security and will abort the investigation if insufficient of neighbouring node exist.

3. SYSTEM MODEL

In this section, the assumptions and the terminologies of this paper will be described and the attack model will be explained as well.

3.1. Assumptions and Terminologies

We assume the wireless ad hoc network is well established and all the nodes interested to communicate with the base station for some reason e.g., Internet access. Since most of the packets flowing upward to the base station, we can assume it resembles some type of hierarchical network. In addition, the central authority can be trusted absolutely and it has no incentive to misbehave. The base station is the central authority of the network and it has good knowledge of the topology of the network. Besides, we assume all the missing packets are mainly caused by the misbehaviour of the nodes. *Misbehaved node* and *misbehaver* are used interchangeably to refer to the node who does not forward packet properly and/or posses a bad intention in defaming other innocent nodes by exploiting the existing protocol.

3.2. Attack Model

We consider the packet losses are mainly due to the misbehaved nodes in network. Thus, we further classify misbehaved nodes as the *selfish nodes* and the *malicious nodes*. Selfish nodes always consider about their own benefits only and refuse to forward legitimate packets from others. Normally, we term this kind of misbehaviour as *packet drop attack*. Malicious nodes are spiteful

nodes with intention to degrade the network performance by defaming other innocent nodes. We name this type of attack as *reputation attack*.

4. SDA DESIGN

In this section we will describe our SDA design consideration and its operation flow.

4.1. Design Consideration

Most of the detection mechanisms fail to serve the primary purpose due to their improper penalization method and the ambiguous accusation. For simple illustration, 16 nodes are deployed in grid and all the nodes in the network send data packet to the base station (node a) periodically (see Figure 1). The links in the network represents the connectivity of the nodes. Node k is a misbehaved node that drops node p's legitimate packets. In this case, only node l, o and p are able to detect node k's misbehaviour (by using promiscuous listening) and they will penalize node k by dropping its packet in return. However, their penalization is useless as node k relies its upstream nodes (node j, f, and g) to forward its data packets to node a. For convenience, we name this issue as improper penalization as the penalization is not executed by all the surrounding nodes of the misbehaved node.
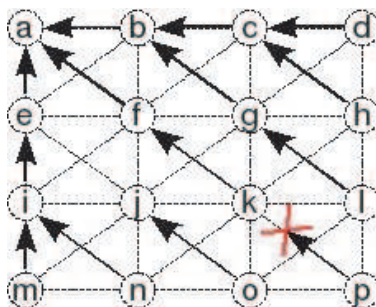

Figure 1: Simple 16 nodes in grid.

Another similar issue is the ambiguousness of the accusation where a node is unsure about the truthfulness of the accusation report sent by a neighbouring node. The accusation may be sent by a malicious node who intents to disgrace other innocent node(s). Assuming node l is penalizing node k as it is dropping node p packet. Node k can revenge back by telling node g and h that node l is a misbehaved node. In this case, node l hardly can defend for itself as node g and h are not aware of the node k misbehaviour at downstream. Example protocols that have these mentioned issue are SORI [2], two-hops acknowledgement [3], PIFA [5], and [6].

In order to overcome the issues mentioned above, we propose SDA approach in the misbehaviour detection where all the accusation will be verified by three different detection authorities. In addition, the SDA approach promotes global penalization where a detected misbehaved nodes will be published by both upstream and downstream nodes.

4.2. Detection and Accusation

Basically, SDA algorithm can be divided into three parts — the victim responsible for lodging the report, the agents investigate the accusation, and the central authority that concludes the judgement.

Firstly, the victim node accuses a misbehaved node by sending a secret accusation report to the base station through a steady route and subsequently the base station will assign a set of random k agents which are the neighbouring node of the accused node (except the accuser itself) to investigate the accusation. These agents investigate the suspected nodes by sending dummy packet with the accuser identity such that the suspected node could not aware of the investigation process. Then, these investigation agents will observe the response of the suspected node and send back the result to the base station for further action secretly. The base station will gather sufficient feedbacks and the conviction is based on the majority vote. Once the misbehaved node is convicted, its identity will be included in the base station blacklist table and sent to all the nodes in the network. Eventually, the detected misbehaving node(s) will be isolated from the network until the penalization period over.

Our proposed approach has overcome the issues mentioned above in the sense that the accusation of the victim is taken as the reference whereas the final conviction is based on the feedback of the

random appointed investigation agents. Hence, the probability of the reputation attack is kept to a very low degree. Moreover, only the central authority can issue the blacklist table for all the nodes to execute the penalization together and thus the detected misbehaver(s) will be recognized and isolated at network-wide.

5. PERFORMANCE EVALUATION

We use OMNet++ [12] to simulate 50 static nodes sending data in a network of size 1000×1000 meter. We assume all the dropped packets are mainly caused by the misbehaved nodes instead of the link error. The shortest path algorithm was used to search for the next route to forward the data to base station.

We run two types of simulations to evaluate the SDA detection effectiveness against the selfish nodes and malicious nodes threats. SDA could achieve a high correct detection percentage in an ideal network where only a small number of selfish nodes existed in the network (Figure 2(a)). As the number of selfish nodes increases, the SDA's correct detection degrade but the false positive detection is still kept to zero percentage.

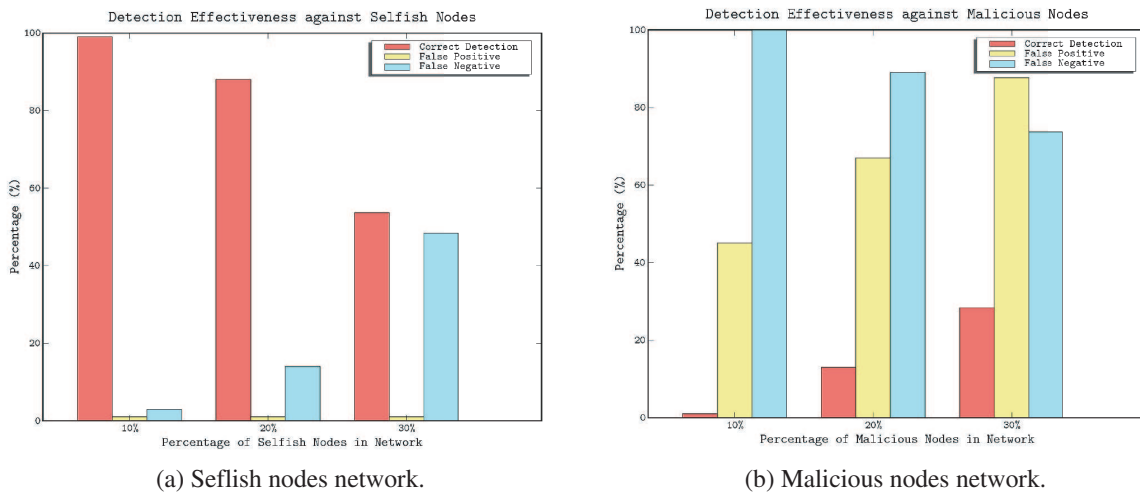(a) Seflish nodes network.　　　　　　　　　　(b) Malicious nodes network.

Figure 2: Detection effectiveness against different percentages of misbehaviour network.

In real world, multiple of malicious nodes might exist in network that threatened the innocent nodes. The increasing number of malicious nodes will augment the false positive detection in the network (Figure 2(b)). In other words, more of innocent nodes is defamed by malicious nodes.

(a) Seflish nodes network.　　　　　　　　　　(b) Malicious nodes network.
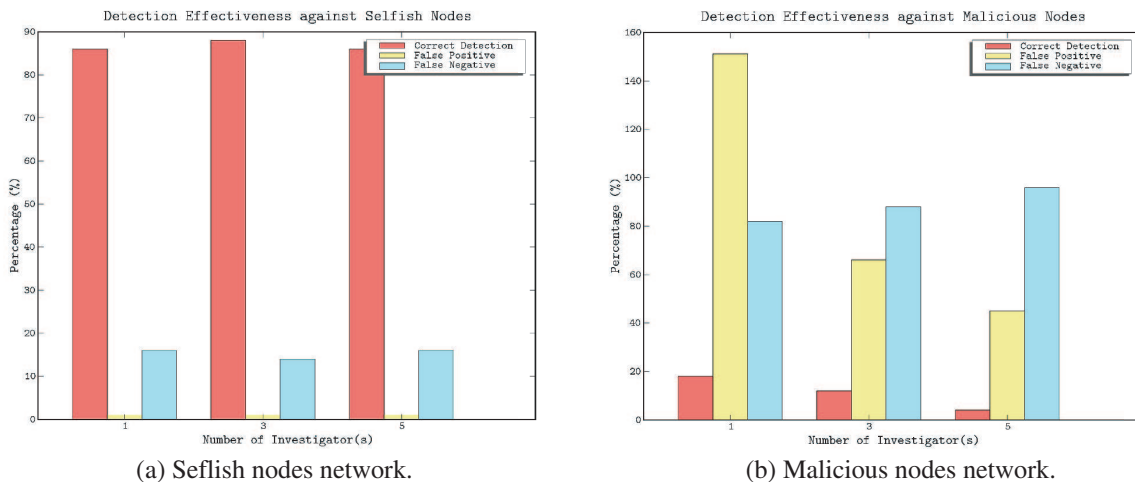
Figure 3: Detection effectiveness of different number of investigators.

Meanwhile, the correct detection percentage is increasing too as the independent malicious nodes mistakenly accuses each other.

Next, we examine the influence of the amount of investigators in SDA approach for the detection effectiveness. In an ideal selfish nodes existence network, the number of the investigators has no significant influence to the SDA's detection effectiveness as selfish nodes do not defame other innocent nodes (Figure 3(a)). However, in a malicious nodes existence network, we observed that the higher number of investigator agents could reduce the false positive detection percentage in the network (Figure 3(b)). The correct detection percentage is slightly reduced as some part of the network may have insufficient agents to complete the investigation.

6. CONCLUSION

SDA utilized a democratic approach in detection process. We believed that our SDA approach is a robust and unambiguous scheme as each accusation will be processed by different authorities and the penalization decision is recognized network-wide.

REFERENCES

1. Felegyhazi, M., J.-P. Hubaux, and L. Buttyan, "Nash equilibria of packet forwarding strategies in wireless ad hoc networks," *IEEE Transactions on Mobile Computing 2006*, 463–476, 2006.

2. He, Q., D. Wu, and P. Khosla, "SORI: A secure and objective reputation based incentive scheme for ad hoc networks," *Proc. of IEEE Wireless Communications and Networking Conference (WCNC2004)*, 2004.

3. Djenouri, D. and N. Badache, "Cross-layer approach to detect data packet droppers in mobile ad-hoc networks," *IWSOS/EuroNGI 2006*, 163–176, 2006.

4. Fahad, T., D. Djenouri, R. Askwith, and M. Merabti, "A new low cost sessions-based misbehaviour detection protocol (SMDP) for MANET," *AINA Workshops*, Vol. 1, 882–887, 2007.

5. Yoo, Y., S. Ahn, and D. P. Agrawal, "A credit-payment scheme for packet forwarding fairness in mobile ad hoc networks," *IEEE International Conference on Communications (ICC 2005)*, May 16–20, 2005.

6. Gonzalez Duque, O. F., G. Ansa, M. Howarth, and G. Pavlou, "Detection and accusation of packet forwarding misbehavior in mobile ad-hoc networks," *Journal of Internet Engineering*, Vol. 2, No. 1, 2008.

7. Buchegger, S. and J. Y. Le Boudec, "Performance analysis of the CONFIDANT protocol: Cooperation of nodes fairness in dynamic ad-hoc networks," *Proceedings of IEEE/ACM Symposium on Mobile Ad Hoc Networking and Computing (MobiHOC)*, Lausanne, CH, June 2002.

8. Michiardi, P. and R. Molva, "CORE: A collaborative reputation mechanism to enforce node cooperation in mobile ad hoc networks," *IFIP — Communication and Multimedia Security Conference 2002*, 2002.

9. Wang, Y. and M. Singhal, "A light-weight solution for selfish nodes problem considering battery status in wireless ad-hoc networks," *IEEE International Conference on Wireless and Mobile Computing, Networking and Communsications, (WiMob'2005)*, Vol. 3, 299–306, August 22–24, 2005.

10. Tan, C.-W. and S. K. Bose, "Enforcing cooperation in an ad hoc network using a cost-credit based forwarding and routing approach," *Wireless Communications and Networking Conference, WCNC 2007*, IEEE, 2935–2939, 2007.

11. Luo, H., J. Kong, P. Zerfos, S. Lu, and L. Zhang, "URSA: Ubiquitous and robust access control for mobile ad-hoc networks," *IEEE/ACM Transactions on Networking*, Vol. 12, No. 6, 1049–1063, December 2004.

12. Drytkiewicz, W., S. Sroka, V. Handziski, A. Koepke, and H. Karl, "A mobility framework for OMNeT++," *3rd International OMNeT++ Workshop*, Budapest University of Technology and Economics, Department of Telecommunications Budapest, Hungary, January 2003.

Optimal Power Allocation Algorithm for AF BAT Relaying

Li Li and Yang Du

The Electromagnetics Academy at Zhejiang University
Zhejiang University, Hangzhou 310027, China

Abstract— The Amplifier-and-Forward (AF) Bi-directional Amplification of Throughput (BAT) relaying is one of the promising techniques to enhance the throughput of the relay-based bi-directional transmission. However, the AF BAT protocol only works effectively at symmetric channels in the bi-directional scenarios. In order to further enhance the system performance in various channel propagations, we study the optimal power allocation algorithm for the AF BAT protocol in this paper. The numerical results show that significant performance improvement is achieved for the AF BAT relaying after the power allocation algorithm is applied, especially for the transmission in asymmetric channels. Furthermore, compared with the DF BAT relaying, the AF BAT protocol with optimal power allocation achieves higher spectrum efficiency and is more robust to the various channel propagations.

1. INTRODUCTION

The relay technique [1], which takes the advantage of cooperative diversity, is agreed to be one of the most attractive and practical way to enhance the system performance with higher spectrum efficiency [2] and low interference. The chief limitation of the relay protocol is the shortage at throughput, since the relay node cannot transmit and receive simultaneously. Compared with the traditional bi-directional relay (TBR) protocol, the Decode-and-Forward (DF) Bi-directional Amplification of Throughput (BAT) [3] is an effective way to improve the transmit rate by utilizing the network coding algorithm. To further improve the peak normalized throughput, later in [4], P. Popovski and H. Yomo proposed the Amplifier-and-Forward (AF) BAT, which utilize the inherent packet combining from multiple access channel. However, the AF BAT algorithm only works effectively when the traffic intensity is symmetric in both $S1$-R and $S2$-R channels. When these two channels are asymmetric, the throughput of the AF BAT relaying will be reduced significantly. Thus proper modification should be concerned for the AF BAT algorithm.

In this paper, we study the optimal power allocation (PA) algorithm for the AF BAT relaying transmission. The PA algorithm is a promising and challenge work to further improve the relay system capacity, especially for the condition when the two channels are asymmetric. The rest of the paper is organized as follow, Section 2 introduces different protocol based on relay algorithm; Section 3 gives the PA algorithm for the AF BAT relaying. The performance is evaluated in Section 4, and Section 5 gives the conclusion.

2. PROTOCOL DESCRIPTION

We consider a three-node wireless network as the system model, which includes two terminals $S1$ and $S2$ and the relay node R, and assume each node is equipped with a single antenna. As illustrated in Figure 1, the parameters $G_{S1,R}$, $G_{S2,R}$ and $G_{S1,S2}$ represent the large scaled path fading of links $S1$-R, $S2$-R, and $S1$-$S2$ respectively. During the bi-directional relaying transmission, the signals x_1 and x_2 are destined from $S1$ to $S2$ and $S2$ to $S1$, respectively.

The TBR is illustrated in Figure 2(a). During the first two slots, $S1$ and $S2$ transmit signals x_1 and x_2 to R separately. Then R forwards x_1 and x_2 to $S2$ and $S1$ during the next two slots. Thus the TBR relaying spends 4 time slots for a complete circuit of transmission, and the transmit efficiency for the TBR is only 1/2.
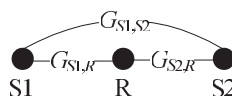

Figure 1: A three-node relay model.

The DF BAT relaying, as illustrated in Figure 2(b), applies the network coding algorithm and broadcasts the signal $x_R = x_1 \oplus x_2$ at the 3rd time slot, where $\oplus$ denotes the bitwise XOR operation. Since $S1$ already has x_1, it can extracts x_2 through $x_2 = x_1 \oplus x_R$. Analogously, $S2$
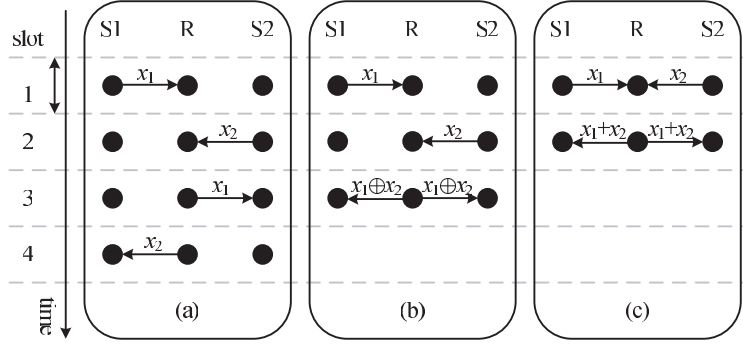
Figure 2: Different protocols for bi-directional relaying: (a) TBR; (b) DF BAT; (c) AF BAT.

extracts $x_1 = x_2 \oplus x_R$. The DF BAT protocol requires 3 time slots for the complete bi-directional transmission, thus the transmit efficiency rises to 2/3.

The AF BAT relaying, as shown in Figure 2(c), both $S1$ and $S2$ transmit x_1 and x_2 to R simultaneously in the first time slot, R receives $r_R = x_1 + x_2$. Then R broadcasts r_R in the second slot. Since $S1$ and $S2$ have x_1 and x_2 in buffer, x_2 and x_1 can be extracted by $x_2 = r_R - x_1$ and $x_1 = r_R - x_2$ separately. There're only two time slots in AF BAT protocol, thus the transmit efficiency is up to 1.

3. POWER ALLOCATION FOR AF BAT

3.1. System Throughput

In this subsection, we give the throughput of the AF BAT relaying. Without loss of generality, assume the information signals x_1 and x_2 are zero mean and with unit variance, as defined in [4].

During the first time slot, both $S1$ and $S2$ transmit x_1 and x_2 to R, thus the received signal r_R at R is

$$r_R = \sqrt{G_{S1,R}E_{S1}}x_1 + \sqrt{G_{S2,R}E_{S2}}x_2 + z_R, \tag{1}$$

where E_{S1} and E_{S2} are the transmit powers for signals x_1 and x_2 at nodes $S1$ and $S2$, and z_R is the additive Gaussian white noise at R with variance N_0.

In the second time slot, R amplifies the received signal z_R by a factor A_R and forwards to $S1$ and $S2$. The signal received at the terminal $S1$ is

$$
\begin{aligned}
r_{S1} &= \sqrt{G_{S1,R}}A_R r_R + z_{S1} \\
&= G_{S1,R}\sqrt{E_{S1}}A_R x_1 + \sqrt{G_{S1,R}G_{S2,R}E_{S2}}A_R x_2 + \sqrt{G_{S1,R}}A_R z_R + z_{S1}, \tag{2}
\end{aligned}
$$

where z_{S1} is the Gaussian noise with variance N_0 at $S1$, and the amplify factor A_R is defined as

$$A_R = \sqrt{\frac{E_R}{G_{S1,R}E_{S1} + G_{S2,R}E_{S2} + N_0}}, \tag{3}$$

where E_R is the transmit power for signal r_R at R.

Assume each node has explicit channel state information (CSI) [5], $S1$ can eliminate the interference by x_1 completely, such as

$$
\begin{aligned}
\bar{r}_{S1} &= r_{S1} - G_{S1,R}\sqrt{E_{S1}}A_R x_1 \\
&= \sqrt{G_{S1,R}G_{S2,R}E_{S2}}A_R x_2 + \sqrt{G_{S1,R}}A_R z_R + z_{S1}. \tag{4}
\end{aligned}
$$

The received SNR for x_2 at $S1$ is

$$SNR_{S1} = \frac{G_{S1,R}G_{S2,R}\gamma_{S2}\gamma_R}{G_{S1,R}\gamma_R + G_{S1,R}\gamma_{S1} + G_{S2,R}\gamma_{S2} + 1}, \tag{5}$$

where $\gamma_{S1} = \frac{E_{S1}}{N_0}$, $\gamma_{S2} = \frac{E_{S2}}{N_0}$ and $\gamma_R = \frac{E_R}{N_0}$. Analogously, the received SNR for x_1 at $S2$ is

$$SNR_{S2} = \frac{G_{S1,R}G_{S2,R}\gamma_{S1}\gamma_R}{G_{S2,R}\gamma_R + G_{S1,R}\gamma_{S1} + G_{S2,R}\gamma_{S2} + 1}. \tag{6}$$

According to the description of AF BAT relaying, the overall system throughput is defined as

$$C = \frac{1}{2}\{\log_2(1 + SNR_{S1}) + \log_2(1 + SNR_{S2})\}. \tag{7}$$

3.2. Power Allocation Algorithm

The protocol in [4] is denoted as the equal power allocated (E-PA) AF BAT, since each node transmits at equivalent power E_s. The E-PA AF BAT relaying only works effectively at the symmetric channels, i.e., $G_{S1,R} = G_{S2,R}$, and its performance will be significantly reduced when the bi-directional channels become asymmetric. In order to eliminate the performance attenuation in various channel propagations, we study the optimal power allocation (O-PA) algorithm in this subsection.

During the analysis, we define the power constraint as

$$\frac{\gamma_{S1} + \gamma_{S2} + \gamma_R}{3} = \gamma_s = \frac{E_s}{N_0}, \tag{8}$$

where E_s represents the average transmit power, which is the same as that in E-PA one. Denote $\alpha_{S1} = \frac{3G_{S1,R}\gamma_s + 1}{G_{S2,R} - G_{S1,R}}$, $\alpha_{S2} = \frac{3G_{S2,R}\gamma_s + 1}{G_{S1,R} - G_{S2,R}}$, and $G = \frac{G_{S1,R}G_{S2,R}}{G_{S2,R} - G_{S1,R}}$, thus $\alpha_{S1} + \alpha_{S2} = -3\gamma_s$.

The optimal power allocation algorithm is designed to maximize the system throughput of the AF BAT relaying, such as

$$C_{\max} = \max_{\gamma_{S1}, \gamma_{S2}} \frac{1}{2}\{\log_2(1 + SNR_{S1}) + \log_2(1 + SNR_{S2})\}, \tag{9}$$

under the constraint (8), where

$$SNR_{S1} = \frac{G\gamma_{S2}(3\gamma_s - \gamma_{S1} - \gamma_{S2})}{\gamma_{S2} + \alpha_{S1}}, \tag{10}$$

$$SNR_{S2} = \frac{-G\gamma_{S1}(3\gamma_s - \gamma_{S1} - \gamma_{S2})}{\gamma_{S1} + \alpha_{S2}}. \tag{11}$$

The system throughput C could be further expressed as

$$C = \frac{1}{2}\log_2(1 + T_1 + T_2), \tag{12}$$

where $T_1 = SNR_{S1}SNR_{S2}$ and $T_2 = SNR_{S1} + SNR_{S2}$.

We study the O-PA algorithm for term T_1 first. The maximum T_1 is achieved at

$$\begin{cases} \frac{\partial T_1}{\partial \gamma_{S1}} = 0 \\ \frac{\partial T_1}{\partial \gamma_{S2}} = 0 \end{cases}. \tag{13}$$

Through numerical analysis, (13) could be further simplified as

$$\begin{cases} (3\gamma_s - \gamma_{S1} - \gamma_{S2})\alpha_{S2} = 2\gamma_{S1}^2 + 2\gamma_{S1}\alpha_{S2} \\ (3\gamma_s - \gamma_{S1} - \gamma_{S2})\alpha_{S1} = 2\gamma_{S2}^2 + 2\gamma_{S2}\alpha_{S1} \end{cases}. \tag{14}$$

Thus optimal power allocation for T_1 could be achieved at $(\gamma_{S1}^{T_1*}, \gamma_{S2}^{T_1*}, \gamma_R^{T_1*})$ by solving (14). Similarly, the optimal T_2 can be achieved at

$$\begin{cases} \frac{\partial T_2}{\partial \gamma_{S1}} = 0 \\ \frac{\partial T_2}{\partial \gamma_{S2}} = 0 \end{cases}. \tag{15}$$

Through numerical analysis, (15) could be further simplified to

$$\alpha_{S1}\alpha_{S2}(3\gamma_s - 2\gamma_{S1} - 2\gamma_{S2}) = \alpha_{S1}\gamma_{S1}^2 + \alpha_{S2}\gamma_{S2}^2, \tag{16}$$

which could be satisfied at $(\gamma_{S1}^{T_1*}, \gamma_{S2}^{T_1*}, \gamma_R^{T_1*})$. Thus the optimal power allocation for AF BAT relaying is achieved at

$$(\gamma_{S1}^*, \gamma_{S2}^*, \gamma_R^*) = (\gamma_{S1}^{T_1*}, \gamma_{S2}^{T_1*}, \gamma_R^{T_1*}). \tag{17}$$

4. SIMULATION RESULT

We follow the model in Figure 1 for transmission. A common large-scaled path fading model $G_{i,j} = d_{i,j}^{-3}$ is used, where $d_{i,j}$ is the normalized distance from node i to node j, $(i, j \in 0, 1, 2)$. We normalize the distance between $S1$ and $S2$ as $d_{S1,S2} = 1$. Denote the distance from $S1$ to R as $d_{S1,R} = dist$, and the distance between $S2$ and R is $d_{S2,R} = 1 - dist$, where $0 < dist < 1$.

As illustrated from Figure 3, we first compare the performance of the AF BAT protocols with E-PA and O-PA algorithms at $\gamma_s = 0$ dB. As we can find from the simulation, the throughput for the E-PA AF BAT relaying is maximized when the bi-directional channels are symmetric at $dist = 0.5$, in this case $G_{S1,R} = G_{S2,R}$. We can expect that the throughput comes down when the bi-directional channels become asymmetric. Especially when R is close to $S1$ ($dist = 0.1$) or $S2$ ($dist = 0.9$), the E-PA AF BAT relaying achieves the throughput nearly half of that at $dist = 0.5$.

The O-PA algorithm, given in Section 3, is introduced to further improve the performance of the AF BAT relaying. Compared with the E-PA AF BAT protocol, the O-PA AF BAT relaying not only enhances the maximal throughput at $dist = 0.5$, but also improves the performance at the asymmetric channels significantly. Thus we can expect that the O-PA AF BAT relaying is more robust for various channel propagations.

Then we compared the performance of the DF BAT and AF BAT protocols at different channel conditions. Figure 4 demonstrates the throughput at $dist = 0.5$. We may find that slight PA gain (about 0.5 dB) is achieved through O-PA algorithm. Both the AF BAT protocols work not as good as the DF BAT protocol at low γ_s, that's because the combining operation "+" in AF BAT reduces the transmit power at R for each signal x_1 and x_2. However, when γ_s increases, the high transmit efficiency guarantees the AF BAT protocol performs far beyond the DF BAT one.
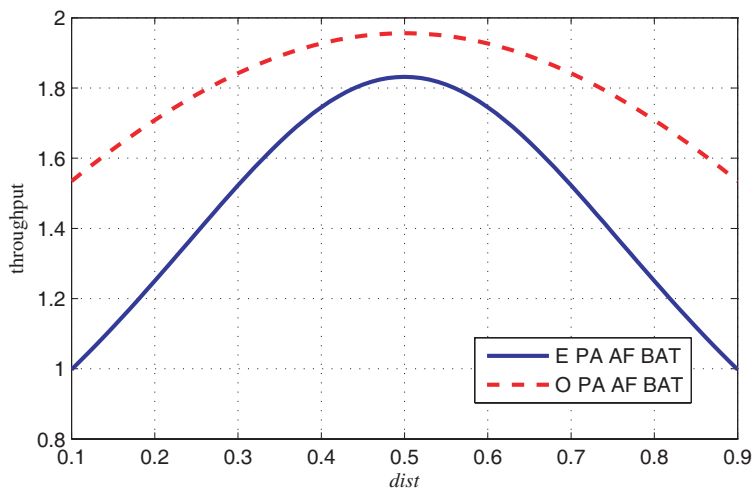

Figure 3: Throughput for AF BAT relaying with E-PA and O-PA at $\gamma_s = 0$ dB.
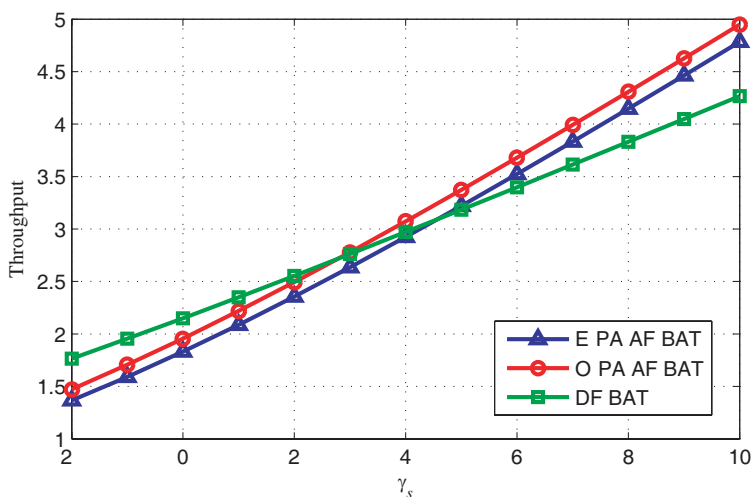

Figure 4: DF and AF BAT performance at $dist = 0.5$.

When the system works in the asymmetric scenarios, the O-PA AF BAT relaying takes its superiority for being adjustable to various channel propagations. As illustrated from Figure 5, in this case ($dist = 0.2$ or 0.8) the PA gain rises up to around $2\,\mathrm{dB}$. Compared with the DF BAT relaying, the E-PA AF BAT protocol also shows its robustness in asymmetric channels. The E-PA AF BAT relaying not only achieves significant throughput improvement at high γ_s, but also performs similarly at low γ_s.
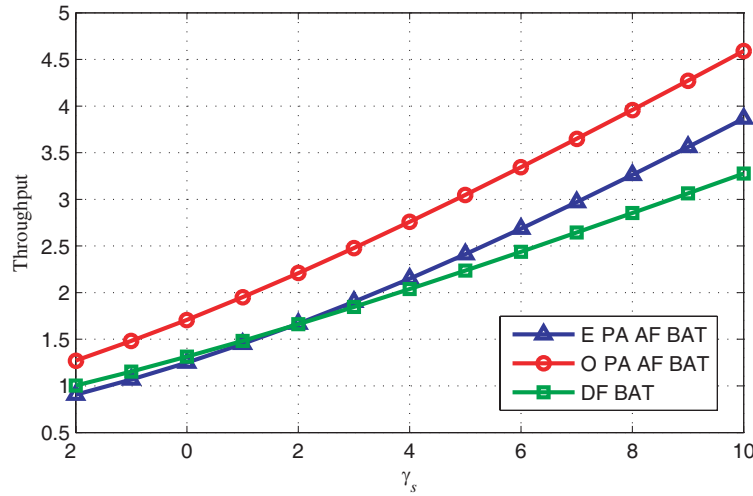
Figure 5: DF and AF BAT performance at $dist = 0.2$ or 0.8.

5. CONCLUSION

Compared with the traditional relaying protocol, the AF BAT relaying is a promising way to improve the spectrum efficiency during the bi-directional transmission. However, the performance of the AF BAT relaying is greatly attenuated in asymmetric channels. In order to enhance the robustness of AF BAT protocol, we study the O-PA algorithm in this paper. Numerical results show that the O-PA AF BAT relaying not only achieves significant PA gain at asymmetric channels, but also has slight performance improvement at symmetric channels.

REFERENCES

1. Fitzek, F. H. P. and M. D. Katz, *Cooperative in Wireless Networks: Principles and Applications*, Springer, 2006.
2. Laneman, J. N. and G. W. Wornell, "Energy-efficient antenna sharing and relaying for wireless networks," *Proc. WCNC*, 2000.
3. Larsson, P., N. Johansson, and K.-E. Sunell, "Coded bi-directional relaying," *5th Scandinavian Workshop on Ad Hoc Networks, (ADHOC'05)*, 2005.
4. Popovski, P. and H. Yomo, "Bi-directional amplification of throughput in a wireless multi-hop network," *Proc. IEEE VTC 2006 Spring*, 2006.
5. Host-Madsen, A. and J. Zhang, "Capacity bounds and power allocation for wireless relay channels," *IEEE Trans. Inf. Theory*, Vol. 51, No. 6, 2020–2040, 2005.

A Compact and Low Loss V-band Lowpass Filter Using Coplanar Waveguide Structure

Hwann-Kaeo Chiou and I-Shan Chen

Department of Electrical Engineering, National Central University, Taiwan

Abstract— This paper presents a compact finite width ground planes coplanar waveguide (FG-CPW) low-pass filter (LPF) for V-band wireless personal area network (WPAN) communication systems. The FG-CPW equivalent circuit is constructed by a series inductor at both ends with two shunted inductor-capacitor (LC) resonators at the center of the circuit, which the series inductor can be realized as a series-short-stub in CPW line. The shunted series LC resonator is realized as a high impedance CPW line series with a low impedance CPW open stub. The full wave EM simulation was performed by Agilent Momentum$^{\text{TM}}$ to predict the frequency response. The CPW LPF was implemented in WIN-semiconductor$^{\text{TM}}$ 0.15-µm pHEMT technology. The simulated results show a good agreement with measurement, the obtained insertion loss is small than 0.5 dB with a return loss better than 20 dB. The 3-dB bandwidth is from 0 to 70 GHz. A better than 20-dB rejection is obtained from 95 to 110 GHz. The measured frequency responses agree well with the simulated ones. The chip size is very compact of $0.45 \times 0.4 \, \text{mm}^2$.

1. INTRODUCTION

The 60-GHz band is of much interest for wireless personal area network (WPAN) applications since it offers a bandwidth up to 7 GHz for multimedia wireless transmission [1]. MMIC-integration for a 60-GHz RF system of a potentially high volume demands the commercial manufacture of a small, low-cost and highly integrated transceiver. Compact FG-CPW LPF appears to be a key component to meet this requirement. FG-CPW LPF with harmonic-suppressed and wide stop-band properties has been highly required in millimeter systems to reject the spurious responses caused by nonlinear devices, such as high power amplifiers, mixers, and oscillators, etc. For this purpose, a lumped element such as an on-wafer capacitor or a resistor has been incorporated in the distributed line circuits in order to break their periodicity with respect to frequency. Another approach is to employ a periodic band-gap (PBG) structure [2] or a defected ground structure (DGS) [3]. Especially, since a DGS has a simple equivalent circuit model and yields a low-pass property with a wide stop-band, many research activities have been performed in order to apply it to LPF design [4]. The CPW composite structures are being used widely to design the millimeter components, such as lowpass filters [5]. However, the CPW composite structures have rarely applied to the LPF with harmonic-suppressed and wide stop-band properties. In this paper, a newly harmonic-suppressed and wide stop-band FG-CPW LPF with sharp rejection slope is proposed. Additionally, the designed LPF is very compact whose area on-wafer is $0.45 \times 0.4 \, \text{mm}^2$.

2. DESIGN FG-CPW LPF

The FG-CPW LPF structure has demonstrated the slow-wave passband and stopband characteristics, making it a lowpass performance with compact size. Figure 1(a) illustrates the three-dimensional view structure of the FG-CPW LPF. For design verification, the FG-CPW LPF is fabricated on a GaAs substrate with a thickness of 500 µm, a dielectric constant $\varepsilon_r = 12.9$, and a low loss tangent $(\tan \delta)$ of 0.005 in WIN$^{\text{TM}}$ 0.15-µm pHEMT technology. Figure 1(b) displays the cross-sectional view and Figure 1(c) shows the top view dimensions. Each element in the circuit model has a definite connection with the physical dimension of the EBG structure such that the passband and stopband characteristics are easy to control [6]. From Figure 2, the ABCD matrix of the equivalent circuit can be determined

$$
\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} 1 + \dfrac{Z_1(Z_1 + 2Z_2 + Z_3)}{Z_2 Z_3^2} & Z_1\left(2 + \dfrac{Z_1}{Z_2}\right) \\ \dfrac{(Z_1 + Z_2)^2 + 2Z_2(Z_1 + Z_3)}{Z_2 Z_3^2} & 1 + \dfrac{Z_1(Z_1 + 2Z_2 + Z_3)}{Z_2 Z_3} \end{bmatrix} \quad (1)
$$

where

$$
Z_1 = j\omega L_S, \quad Z_2 = j\omega \frac{L_g}{2} + \frac{1}{j\omega 2 C_g}, \quad Z_3 = \frac{1}{j\omega 2 C_g}
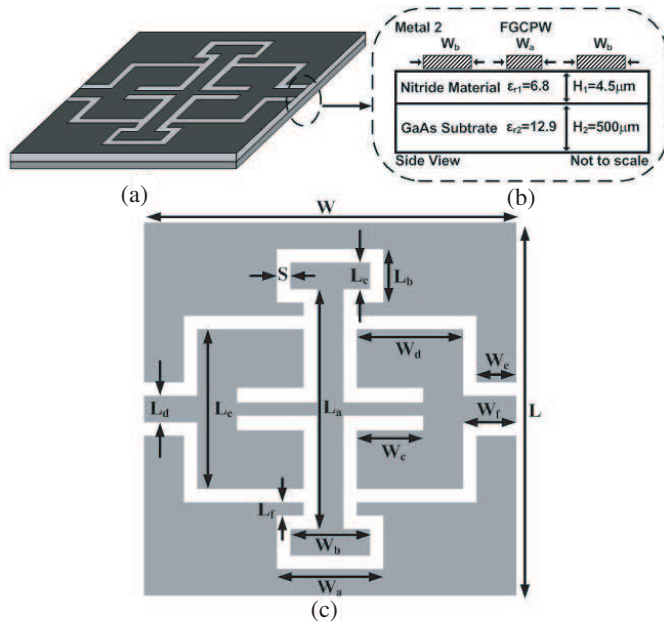$$

Figure 1: Configuration of the FG-CPW LPF: (a) three-dimensional view, (b) cross-sectional view, and (c) top view dimensions: ($W = 400\,\mu m$, $W_a = 150\,\mu m$, $W_b = 110\,\mu m$, $W_c = 100\,\mu m$, $W_d = 155\,\mu m$, $W_e = 20\,\mu m$, $W_f = 50\,\mu m$, $L = 400\,\mu m$, $L_a = 290\,\mu m$, $L_b = 80\,\mu m$, $L_c = 40\,\mu m$, $L_d = 50\,\mu m$, $L_e = 160\,\mu m$, and $L_f = 30\,\mu m$).
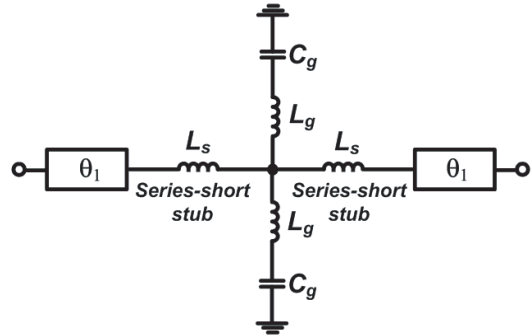
Figure 2: Lumped-element equivalent-circuit models for the FG-CPW LPF structures.

The transmission coefficient and its phase angle can then be obtained by using the following expressions:

$$S_{21} = \frac{2}{A + B/Z_0 + CZ_0 + D} \tag{2}$$

$$\theta_{21} = \tan^{-1}\left(\frac{\mathrm{Im}(S_{21})}{\mathrm{Re}(S_{21})}\right) \tag{3}$$

Based on the equivalent circuit, the characteristic of the compact lowpass resonators is developed. The layout of the proposed lowpass resonator, which is connected by series inductor at both ends with two shunted inductor-capacitor (LC) resonators at the center of the circuit, is shown in Figure 1(c). As can be seen, the series inductor can be realized as a series-short-stub in CPW line. The shunted series LC resonator is realized as a high impedance CPW line series with a low impedance CPW open stub. The FG-CPW LPF is characterized by the equivalent circuit shown in Figure 2 and the analytical phase angle of (3). Moreover, the previously designed LPF has to be modified to achieve a more advanced property. In this design, a 50 ohms impedance match in the input/output pads of the purposed LPF, this can increase the stop-band rejection.

3. FG-CPW LPF SIMULATION AND MEASUREMENT RESULTS

The obtained frequency responses of the improve LPF which has been tuned and optimized are plotted as shown in Figure 3. In the pass-band of the LPF, i.e., below the cut-off frequency of 75 GHz, the insertion loss is less than 1.0 dB, the return loss is higher than 15 dB. In the transition band, there is a sharp slop from 75 GHz to 95 GHz, whose attenuation is reduced from 3 dB to 25 dB. In the stopband range, we have observed a very wide rejection property up to 100 GHz, whose attention values are below 25 dB. Therefore, the improved LPF has a more excellently harmonic-suppressed and spurious-suppressed properties which are necessary to suppress the high-order harmonics and spurious response in the RF/Microwave high power amplifiers and mixers. The overall structure of the presented filter is symmetrical distribution about the central cross-junction section, and the photograph of the fabricated FG-CPW LPF is shown in Figure 4.
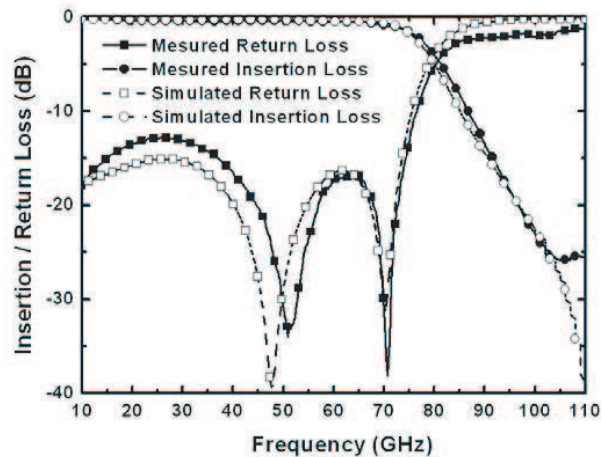
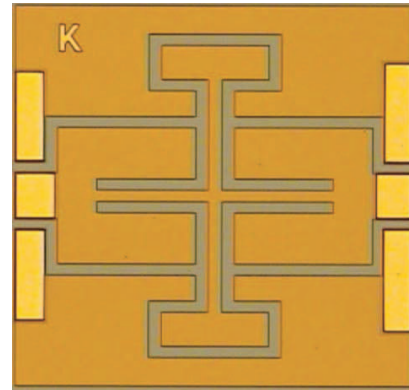Figure 3: The simulated and measured S-parameter of the FG-CPW LPF.

Figure 4: The photograph of the fabricated FG-CPW LPF.

4. CONCLUSION

In this paper, a novel and compact low-pass filter with harmonic-suppressed, sharp rejection slope and wide and deep stop-band rejection properties has been presented. The proposed FG-CPW LPF has shown a wider and deeper stop-band rejection, whose attenuation is below 25 dB more than 100 GHz wide frequency range. Besides, there is very sharp transition slope from 75 GHz to 95 GHz, whose attenuation is reduced to 25 dB. Therefore, the proposed LPF can be widely used to suppress the high-order harmonics and spurious responses which are usually generated in RF/Microwave nonlinear devices, such as high power amplifiers, mixers and oscillator.

REFERENCES

1. Smulders, P., "Exploring the 60 GHz band for local wireless multimedia access: Prospects and future directions," *IEEE Commun. Mag.*, Vol. 40, No. 1, 140–147, Jan. 2002.
2. Yang, F.-R., Y. Qian, and T. Itoh, "A novel uniplanar compact PBG structure for filter and mixer applications," *IEEE MTT-S Int. Dig.*, Vol. 3, 919–922, Jun. 1999.
3. Ahn, D., J.-S. Park, C.-S. Kim, J. Kim, Y. Qian, and T. Itoh, "A design of the low-pass filter using the novel microstrip defected ground structure," *IEEE Trans. Microw. Theory Tech.*, Vol. 49, No. 1, 86–93, Jan. 2001.
4. Chen, H.-J., T.-H. Huang, C.-S. Chang, L.-S. Chen, N.-F. Wang, Y. H. Wang, and M.-P. Houng, "A novel cross-shaped DGS applied to design ultra-wide stopband low-pass filters," *IEEE Microw. Wireless Compon. Lett.*, Vol. 16, No. 5, 252–254, May 2006.
5. Hettak, K. and G. Delisle, "A new miniature uniplanar lowpass filter using series resonators," *IEEE MTT-S.*, Vol. 3, No. 3, 1193–1196, Jun. 7–12, 1998.
6. Mao, S.-G. and M.-Y. Chen, "Propagation characteristics of finite-width conductor-backed coplanar waveguides with periodic electromagnetic bandgap cells," *IEEE Trans. Microw. Theory Tech.*, Vol. 50, No. 11, 2624–2628, Nov. 2002.

Introduction to the System-level Susceptibility Assessments for HEMP and HPEM

Congguang Mao[1], **Hui Zhou**[1], **Jiwei Fu**[2], **Beiyun Sun**[1], **and Haitao Yu**[2]

[1]Northwest Institute of Nuclear Technology, Xi'an 710024, China
[2]Systems Engineering Division of CALT, Beijing 100076, China

Abstract— The system-level susceptibility assessments for HEMP and HPEM are believed to be the most complex and hardest task to fulfill of the HPEM problems. The objective of this paper is to systematically expound the problem and clarify the relevant concepts. The core contents of the assessment activities are formulated utilizing the model based on the electromagnetic topology (EMT) and network graph theory. The necessary parameters, their characteristics and corresponding testing types are presented and discussed in details. We believe that it is very significant and helpful to its reduction and solution in the first step to well illuminate the complicated systematic problem.

1. INTRODUCTION

With the remarkable progress of the high-power electromagnetic (HPEM) technologies, the susceptibility problem of the electronic systems and the civilian infrastructures has abstracted increasing attentions [1, 2]. The term, HPEM, generally describes a set of transient electromagnetic environments (EME) where the peak electric and magnetic fields can be very intense, such as HEMP (high-altitude electro-magnetic pulse), HPM, UWB, and so on. It should be noticed that the terms, HEMP and HPEM, are also applied equally, where the different production mechanisms are considered, i.e., the former is produced by the nuclear explosion, but the latter is radiated by the special antenna and source. The intense electromagnetic energy can induce the permanent damage or temporary logic upset of the electrical systems.

Nowadays many scientists and international organizations have applied themselves to the study how to evaluate and assure the safety of the electronic systems in the rigorous EME. A special issue on HPEM and IEMI (intentional electromagnetic interferences) was published in IEEE transactions on EMC in August 2004. And a series of standards on HPEM and HEMP have been developed by the International Electrotechnical Commission (IEC) [3, 4]. The electromagnetic damage is evaluated by the probabilistic model in [5, 6], while only the collection of tests is recommended for the practical system assessments in [7]. Obviously no general idea has been achieved, and there are fewer open papers to systematically expound the problem of the system-level susceptibility assessments for HEMP and HPEM, which is believed to be the most complex and hardest task to fulfill of the HPEM problems.

We believe that it is very significant and helpful to its reduction and solution in the first step to well illuminate the complicated systematic problem. Firstly the concepts of susceptibility assessment are introduced. Then the core contents of the assessment activities are formulated based on the electromagnetic topology (EMT) and network graph theory. The necessary parameters, their characteristics and corresponding testing types are presented and discussed in details. The further study work needed to do is discussed finally.

2. CONCEPTS OF SUSCEPTIBILITY ASSESSMENTS

A representative electronic system impacted by the intense electromagnetic pulse (EMP) is shown in Fig. 1, which generalized the essential coupling mechanism. The core equipment and three subsystems are enclosed in a slotted shielded enclosure. The exterior conductors and antenna can pick the incident EM energy and lead it to the core equipment. The EM wave can also penetrate into the shield enclosure through the slots or apertures on the wall.

The system is called to be susceptible to HPEM if it can not function normally in the HPEM environments. Before analysis and tests, however, no one knows whether the system can withstand the severe interference. So the purpose of susceptibility assessment is to conduct the effect analysis and tests at the system-level, diagnose and analyze the component faults induced by the intense EM energy and then present the suggestions to harden the system. In this process, one must acquire two foundational quantities, the responses and interference threshold (IT). The responses are referred

as the EM stresses (EMS) on the critical components, i.e., currents, voltages, power, energy, etc. Both the conducted and radiated disturbances (in Fig. 1) can pose EMS on the subsystems and equipment. And IT is the maximum strength of the device withstanding the EMS. If the stress is greater than the strength, the device could be damaged or the function upset. The comparison and fusion of the two quantities can quantify the impact of HPEM on the system, based on which some effective measurements can be adopted to optimize the protective design of the systems.
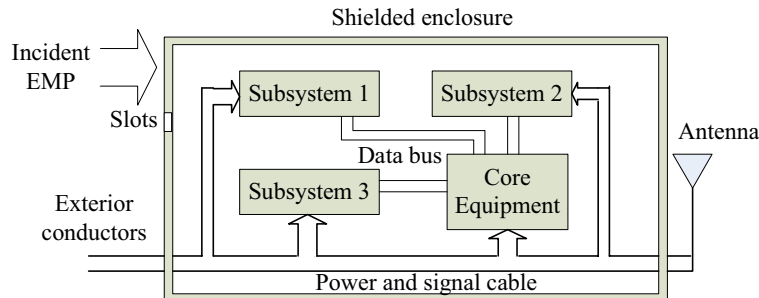
Figure 1: Diagram of electronic system impacted by EMP.

Unfortunately, this is not a task to readily fulfill because it depends on the attributes of both the EM environment and the electronic systems. They are following: 1) The wide band frequency spectrum of EMP often require the system attributes out of the operational band, which usually are unknown; 2) The high level of the electric and magnetic field can lead to the nonlinear effects on the components such as breakdown of the hardware and logic upset of the software, which is difficult to predict and analyze beforehand; 3) The characteristics of HPEM and systems usually vary largely compared with that of the continuous wave (CW), which leads to the great uncertainty in the analysis and testing; 4) An electrical system is a collection of many parts and components. Sometimes the number of elements, coupling paths and effects is so large that even makes the system too complex for the analysis or numerical simulation. 5) Because of the difficulties in the analysis, the tests play an important role in the assessments. However large-scale tests are usually of high cost, even which can not be afforded, such as aircraft, radar and C^{4}I systems.

These features determine that the assessment procedure involves several subprograms, i.e., the environment understanding, analysis and computational simulation, tests and fault diagnosis and proposals to optimize the protective design.

3. INTERACTION ILLUMINATION

In order to understand the complex interactions of EM energy with electronic systems, the concepts of the EM topology (EMT) are proposed [2, 5, 7, 8]. The main idea of EMT theory is to decompose the complex system into a set of volumes and surfaces, which are interconnected by the propagation paths of the radiated and conducted disturbances. The interactions in Fig. 1 can be illustrated by EMT in Fig. 2. The enclosure, system and subsystems are denoted by the different levels of
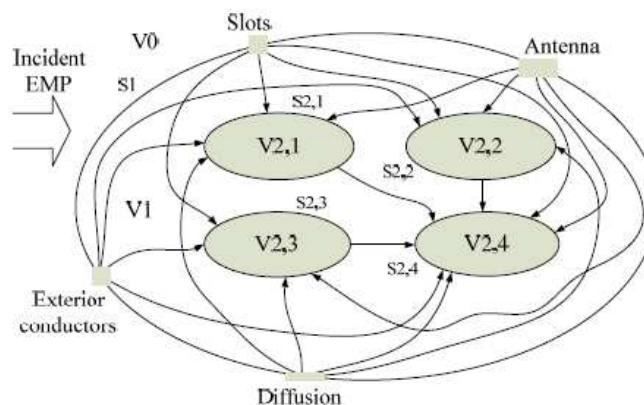

Figure 2: Electromagnetic topology of interaction.

volumes enclosed by corresponding surface. And the arrows stand for the propagating paths and directions of the EM interferences (see Fig. 2).

The interaction can further be abstracted as a network graph consisting of the nodes and the branches (see Fig. 3). The nodes (open, black and gray circle) denote respectively the spaces of the exterior and interior enclosure, volumes (V2,1–V2,4), conductors, antenna and protective devices, i.e., $v1$ is the outer space; $v2$, $v3$, $v4$ and $v5$ respectively the exterior conductors, antenna, slots and diffusion; $v6$, $v7$, $v8$ and $v9$ the three subsystems and core equipment. The branches connecting two nodes mean that there is disturbing EM energy propagating from the start to the end of the arrow. On the branches of $v1$–$v2$ and $v1$–$v3$ there are conducted protective devices such as filters, and the shields are placed on the branches of $v1$–$v4$ and $v1$–$v5$ to suppress the radiated disturbance.

The graph displays clearly the interaction sequence. According to the graph theory the network can be described by the adjacent matrix

$$G_f = [\text{SE}_{i,j}]_{N \times N} \tag{1}$$

where f, frequency, is the independent variable; N is the total number of nodes; the index i (or j) is the node code, where all the nodes are coded from the outer layer to the inner layer for the reference always from the exterior to the interior, and such code can simplify the matrix G_f; there are three types values of the elements $\text{SE}_{i,j}$, 0 shows no attenuation from the node to itself, ∞ stands for no branches between two node, and the elements of non-0 and non-∞ are the shielding effectiveness from node i to node j. This is obviously based on the linear-system theory. So G_f is the matrix style of the interactions sequence. The establishment of the matrix G_f is the process of decomposition of the system and acquirement of $\text{SE}_{i,j}$. The acquirement means are usually the low-power transient and CW tests or computations. The so-called "low-level" implies that no nonlinear effects occur during the testing.
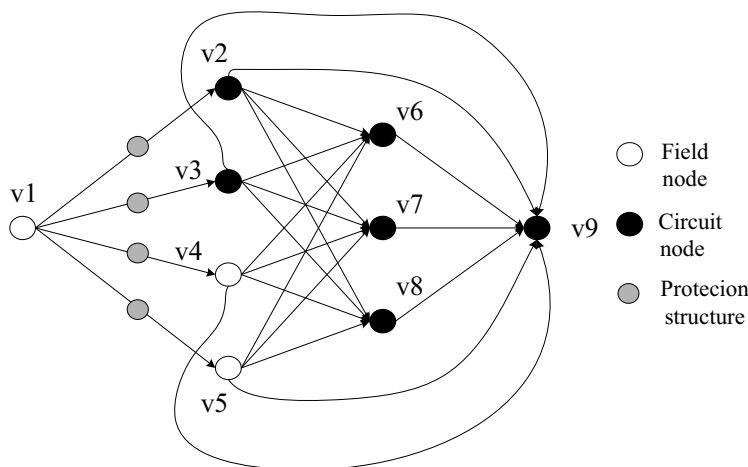

Figure 3: Network graph of EMT.

4. FORMULATION OF SUSCEPTIBILIT ASSESSEMENT ACTIVITIES

The EM stress (EMS) on the critical component is determined by subtracting the total shielding effectiveness (TSE) of the paths from the EME value, i.e.,

$$\text{EMS} = \text{EME} - \text{TSE} \quad (\text{in dB}) \tag{2}$$

The ambient field propagating by one path can be attenuated by several shields or filters. Thus

$$\text{TSE}(\omega) = \prod_{j=1}^{M} \text{SE}_{i,j}(\omega) \tag{3}$$

where M is the total number of SE in the path. In addition, the EM disturbances can reach one node by several paths, so EMS can further be expressed in frequency domain by (3) [1].

$$\text{EMS}(\omega) = \sum_{i=1}^{P} \left(\prod_{j=1}^{M} \text{SE}_{i,j}(\omega) \right) \cdot \text{TAF}(\omega) \tag{4}$$

where P is the total number of interference paths to one victim component.

Every gray and black in Fig. 3 stands for an interference threshold (IT). Once EMS and the IT are determined, the failure indication (FI) is calculated only by subtracting EMS with IT,

$$\text{FI} = \text{EMS} - \text{IT} \quad (\text{in dB}) \tag{5}$$

The positive FI indicates the system is susceptible, i.e., very probable to be destroyed or upset. Then the design requirements at one or more locations in the network must be modified in order to decrease the probability of failure. IT usually is acquired by means of the test of the pulsed current injection (PCI). The testing types to obtain the parameters in (1)–(5) are listed in Table 1.

Table 1: Parameters and test type in assessment.

Parameter	Test type	Comment
EMS	illumination	high-power
IT	PCI	ruinous
SE	CW, transient	low-level

For the reasons discussed earlier, some of the physical quantities of concern can not be acquired readily in practice, all the information available about the parameters in (5) should be collected and utilized, including the specifications, past experience, database and experts' judgments. It is the complete test during the normal operation in the real or simulated HPEM environments that can examine the survival ability of the real systems. The analysis can not take the place of the tests, but play the absolutely necessary assistant role in the whole assessment procedure.

What's more important is that (1)–(4) can play as a guidance to search the critical interference paths and susceptible components. Here the critical component is the core equipment $v9$. Hypothetically given G_f at $f = 10\,\text{kHz}$,

$$G_{f=10\,\text{kHz}} = \begin{bmatrix} 0 & 1 & 1.2 & 2 & 20 & \infty & \infty & \infty & \infty \\ \infty & 0 & \infty & \infty & \infty & 30 & 25 & 20 & 15 \\ \infty & \infty & 0 & \infty & \infty & 8 & 10 & 6 & 5 \\ \infty & \infty & \infty & 0 & \infty & 15 & 10 & 15 & 5 \\ \infty & \infty & \infty & \infty & 0 & 5 & 3 & 8 & 3.5 \\ \infty & \infty & \infty & \infty & \infty & 0 & \infty & \infty & 1 \\ \infty & \infty & \infty & \infty & \infty & \infty & 0 & \infty & 2 \\ \infty & \infty & \infty & \infty & \infty & \infty & \infty & 0 & 1.5 \\ \infty & \infty & \infty & \infty & \infty & \infty & \infty & \infty & 0 \end{bmatrix} \tag{6}$$

The poor paths less than 20 dB find by the node-searching algorithm [10] realized by the Matlab language. They are:

$$v_1 \rightarrow v_2 \rightarrow v_9, \ v_1 \rightarrow v_3 \rightarrow v_9, \ v_1 \rightarrow v_4 \rightarrow v_9,$$
$$v_1 \rightarrow v_3 \rightarrow v_6 \rightarrow v_9, \ v_1 \rightarrow v_3 \rightarrow v_7 \rightarrow v_9,$$
$$v_1 \rightarrow v_3 \rightarrow v_8 \rightarrow v_9, \ v_1 \rightarrow v_4 \rightarrow v_6 \rightarrow v_9,$$
$$v_1 \rightarrow v_4 \rightarrow v_7 \rightarrow v_9, \ v_1 \rightarrow v_4 \rightarrow v_8 \rightarrow v_9.$$

These are the main paths conducting ambient EM energy to the core equipment, and the results are the same as that in [9]. However, this algorithm uses less concepts of graph theory than the Dijakstra [1] and MPS [9] algorithm.

Moreover, if some subsystems in the EMT need to be analyzed further or may be ignored, the matrix G_f can be extended or reduced as the same procedure as described earlier. This means that once the critical subsystems and paths are located, most of efforts are concentrated on them, thus the scale and cost of tests and analysis can be reduced.

5. DISCUSSION

The concepts of system-level susceptibility assessments for HEMP and HPEM are introduced in this paper. The model based on the EMT and network graph theory is utilized to formulate the assessment activities. The necessary parameters, their characteristics and relationship are presented and discussed in details. In the future work, the test standards and analysis or simulation tools for the susceptible assessment must be developed. The attributes of the damage or upset thresholds of various electrical devices, i.e., the nonlinear effects, should be studied extensively. The component failures need to be explored to quantitatively describe the impact of local faults on the system functions. It can be seen that all the means mentioned above should be integrated properly for the rational and reliable assessments, and any one means itself is not competent for the resolution of the complex and systematic problem.

REFERENCES

1. Vetri, J. L. and G. I. Costache, "An electromagnetic interaction modeling advisor," *IEEE Trans. on Electromagn. Compat.*, Vol. 33, 241–251, Aug. 1991.

2. Radasky, W. A., C. E. Baum, and M. W. Wik, "Introduction to the special issue on high-power electromagnetics (HPEM) and intentional electromagnetic interference (IEMI)," *IEEE Trans. on Electromagn. Compat.*, Vol. 46, 314–321, Aug. 2004.

3. Wik, M. W. and W. A. Radasky, "Development of high-power electromagnetic (HPEM) standards," *IEEE Trans. on Electromagn. Compat.*, Vol. 46, 439–445, Aug. 2004.

4. Radasky, W. A., "2007 status of the development of high-power electromagnetic (HPEM) standards in the IEC," *International Symposium on Electromagnetic Compatibility*, 6–9, Oct. 2007. http://www.iec.ch

5. Cabayan, H. S., "Phenomenological investigations in high-altitude EMP (HEMP)," July 1984.

6. Tan, Z., S. Liu, Y. Lin, and R. Zhang, "Study of evaluation methods on electromagnetic damage of electronic systems," *Chinese Journal of Radio Science*, Vol. 22, No. 5, 821–823, Oct. 2007.

7. Baum, C. E., "From the electromagnetic pulse to high-power electromagneics," *Proceedings of the IEEE*, Vol. 80. No. 6, 789–817, June 1992.

8. Lee, K. S. H., *EMP Interaction: Principles, Techniques, and Reference DATA (A Compleat Concatenation of Technology from the EMP Interaction Notes)*, Dec. 1980

9. Fu, J., C. Hou, and L. Dou, "An analytical method of the electromagnetic topological diagram," *Proceeding of 2004 Academic Forum for Chinese Doctoral Candidates*, 247–251, 2004.

10. Jin, X. and Y. Hong, *Assessment Methods of System Reliability*, National Defense Industry Press, Beijing, June 2005.

A Viewpoint of Time Variant Dielectric Effect in Vital Sign Detection Using Microwave Radar

Jingzhou Luo[1] and Kemin Sheng[2]

[1]Information Science and Engineering College
Zhejiang University, Hangzhou, Zhejiang, China
[2]Southwest Jiaotong University, Chengdu, China

Abstract— Through theoretical analysis and data examinations for a continuous microwave radar system, it is pointed out that there are some inconsistencies when using Doppler radar theory to explain the mechanism of vital sign detection. Meanwhile a time variant dielectric hypothesis viewpoint based on bioelectrical effects is presented, and a simple system model is established. Comparing to the traditional Doppler radar theory explanation, this viewpoint is more reasonable.

1. INTRODUCTION

As early as 1970's, Doppler radar system was applied to detect human body's vital sign like respiration and heartbeat. Up to now, the main principle of microwave radar for vital sign detection is still based on Doppler theory [1]. It says that when a human body is exposed under the incidence of a microwave, the reflected signal will be phase modulated (PM) due to movements caused by respiration and heartbeat, and the frequency or phase of the incident wave can be changed. So by appropriate demodulation techniques, one can obtain the vital sign signal from the change of reflected wave.

Most of us take it for granted that Doppler radar theory is the foundation of vital sign detection. However, with further researches we find there are some inconsistencies between the theory and measurements. So we try to write this paper to show our considerations and give a hypothesis. Firstly, after a short review of Doppler radar theory we analyze two significant contradictions between this theory and measurements. Secondly, we present a new viewpoint or a hypothesis based on time variant dielectric effects in human body and in virtue of electromagnetic scattering theory. Finally, we give a comparison for the viewpoint to traditional Doppler radar theory explanation to show that it can be more reasonable to explain the mechanism.

2. DISCUSSIONS FOR DOPPLER RADAR THEORY OF VITAL SIGN DETECTION

2.1. A Short Review of Vital Sign Detection Based on Doppler Radar Theory

The classic explanation for this topic is Doppler theory [1]. According to this theory, the incident wave that is reflected off human's chest with a periodical varying displacement would result in a narrowband PM reflected wave:

$$S(t) = \cos(\omega_0 + \phi(t)) \tag{1}$$

where ω_0 is the angular frequency and the phase shift is:

$$\phi(t) = 4\pi \cdot x(t)/\lambda, \tag{2}$$

where $x(t)$ is the displacement of the target to be measured. In receiver part of the radar $\phi(t)$ can often be obtained by coherent detection. The principle is shown in Figure 1. To obtain $\phi(t)$, one should adjust the phase of phase-shifter to make $\theta = 90°$. In a small phase shift condition, we have $\sin(\phi(t)) \sim \phi(t)$. So by low pass filtering and baseband amplifying, $x(t)$ will be displayed on the oscillograph. When the phase shift is large enough, this linear approximation is not valid. So this case has led researchers to further deeper study [2, 3]. However, their works are still based on a cosine function baseband signal model that is derived from classical Doppler radar theory. Although they can improve the detection accuracy to some extent, the following doubt points we put forward still keep remained.
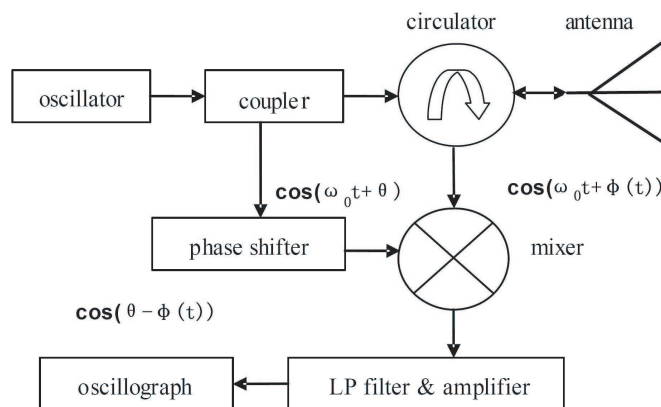
Figure 1: The simple diagram of continuous wave Doppler radar.

2.2. The Baseband Signal Difference between Theory and Experiments

The first contradiction is that the baseband signal detected in lab experiment is different from the theory. Referring to Figure 1, to obtain $\phi(t)$ from $\cos(\phi(t))$, two conditions must be satisfied: (a) Coherent local oscillation signal should be orthogonal to the received RF signal; (b) The reflected signal should be modulated with small phase shift to the carrier wave, that is:

$$\max \phi(t) \ll \pi/6 \tag{3}$$

If these two conditions cannot be satisfied, ordinarily an inverse cosine operation to the output signal will be needed. If condition (b) is satisfied but condition (a) is not, we can obtain the approximate results by differentiating to the output signal. Otherwise the demodulation baseband signal is just $\cos(\theta - \phi(t))$, not the target movement signal $\phi(t)$ (so is $x(t)$). But the fact is contrary: almost all the experiment results show that one can directly obtain the heartbeat and/or respiration signal $x(t)$ waveform after demodulation. Actually, most of the experiments didn't meet the two conditions described above (we have confirmed this). Here we cite a typical test results in Figure 2 to explain the problem [4]. We can clearly verify this observation from the figures.
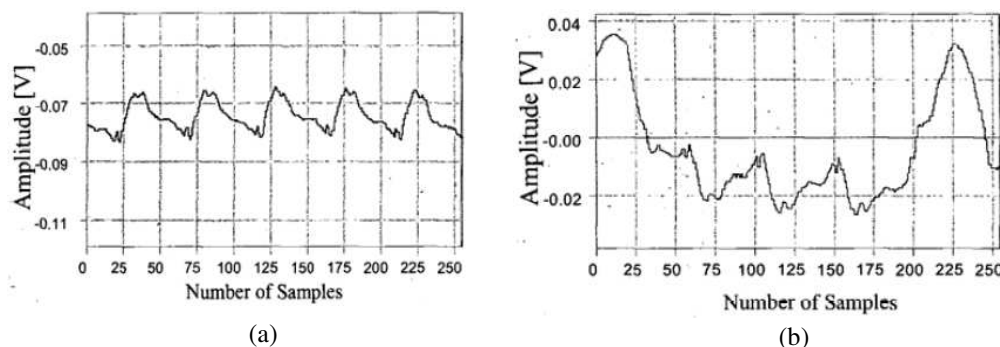

Figure 2: The test results [4], $f = 10.353\,\text{GHz}$, 50 samples per second. (a) The heartbeat signal; (b) The combination signal of the respiration and heartbeat.

2.3. The Amplitude Difference between Heartbeat and Respiration Signals

The second important problem is that Doppler radar theory cannot explain the amplitude difference between heartbeat and respiration signals. We can see also from Figure 2 and test results mentioned in other papers that they are basically in a same decimal order [5]. But according to Doppler theory, the difference should be much bigger, as we prove as follows.

Firstly, we consider this problem from the surface movement made by heartbeat and respiration. According to Equation (2), the maximum amplitude of modulated signal should be determined by maximum target's displacement $x(t)$. Directly comparing the displacement of respiration and heartbeat, we can find the displacement of breast and abdomen caused by respiration is much

larger than that by heartbeat. For a healthy adult man, the displacement caused by respiration can reach up to 10 millimeters even more. However, for displacement caused by heartbeat, we can hardly observe it even the testee hold the breath. Therefore, in the view point of Doppler theory, respiration signal should be much larger than that of heartbeat signal.

Secondly, if considering that the movement displacement caused by heartbeat inside chest is larger than that on the surface, we can analyze the RF loss in human body to explain the problem. Here we only present the result because of the paper length limitation: a detailed calculation by ADS simulation shows that heartbeat signal is about 96 dB smaller than the respiration signal at 10 GHz [6].

Therefore, in view of Doppler radar theory, no matter how we analyze the problem from either aspect, respiration signal should be much larger than heartbeat signal. However, the fact isn't like this as we pointed out.

3. THE MECHANISM OF VITAL SIGN DETECTION BASED ON TIME VARIANT DIELECTRIC EFFECTS IN HUMAN BODY

As is known, bioelectricity is an electrical phenomenon of organism. The sinoatrial node controls heartbeat of a healthy people. The bioelectrical pulses produced by sinoatrial node pass through special transmission tissues, atria, and atrioventricular node (the node between atria and ventricle), finally reach the ventricle, and result in contraction of the atria and ventricle. The regular repetitions of this process become the heartbeat and rhythm of heart.

There is another character of bioelectricity that the vital information it carries can be transmitted to outer skin through electric tissues and body fluid around the heart and nerve cells. So testing electrode put on the skin can detect the change of the voltage in human's body. Because of this character, the biological electric fields caused by every action in human's body exist on human's skin. According to electromagnetic theory, the body can be regarded as a time varying dielectric coefficient ε and conductivity σ, which can carry the information of biological electric field:

$$\varepsilon = \varepsilon_0 \varepsilon_r f(t), \quad \text{and} \quad \sigma = \sigma_0 f(t). \tag{4}$$

where ε_0 is the permittivity in vacuum, ε_r and σ_0 is the normally defined human body's dielectric constant and conductivity, respectively; $f(t)$ is a hypothetical function that represents time variant dielectric effects in human body. Comparing with the variation of incident field, $f(t)$ varies much slowly. When a human body is exposed under the incidence of microwave radar shown in Figure 1, the Maxwell equation has the form by (4)

$$\nabla \times H = j\omega\varepsilon(t)E + \sigma(t)E = j\omega\varepsilon_e f(t)E \tag{5}$$

where ε_e is the ordinary complex dielectric constant,

$$\varepsilon_e = \varepsilon_r \varepsilon_0 - j\frac{\sigma_0}{\omega}. \tag{6}$$

The field in (5) is a total field, involving incident field $\mathbf{E}^{inc}$ and scattered field $\mathbf{E}^s$. We can use the concept of equivalent polarization current and integral equation method to solve the scattered field [7]. Therefore we can derive

$$\mathbf{E}^s = \varepsilon_e f(t) e^{j\phi(t)} \left(\delta\omega^2 \mu_0 \varepsilon_0 \iint_S \mathbf{E}^{inc} G(r, r') ds' + \iint_S (n \cdot \mathbf{E}^{inc}) \nabla' G(r, r') ds' \right) / \varepsilon_0 \tag{7}$$

where δ is the thickness of human body's surface layer skin; G is Green function in free space; the integrals are performed both on human's surface with a unit normal vector n. The Equation (7) is derived under following approximations: The field in human body attenuates rapidly, $|\varepsilon_e| \gg \varepsilon_0$, and $|\mathbf{E}^{inc}| \gg |\mathbf{E}^s|$.

When the incident field is scattered by human body, referring to (7), the scattered field will have a product form of $f(t)$ and incident field. This is equivalent to that the incident carrier wave is modulated in amplitude by the bioelectrical signal. Therefore, the scattered field has the character not only of PM wave with Doppler phase shift, but also of amplitude modulated (AM) wave. Namely, the scattered field is an AM-PM signal. So in a radar receiver, we can use coherent demodulation method to get the function $f(t)\cos(\theta - \phi(t))$, and adjust θ only for getting maximum

amplitude, but not for restoring the signal information itself. This is an essential difference from the Doppler theory mentioned in Section 2.1.

The model in Figure 3 shows above work principle. The main difference of Figure 3 and Figure 1 is that the former includes a time variant dielectric function $f(t)$, so the demodulated result is completely different.
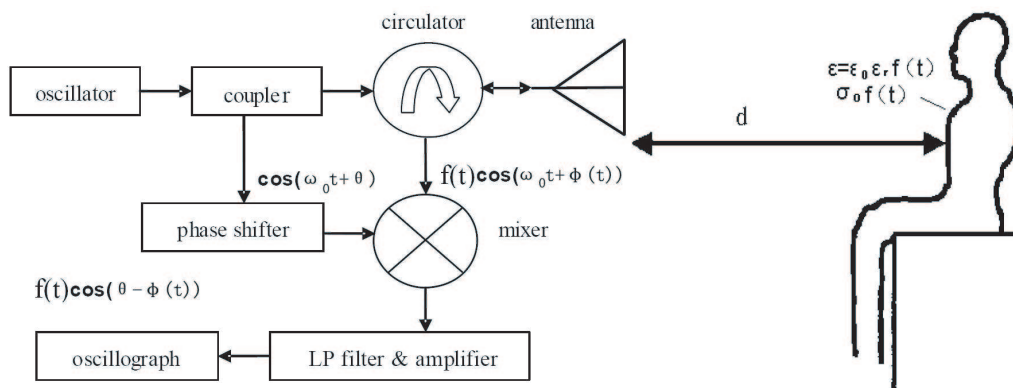

Figure 3: Time variant dielectric effect models of vital sign detection using microwave radar.

Referring to Figure 3, the model has several characters. Firstly, the scattered signal can be equivalent to an AM-PM wave, which has less dependency on the movement of human body. After coherent demodulation to the AM wave we can get the baseband signal both of respiration and heartbeat. Secondly, from Equation (7), since the solution of the scattered field can be carried out only on the surface of human body, the microwave loss problem is avoided. Finally, since AM signal satisfies the principle of linear superposition, this model can process multi-signal detection problem, which is difficult to solve in Doppler radar theory.

It is worthy to indicate that in the past, people used to consider organism as an inhomogeneous loss medium only, and the dielectric characteristics of all biological organization do not change along with organism's status. In fact, as human body is a complex organism, its metabolism process is fulfilled by a series of complex biochemical response. Yan Liping et al., who are in Sichuan University, China, have made a research on this topic and published their paper [8]. They made an experiment to show the behavior of time-variable dielectric coefficients and conductivity. According to their research report, the experiment takes about 10 seconds from time of beginning to stable status, and the relative variation rate of complex dielectric coefficients and conductivity in human body can be reached up to 37%. Detailed information of their contribution and our further work will be described in the future.

4. CONCLUSION

This paper analyzed the inconsistencies between Doppler radartheory for vital sign detection and known measurements. It also provides a new mechanism and an AM-PM wave model for microwave radar vital sign detection based on time variant dielectric effects, which is essentially caused by bioelectrical phenomena of human body in our opinion. Comparing with the traditional Doppler radar model, the advantages of this model are:

1) It avoids the inconsistencies between theoretical analysis and experiment results.
2) Because a variety of time variant dielectric effects is on human's skin, so whether we detect any vital sign, we just need to consider the reflection of the incident wave on human's skin surface, and the detection will have less dependency on movement. Besides, as vital sign signals satisfy the principle of linear superposition, it makes easy to detect the heartbeat and respiration at same time.
3) The simulation model can be more similar to the reality, and can be constructed by more rigorous electromagnetic theoretical analysis.

ACKNOWLEDGMENT

REFERENCES

1. Boric-Lubecke, O., et al., "Wireless IC Doppler radars for sensing of heart and respiration activity," *6th International Conference on Telecommunications in Modern Satellite, Cable and Broadcasting Service*, Vol. 1, 337–344, Oct. 2003.

2. Li, C., et al., "Experiment and spectral analysis of a low-power Ka-band heartbeat detector measuring from four sides of a human body," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 54, No. 12, Dec. 2006.

3. Li, C. and J. Lin, "Non-contact measurement of periodic movements by a 22–40 GHz radar sensor using nonlinear phase modulation," *IEEE MTT-S International Microwave Symposium Digest*, 579–582, Jun. 2007.

4. Boric Lubecke, O., et al., "10 GHz Doppler radar sensing of respiration and heart movement," Bell Laboratories, Lucent Technologies.

5. Chen, K.-M., et al., "Microwave life-detection systems for searching earthquake rubble or behind barrier," *IEEE Transactions on Biomedical Engineering*, Vol. 27, No. 1, Jan. 2000.

6. Luo, J., et al., "Discussion on the mechanism of vital sign detection with CW radars," *Telecommunication Engineering*, (A Chinese Monthly Journal), Vol. 48, No. 7, Jul. 2008.

7. Liu, W., et al., "Characteristics of a half-wave length dipole on a dielectric slab," *Proc. of 3rd International Symposium on Antenna and EM Theory*, 136–139, Sept. 1993.

8. Yan, L., et al., "The non-linear characteristics of equivalent complex dielectric coefficient in biochemical reaction process," *The Proc. of 2nd International Electromagnetic Radiation and Health Research Symposium*, 244-246, China Institute of Biomedical Engineering, Beijing, 2000.

A Compact 5.8 GHz Rectifying Circuit Design and Experiments

Qijuan He, **Kama Huang**, and **Changjun Liu**

School of Electronics and Information Engineering, Sichuan University, Chengdu 610064, China

Abstract— A compact rectenna circuit at 5.8 GHz is presented. A matching inductor and two series diodes are introduced to the rectifying circuit so as to achieve well impedance matching. The circuit simulation and layout electromagnetic (EM) simulation are compared hereby to confirm this design. The layout EM simulation and circuit simulation are close to each other. There is a RF-DC conversion efficiency difference of about 3%–5% between them. The experiment DC output voltage of the rectifying circuit with this novel design achieves 9.80 V at a DC load 2000 Ohm, when the input RF power is at 20 dBm. The highest RF-DC conversion efficiency is 60%. The capacitors used in the rectifying circuit are also studied, and by adjusting their values, the RF-DC conversion efficiency at desired frequencies could be optimized.

1. INTRODUCTION

A rectenna is a RF power receiver that converts the received power into DC power which is consumed by the DC load components, e.g. an active RFID. Rectenna composed of receiving antennas and rectification circuits, is used to transmit microwave power without transmission lines. Rectenna at 5.8 GHz is extensively studied [1, 2], and rectenna array is also under research [3]. Differently from the single diode, double voltage diode rectifying circuit and the μ-power rectifier circuit [4], this paper present a novel rectifier operation at 5.8 GHz, shown as Fig. 1, combining the double voltage rectifier circuit and single diode circuit, and also being different from the μ-power rectifier circuit.
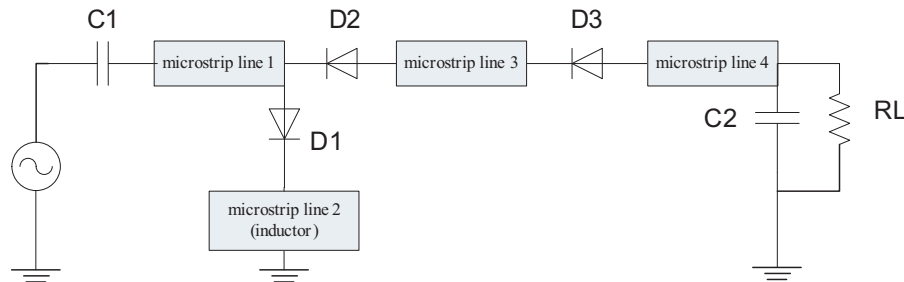

Figure 1: Proposed rectifying circuit scheme.

2. RECTIFYING CIRCUIT DESIGN

In our experiment, the Schottky diode used is HSMS8202, and its parameters are $R_s = 4$ Ohm, $V_{br} = 4$ V, $I_s = 9e - 8$ A, $n = 1.08$, $C_{j0} = 0.23$ pF. The Schottky diode model is studied and experimented in [5]. According to that diode nonlinear model, the input impedance of the diode HSMS8202 is not a pure real resistance, but has a capacitive reactance. So the matching circuit series connected to the diode D1 is expected to conjugate matching the imaginary part of the diode. The voltage of the rectifier of D1 is:

$$V_d = \begin{cases} V_{d0} + V_{d1}\cos(\omega t + \theta) \\ V_f \end{cases} \tag{1}$$

where, V_f is the turn on voltage. When the diode D1 is turned on, V_f is reverse bias the two diodes D2 and D3. Compared with the double voltage circuit, V_f is reverse bias two diodes. When D1 is turned off, the voltage V_d falls on the two diodes D2 and D3, and the turn on angle of the two diodes decrease compared with one diode, the power loss due to the R_s of the diodes decrease. In the double voltage circuit the power consumed by R_s is $(V_{d0} - V_{out})^2/R_s$, while $(V_{d0} - V_{out})^2/2R_s$ is totally consumed by the two diodes D2 and D3 in this design. $V_{out} = \frac{V_{d0}}{\frac{R_s}{R_L} + 1}$ is the output DC voltage. Parallel of D2 and D3 also make the conjunction capacitance to half of the single diode.

Before and after each diode microstrip line is added to make the circuit match to the source and the load impedance. Here, the source impedance remains 50 Ohm, but the load resistance varies from 100 Ohm to 2000 Ohm to achieve best RF-DC conversion efficiency. The capacitance C1 and C2 are chosen both 6.8 pF initially for the scheme and lay out simulation and in the measurement those capacitors are adjusted for best RF-DC performance. The layout scheme is shown in Fig. 2. The microstrip line parameters are as follows: The length of line1 is 4 mm and the width is 2.7 mm and 1 mm, respectively. The line2 is 8.04 mm long and 1.5 mm wide, the line3 is 1.76 mm long and 0.2 mm wide, and the last is 12 mm long and 1.2 mm wide.
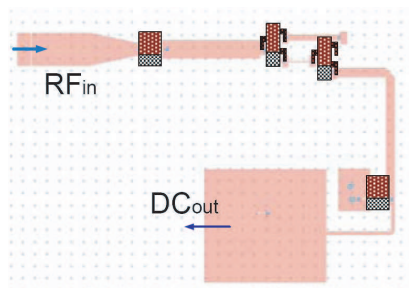
Figure 2: Layout of proposed rectifying circuit.

3. RECTENNA SIMULATION AND MEASUREMENT

The schematic simulation and the layout simulation both confirm this design, and the two simulations in ADS agree well with each other. The RF-DC conversion efficiency of layout simulation (EM simulation) is about 3%–5% lower than the schematic simulation as expected. The simulated and measured DC output voltage and RF-DC efficiency are shown in Fig. 3 and Fig. 4. Compared with simulated results, measured DC output voltage is better when RL becomes larger than 1000 Ohm, but worse when RL between 500 Ohm and 1000 Ohm. As to the RF-DC conversion efficiency at fixed DC load, it is more sensitive to C1 than C2 during our experiments. The optimal value of C1 and C2 here is 2.2 pF and 3.9 pF. The RF-DC efficiency increases by 3%–4% shown in Fig. 4. The RF-DC efficiency at large load resistance is better than simulation, which is properly because the radiation loss isn't calculated in the simulation. The measured efficiency is the ratio of the DC load power to the RF input power, the radiation loss and reflection loss aren't taken out from the RF input power.

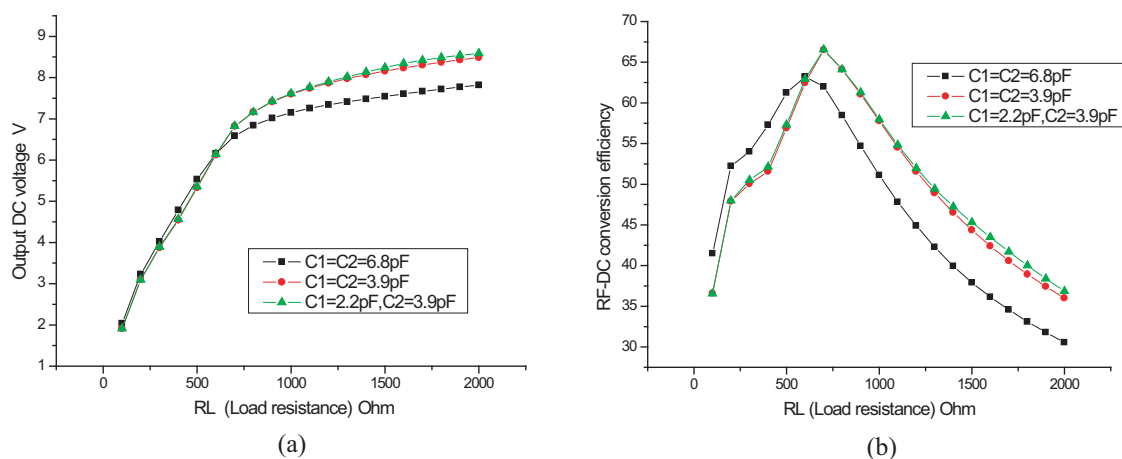(a)   (b)

Figure 3: (a) Simulated DC voltage. (b) Simulated RF-DC efficiency at 20 dBm input power.

The best RF-DC conversion efficiency of desired frequency could also be adjusted by change the value of C1 here. Fig. 5 shows RF-DC conversion efficiencies of several frequencies near the desired the frequency 5.8 GHz. In the simulation, schematic simulation and layout simulation under the three cases of $C1 = C2 = 6.8$ pF, $C1 = C2 = 3.9$ pF, $C1 = 2.2$ pF, $C2 = 3.9$ pF, are studied, the efficiency curves of which are almost the same as Fig. 5. In our experiment shown in Fig. 6,

however, the RF-DC efficiency of frequency 5.59 GHz is better than the desired frequency 5.8 GHz when RL lower than about 1200 Ohm. Due to our knowledge of the RC charge and recharge, one can make 5.8 GHz have the best perform but the corresponding C1 and C2 may not be found in the existed SMD capacitors.
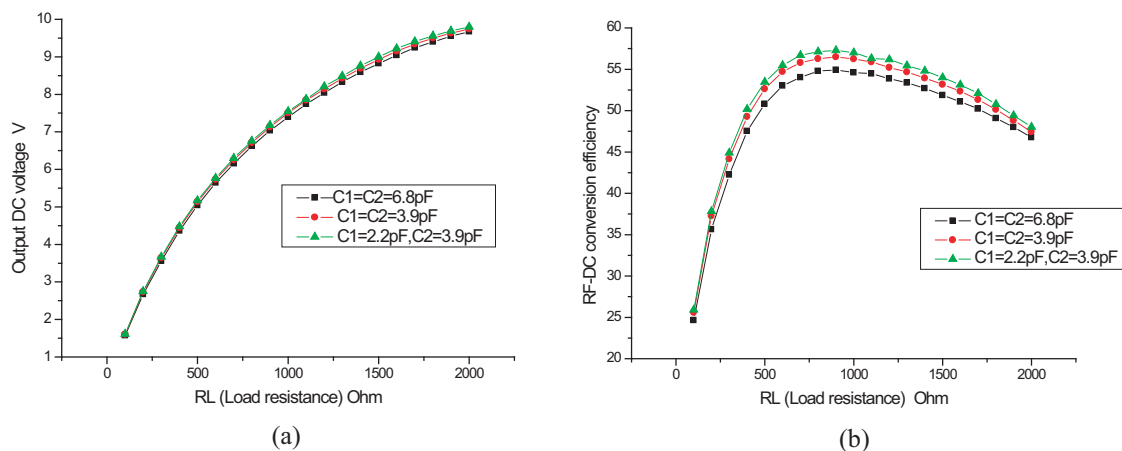

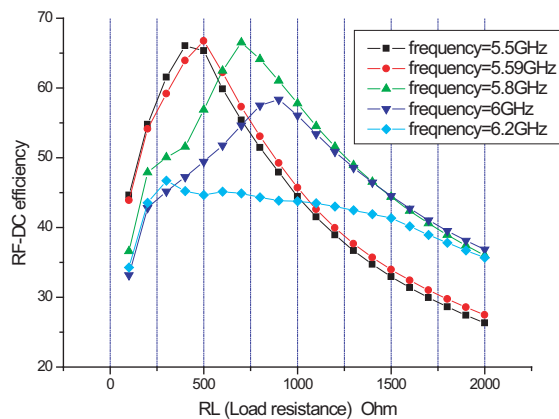Figure 4: (a) Measured DC voltage. (b) Measured RF-DC efficiency at 20 dBm input power.

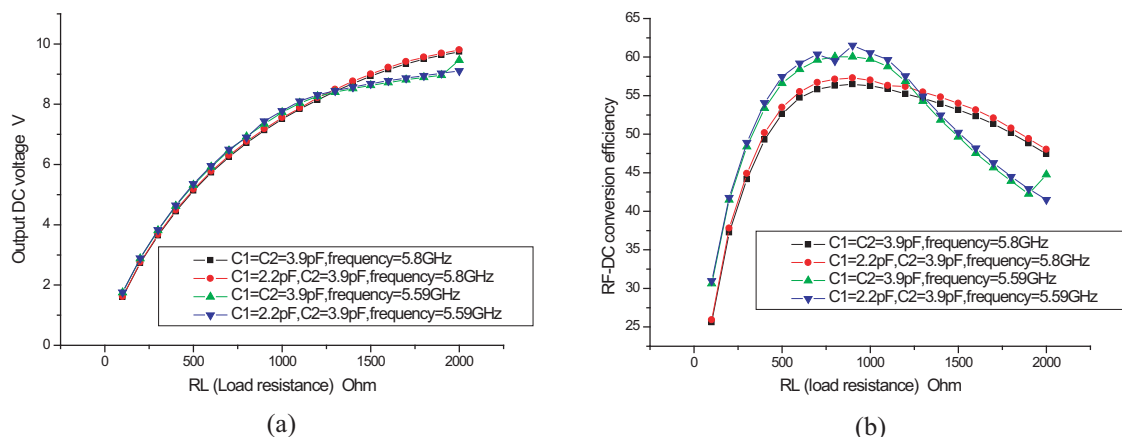Figure 5: Simulated varied frequency vs. varied DC loads when input power is 20 dBm.

Figure 6: (a) Measured DC voltage comparison. (b) Measured RF-DC efficiency comparison when input power is 20 dBm.

4. CONCLUSIONS

This paper proposes a novel rectifying circuit design and studies the capacitors connected in the circuit with its performance effecting on the RF-DC conversion efficiency and the frequency desired. The schematic lay out simulation and experiment are confirm this design. The value of capacitors are optimized in the experiments and results of the RF-DC efficiencies are compared.

REFERENCES

1. Chin, C. H., Q. Xue, and C. H. Chan, "Design of a 5.8-GHz rectenna incorporating a new patch antenna," *IEEE Antennas and Wireless Propagation Letters*, Vol. 4, 175–178, 2005.
2. Heikkinen, J. and M. Kivikoski, "Low-profile circularly polarized rectifying antenna for wireless power transmission at 5.8 GHz," *IEEE Microwave and Wireless Component Letters*, Vol. 14, No. 4, 162–164, 2004.
3. Ren, Y. J. and K. Chang, "New 5.8-GHz circularly polarized retrodirective rectenna arrays for wireless power transmission," *IEEE Trans. Circuits and Systems Theory Tech.*, Vol. 54, No. 7, 2970–2976, 2006.
4. Cury, J. P. and N. Joeh, "A model for μ-power rectifier analysis and design," *IEEE Trans. Circuits and Systems Theory Tech.*, Vol. 52, No. 12, 2771–2779, 2006.
5. Yoo, T. W. and K. Chang, "Theoretical and experimental development of 10 and 35 GHz rectennas," *IEEE Trans. Circuits and Systems Theory Tech.*, Vol. 40, No. 6, 1259–1266, 1992.

Parallelisation of Implicit Time Domain Methods: Progress with ADI-FDTD

Timothy D. Drysdale and Tomasz P. Stefanski

University of Glasgow, Glasgow G12 8LT, United Kingdom

Abstract— We provide an overview of our work to date on the parallelization of implicit time domain methods, in particular the alternating direction implicit finite-difference time-domain method (ADI-FDTD). First we describe a domain decomposition scheme for parallel ADI-FDTD in three dimensions that is suitable for implementation on widely available high performance computer architectures such as symmetric multiprocessors (SMP) and distributed memory computer clusters (DMCC). We present a selection of benchmark results for parallel solutions for domains of up to 8 billion mesh cells, and compare against standard parallel FDTD. The results indicate that a useful speed-up can be obtained with ADI-FDTD for large, highly oversampled meshes. We also present a formulation and benchmark results for parallel ADI-Body Of Revolution-FDTD. Our demonstration of parallel speed up represents an important step forward for the application of implicit time domain solvers. We expect that our parallelisation approach can be adopted for related implicit FDTD methods.

1. INTRODUCTION

The parallelisation of implicit time domain solvers is essential for applications requiring large, highly over-sampled meshes that exceed the memory limitations of standalone workstations [1]. Prime application examples include large arrays of small antennas, and on-chip interconnects for state of the art integrated circuits (IC) [2, 3]. The need for highly over sampled meshes in these applications can be illustrated by example. Current on-chip IC interconnect widths can be as low as 22 nm or less, yet cover an area $> 100\,(\text{mm})^2$ and transmit signals with fundamental frequencies in the range DC–10 GHz. Typically, transient effects involving high frequencies are of the most interest, but even this can still require meshes with $\Delta x < (\lambda_0/10^6)$, i.e., four to five orders of magnitude smaller than for typical FDTD meshes (neglecting considerations relating to material properties). It is not practical to use explicit FDTD solvers for such meshes, even in parallel, because the Courant-Friedrichs-Lewy (CFL) stability criteria enforces a commensurate reduction in the time step size for numerical reasons. While implicit FDTD solvers, such as ADI-FDTD, offer freedom from the CFL stability criteria, it has been presumed that parallel implementations on all but the most specialised architectures [4] would be of little benefit due to the high communication overhead. However, we have been able to show that this is not the case [5].

In this Paper we present an overview of our work to date on the parallelisation of implicit time domain methods, including parallel ADI-FDTD [1, 5] and parallel ADI-BOR-FDTD [6] on both symmetric multiprocessor (SMP) and distributed memory computer clusters (DMCC). In Section 2 we describe our domain decomposition scheme. In Section 3 we present benchmarking results for domains of up to 8 billion mesh cells. In Section 4 we present a parallel ADI-BOR-FDTD formulation and results. We present our concluding remarks in Section 5.

2. PARALLELISATION APPROACH

The ADI-FDTD iteration comprises two substeps which each require solving a tridiagonal matrix system of equations over the three orthogonal Cartesian planes of the domain. For the sake of brevity we do not include the equations here, because they can be found elsewhere [1, 5]. We do not parallelise the tridiagonal matrix solver itself, because each 2D matrix must have more than 4,000 (40,000) elements per direction before this becomes efficient for SMP (DMCC) machines. This requires 3D domains that are at or beyond current memory limits for state of the art machines. Instead, we decompose the domain in two directions as shown in Figure 1 and solve the resulting multiple, smaller, tridiagonal matrix systems in parallel.

The highlighted region in Figure 1(a) (local sub-domain) is employed for all explicit updates and the implicit updates of electric field components along x and y directions. These updates are based solely on the local sub-domain data so they do not require data stored inside other sub-domains (except for fields on the surface of their nearest neighbours, as for standard FDTD). However, for updates along the z-direction, the sub-domains take the form of the highlighted region in

Figure 1(b), requiring data from within local sub-domains to be communicated between processors. Although careful ordering of the communications ensures that data handling is minimized, this forms the rate-limiting step of the procedure. The overhead is proportional to the volume of the sub-domains.
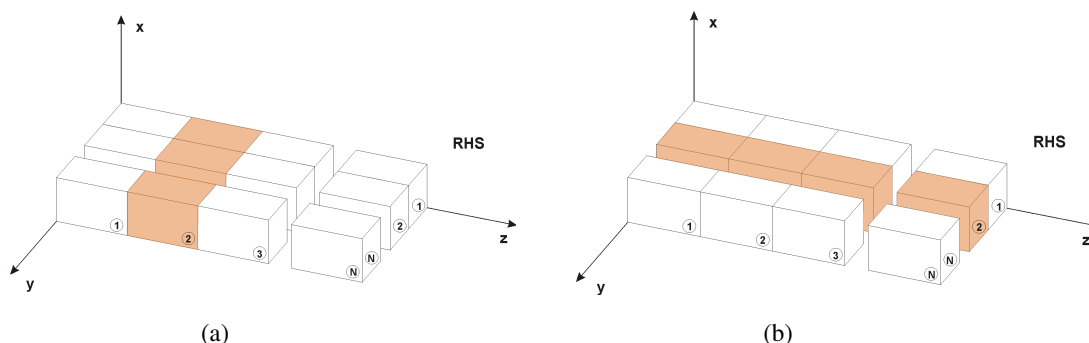

(a)   (b)

Figure 1: Domain decomposition scheme for parallelized ADI-FDTD. The sub-domain is highlighted for (a) explicit updates and implicit updates in x, y directions (b) implicit updates in z direction. [© IEEE]

3. PERFORMANCE BENCHMARKING

The code was developed in the C programming language using a message passing interface (MPI) library. Benchmark simulations were run on an SMP machine (IBM p690, Linux, Power 4+ 1.9 GHz 32 processors, 32 GB of shared RAM) and a DMCC machine (36 nodes with two dual-core Opteron 2222 3 GHz processors, 32 GB of RAM per node, DDR Infiniband interconnects). The computational domain was bounded either by the Mur 1st order ABC or 10 cells of the CPML boundary condition. Our investigations of the ADI-FDTD method resulted in improvements to the standard implementations of the Mur 1st order ABC [7] and CPML boundary condition [8]. Both of our improved boundary conditions treated the intermediate variables as non-physical variables, so that they are consistent with the ADI-FDTD scheme, although we do not repeat the details here for the sake of brevity. In order to aid the comparison of our benchmark results against other serial ADI-FDTD codes, we used the standard boundary conditions referenced in [7, 8].
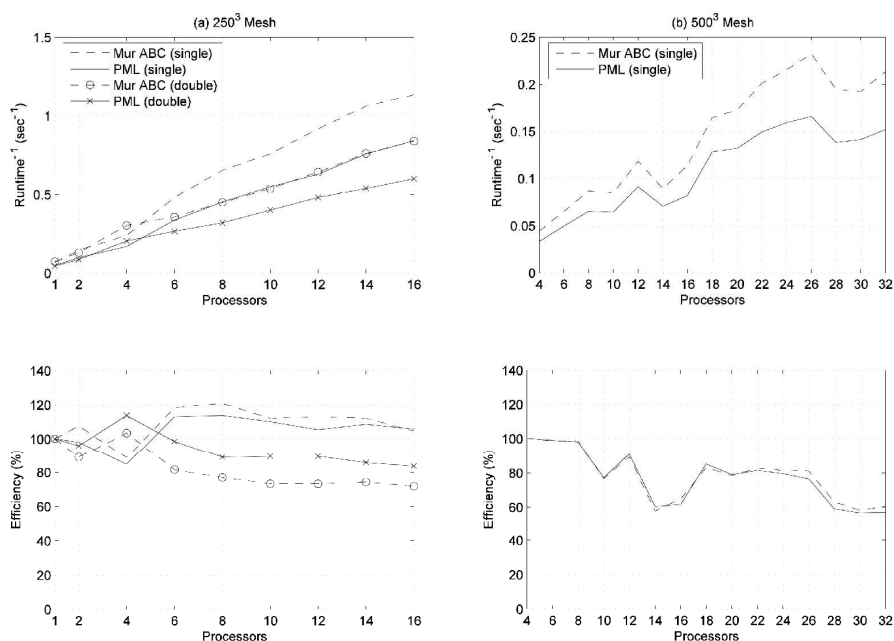

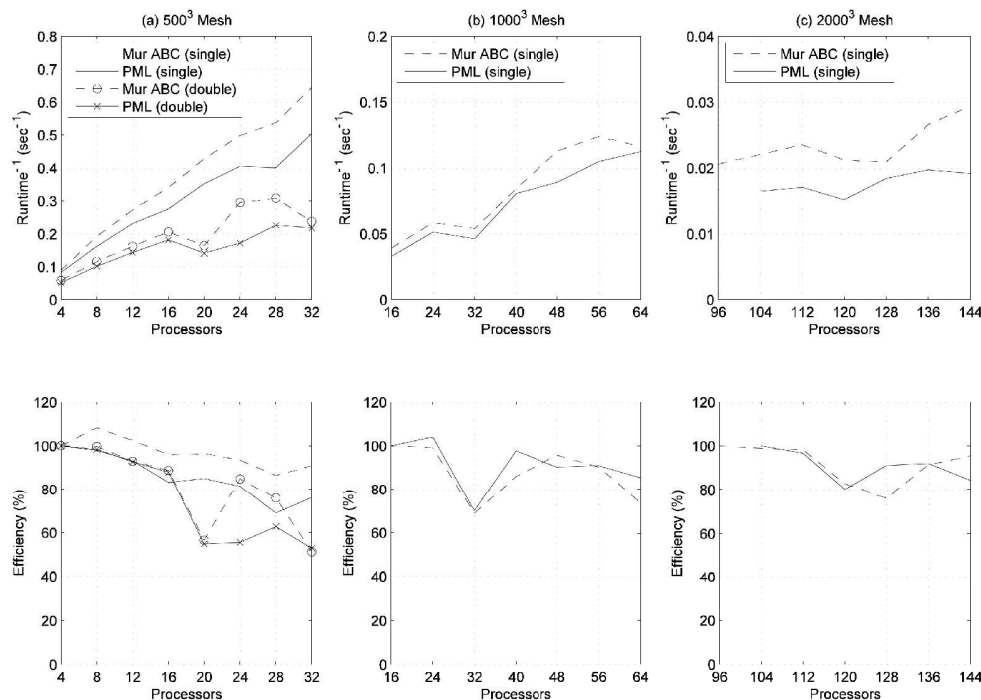Figure 2: Reciprocal of the runtime and scalability for parallel ADI-FDTD on SMP (250^3, 500^3 cells). [© Wiley]

Figure 3: Reciprocal of the runtime and scalability for parallel ADI-FDTD on DMCC (500^3, 1000^3, 2000^3 cells). [© Wiley]

The inverse runtimes and scaling efficiencies for the IBM p690 are presented in Figure 2. For the smallest domain (250^3 cells), there is an almost linear relation between inverse runtime and number of processors for single precision, confirming the good scalability of the algorithm. Efficiencies slightly in excess of 100% are attributed to caching effects. Despite having twice the communication overhead, the double precision simulations only took 40% longer, indicating that the benefit of solving multiple tridiagonal matrix systems of equations in parallel outweighs the communication penalty. For the 500^3 domain, the characteristics suffer from nonlinearities but the computational speedup remains significant, with efficiencies above 60%.

The DMCC gave a similar performance improvement from parallelization, as shown in Figure 3. Since the inter-process memory bandwidth is reduced compared to the SMP machine, linear inverse runtime characteristics are only observed for single precision computations in the smallest domain. The double precision runtimes are about twice as long for large numbers of processors which suggests that communication time outweighs the computation time. Nonetheless, employing parallel ADI-FDTD algorithm for ultra large mesh size computations (1000^3 and 2000^3) still provides a scalable speedup even for ultra large meshes. For the largest domain (2000^3) we attribute the non-linearities to the volume-proportional communication overhead, influence of cache size, load changes due to other cluster users, and non-optimal code placement by the queuing system with respect to the number of switches between communicating processors. Although the application of CPML as a boundary condition results in smaller reflections than for Mur ABC, the runtime increases by 20%. Although not shown here, the effects of cumulative propagation error are such that the agreement of the far field pattern of a Hertzian dipole computed with ADI-FDTD compares well with the analytical solution up to $CFL = 10$.

4. PARALLEL ADI-BOR-FDTD

Analyses involving objects with body-of-revolution (BOR) rotational symmetry, like cylindrical and coaxial waveguides and resonators, can be simplified through projection of the 3-D Yee-cell in cylindrical coordinates onto a 2-D plane with a Cartesian coordinate system [9]. To the best of our knowledge, we were the first to demonstrate a parallel implementation of ADI-BOR-FDTD [6]. In our parallel ADI-BOR-FDTD, we adopted a similar scheme to that described in Section 2 where spatial decomposition was applied in two directions (r, z).

We benchmarked the code by simulating the electromagnetic field inside a cylindrical perfect

electric conductor cavity resonator excited by a differentiated Gaussian pulse point source. The runtime and efficiency is shown in Figure 4 for simulations of 80×120 and 1600×2400 cells on the SMP machine. Surprisingly, the efficiency can be above 100% for large-scale (1600×2400 mesh) parallel BOR-FDTD and ADI-BOR-FDTD simulations due to cache effects [10].
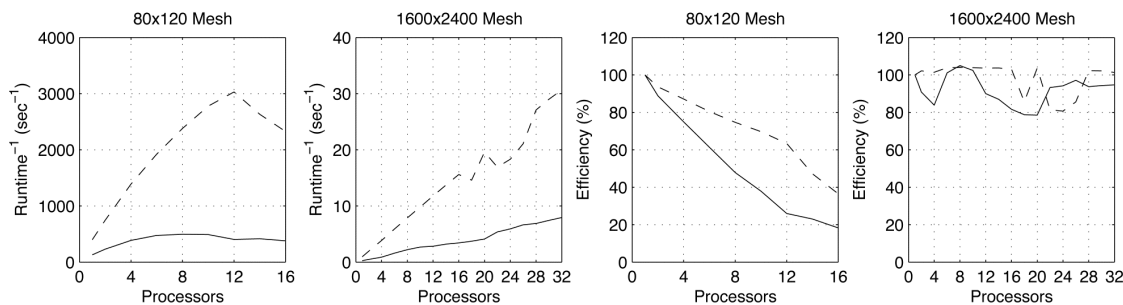

Figure 4: Reciprocal of the single step runtime and efficiency for the parallel ADI-BOR-FDTD (—) method compared against parallel BOR-FDTD (- - -) on SMP. [© IEEE]

5. CONCLUSIONS

Practical applications of implicit time domain methods such as ADI-FDTD are likely to require large amounts of memory that can be most economically accessed by exploiting parallel computer architectures. We have demonstrated good scalability on the widely available SMP and DMCC architectures, for domains of up to 8 billion mesh cells for 3D ADI-FDTD, and also for ADI-BOR-FDTD. We expect that our parallelisation approach can be adopted for related implicit FDTD methods.

ACKNOWLEDGMENT

REFERENCES

1. Stefanski, T. P. and T. D. Drysdale, "Parallel implementation of the ADI-FDTD method," submitted to *Microwave and Opt. Tech. Lett.*, 2008.
2. Yang, Y., R. S. Chen, W. C. Tang, K. Sha, and E. K. N. Yung, "Analysis of planar circuits using an unconditionally stable 3D ADI-FDTD method," *Microwave and Opt. Tech. Lett.*, Vol. 46, 175–179, 2005.
3. Choi, I.-J., S.-I. Yoon, J.-W. Kim, and T. Won, "ADI-FDTD for the modelling and simulation of VLSI interconnects," *Current Appl. Phys.*, Vol. 5, 356–364, 2005.
4. Jordan, H., S. Bokhari, S. Staker, J. Sauer, M. Elhelbawy, and M. Piket-May, "Experience with ADI-FDTD techniques on the cray MTA Supercomputer," *Proc. SPIE*, Vol. 4528, 68–76, Denver, USA, August 2001.
5. Stefanski, T. P. and T. D. Drysdale, "Parallel implementation of ADI-FDTD on shared and distributed memory computers," *Proc. 2nd European Conf. Antennas and Propag.*, Edinburgh, UK, November 2007.
6. Stefanski, T. P. and T. D. Drysdale, "Parallel ADI-BOR-FDTD algorithm," *IEEE Microwave Wireless Compon. Lett.*, Vol. 18, No. 11, 722–724, 2008.
7. Stefanski, T. P. and T. D. Drysdale, "Improved implementation of the Mur first-order absorbing boundary condition in the ADI-FDTD method," *Microwave and Opt. Tech. Lett.*, Vol. 50, No. 7, 1757–1761, 2008.
8. Stefanski, T. P. and T. D. Drysdale, "Novel implementation of the convolution perfectly matched layer in ADI-FDTD method," *IEEE Ant. Propag. Soc. Int. Symp.*, San Diego, USA, July 2008.
9. Chen, H.-L., B. Chen, Y. Yi, and D.-G. Fang, "Unconditionally stable ADI–BOR–FDTD algorithm for the analysis of rotationally symmetric geometries," *IEEE Microw. Wireless Compon. Lett.*, Vol. 17, 304–306, 2007.
10. Yu, W., R. Mittra, T. Su, Y. Liu, and X. Yang, *Parallel Finite-Difference Time-Domain Method*, Artech House, Boston, 2006.

An Improved GE's Method for Calculating Green's Functions in the Shielded Multilayered Structure

Huan Li, **Hao Gang Wang**, and **Hua Zhang**

The Electromagnetics Academy at Zhejiang University
Zhejiang University, Hangzhou 310027, China

Abstract— In this paper, we improve the Ge-Esselle's (GE's) method and apply it to calculate the multilayered Green's functions in the shielded structures. In the improved GE's method, 1) the poles are first extracted using a recursively contour integration method; 2) then the general pencil-of-function (GPOF) is performed to approximate the part of the spectral-domain Green's functions just along the real axis of k_ρ plane instead of the rooftop shaped path defined in the original GE's method; 3) Subsequently, the GE's analytical method is employed to obtain the spatial-domain Green's functions. The numerical results in this paper show that the improved GE's method can accurately and efficiently calculate the Green's functions in the shielded multilayered structure.

1. INTRODUCTION

In recent years, researchers have paid much attention to the multilayered structures encountered in the electromagnetic (EM) analyses of microstrip antennas and radio frequency integrated circuits (RFICs). In modeling the EM characteristics of these structures, we usually employ the method of moments (MoM), which the accurate and efficient calculations of the multilayered Green's functions are very critical.

The discrete complex image method (DCIM) was proposed for efficient calculation of the spatial-domain multilayered Green's functions. The DCIM does work in most conditions, but we found that it can not work well in the shielded structure, that is the first layer and the last layer are both perfect electric conductors (PECs). Recently, Ge and Esselle present a new complex image method that has the closed-form solutions. The main characteristic is to use an analytical expression to evaluate the spatial functions. But the original method has some drawbacks that can't have the results as expected. In this paper, we have some improvements to make it more robust and efficient for the shielded multilayered structure.

2. THE DIFFICULTY ENCOUNTERED BY THE DCIM FOR THE SHIELDED STRUCTURE

The basic idea of the DCIM is to approximate the spectral functions into a series of exponentials using the GPOF method. But we found that in the shielded structure, that is, the first layer interface and the last layer interface are both PECs, the values of spatial function G_{xx} oscillate rapidly around the accurate values performed by the Sommerfeld integrals (SIs). Fig. 1(a) shows the plots of G_{xx} using the DCIM and the SIs and Fig. 1(b) shows the relative error of the DCIM.

For the shielded structure, there should have no branch point for the spectral functions. However, when we perform the DCIM, we usually formulate the spectral functions in the form as formula (1) in order to use the Sommerfeld identity conveniently.

$$\widetilde{G}(k_\rho) = \frac{F(k_\rho)}{2jk_{zi}} \tag{1}$$

where $k_{zi}^2 = k_i^2 - k_\rho^2$, k_i is the wave number in the ith layer.

When $k_\rho = k_i$, it would result in $k_{zi} = 0$. Thus there is an artificial "branch point", which doesn't physically exist. We sample the values of spectral function $F(k_\rho)$ from the real axis of k_ρ plane, and also perform the DCIM along the same path. We can see from Fig. 2 that the DCIM can not accurately simulate the curve of the sampled values.

3. THE IMPROVED GE'S METHOD

The basic idea of the GE's method is as follows:
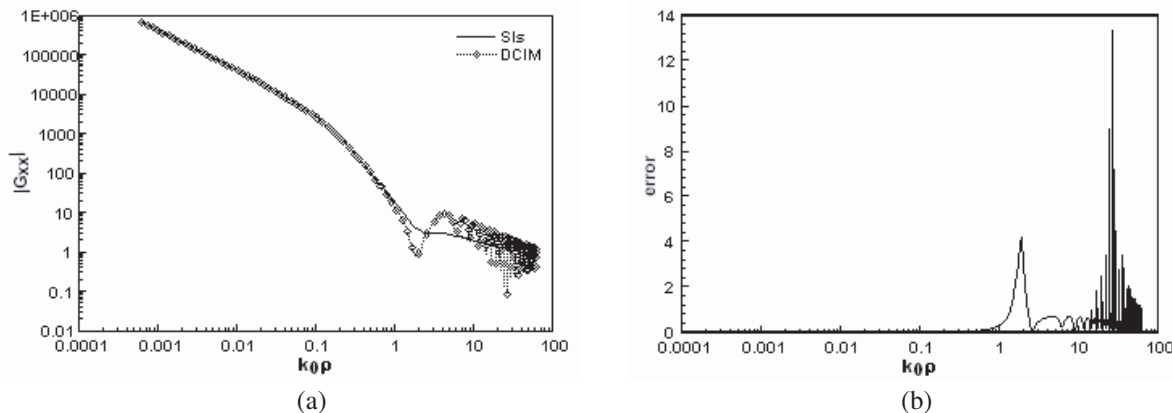
(a)   (b)

Figure 1: (a) The spatial-domain Green's function G_{xx} obtained by the DCIM and the SIs. (b) The relative error of the DCIM compared with the SIs.

1) Change the spectral function $\widetilde{G}(k_\rho)$ instead of $F(k_\rho)$ into a series of exponentials (formula (2)) using the GPOF along a rooftop shaped integration path L_1 (shown in Fig. 3) without extraction of SWPs and rewrite the SIs (formula (3))

$$\widetilde{G}(k_\rho) = \sum_{i=1}^{M} b_i e^{-s_i k_\rho} \tag{2}$$

$$\int_0^\infty \widetilde{G}(k_\rho) J_0(k_\rho \rho) k_\rho dk_\rho = \sum_{i=1}^{M} b_i \int_0^\infty e^{-s_i k_\rho} J_0(k_\rho \rho) k_\rho dk_\rho. \tag{3}$$

2) Use the GE's analytical identity to solve the integrations analytically (formula (4)).

$$\int_0^\infty e^{-s_i k_\rho} J_0(k_\rho \rho) k_\rho dk_\rho = \frac{s_i}{\left(s_i^2 + \rho^2\right)^{1.5}}. \tag{4}$$

The original GE's method didn't take into account the contribution of SWPs, which is always strong. Only performing the DCIM along the rooftop shaped integration path can't easily take into account the contribution of SWPs and guided wave. Fig. 4 shows the results using the original GE's method compared with the SIs.

Here, we do mainly two improvements: first, extract the SWPs and quasi-static terms before applying the GPOF; second, using a new integration path, that is, just along the real k_ρ axis.
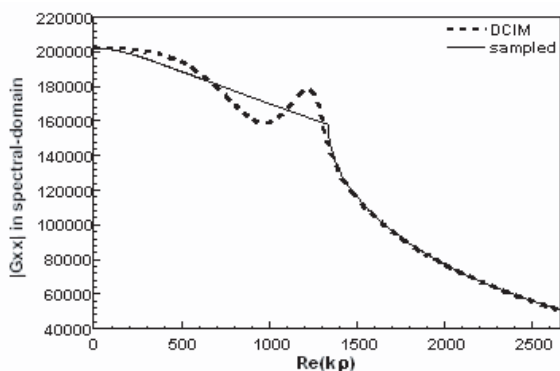

Figure 2: The values of the spectral-domain Green's function $\widetilde{G}_{xx}$ of the DCIM and the sampled one.
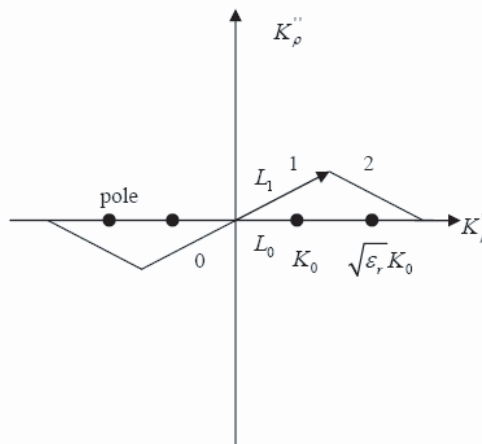
Figure 3: The suggested path in the original GE's method while performing the GPOF.

With the improvements, the GE's method can do well in the shielded multilayered structure. From Fig. 5, we can see that these two methods coincide with each other perfectly.
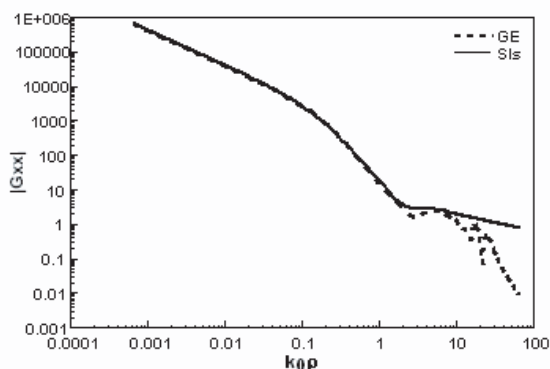

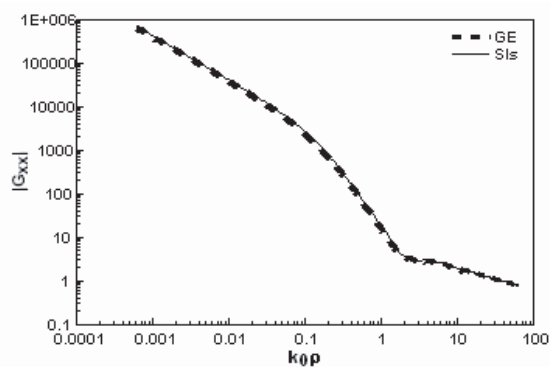Figure 4: The magnitude of G_{xx} using the original GE's method and the SIs.

Figure 5: The coincidence of the improved GE's method and the SIs.

4. CONCLUSION

From former description, some improvements are proposed to make the GE's method more robust and efficient. We extract SWPs before using the GPOF and perform the GPOF along the real k_ρ axis. With these improvements, the GE's method can be used to efficiently calculate the multilayered Green's functions of the shielded stratified structure. The numerical examples are given to verify it.

REFERENCES

1. Ling, F., J. Liu, and J. M. Jin, "Efficient electromagnetic modeling of three-dimensional multi-layer microstrip antennas and circuits," *IEEE Trans. Mircowave Theory Tech.*, Vol. 50, 1628–1635, Jun. 2002.
2. Hua, Y. and T. K. Sarkar, "Generalized pencil-of-function method for extracting poles of an EM system from its transient response," *IEEE Trans. Antennas Propagat.*, Vol. 37, 229–234, Feb. 1989.
3. Ge, Y. and K. P. Esselle, "New closed-form Green's functions for microstrip structure — Theory and results," *IEEE Trans. Antennas Propagat.*, Vol. 50, 1556–1560, Jun. 2002.
4. Fang, D. G., J. J. Yang, and G. Y. Delisle, "Discrete image theory for horizontal electric dipoles in a multilayered medium," *IEE Proc. — H*, Vol. 135, 297–303, Oct. 1988.