FDTD Computational Analysis of X-ray Transmission and Scattering Characteristics in Medical Imaging Diagnosis

Yasumitsu Miyazaki[1], **Koichi Takahashi**[1], **and Nobuo Goto**[2]

[1]Department of Electronics and Information Engineering, Aichi University of Technology
50-2 Manori, Nishihazama-cho, Gamagori 443-0047, Japan
[2]Department of Information and Computer Sciences, Toyohashi University of Technology
1-1 Hibarigaoka, Tempaku-cho, Toyohashi 441-8580, Japan

Abstract— X-ray diagnosis depends on the intensity of transmitting and scattering waves in X-ray propagation in biological media. X-ray is scattered and absorbed by tissues, such as fat, bone and internal organs. However, image processing for medical diagnosis, based on the scattering and absorption characteristics of these tissues in X-ray spectrum is not so much studied. Particularly, scattered waves in biomedical media cause serious deterioration of image quality. In this paper, X-ray transmission and scattering characteristics in biological media are studied using FDTD method with grid computer. Also, to remove the scattered waves and improve the medical image diagnosis, the waveguide-type grid filter is considered.

1. INTRODUCTION

X-ray is indispensably important for medical image diagnosis as a tool to obtain information on the internal part of human body. Recently, automatic X-ray image diagnosis based on digital signal processing technique has been developing rapidly [1, 2]. Using digital apparatus in medical diagnosis, a lot of digital images with high quality are obtained. CAD (Computer-Aided Diagnosis) system has been developing to analyze these medical information efficiently using computer and to extract the useful information for diagnosis automatically. In Fig. 1, CAD system provides advanced signal and image processing technique of medical information and supports medical imaging diagnosis effectively. X-ray diagnosis depends on the intensity of transmitted and scattered waves in X-ray propagation in biomedical media. To obtain precise information on tissues, such as fat, bone and internal organs in biomedical media, the characteristics evaluation of refractive index and absorption depending on X-ray wavelength are indispensable. Particularly, scattering characteristics are important to improve the accuracy of X-ray image diagnosis. For computational analysis for large area, parallel processing is important to study field equations with large number of elements. Numerical analysis using grid computer is one of the most useful parallel processings.

In this paper, X-ray transmission and scattering in biomedical media are studied using two-dimensional FDTD method with grid computer [3–7]. Also, to improve the quality of X-ray image, lossy waveguide-type grid filter is introduced and the filtering characteristics of scattered wave are studied numerically.

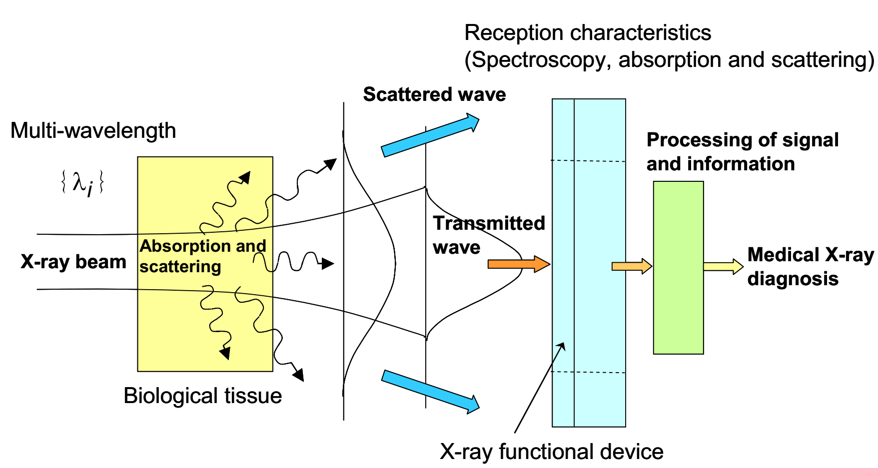


Figure 1: Characteristics of X-ray transmission and scattering and X-ray diagnosis system.

2. X-RAY TRANSMISSION AND SCATTERING ANALYSIS IN HUMAN BODY USING FDTD METHODIN

the human biological body region, Maxwell's equations are

$$\nabla \times \mathbf{E} = -\mu_0 \frac{\partial \mathbf{H}}{\partial t}, \quad \nabla \times \mathbf{H} = \varepsilon(\mathbf{r}) \frac{\partial \mathbf{E}}{\partial t} + \sigma(\mathbf{r})\mathbf{E} \tag{1}$$

where, $\varepsilon(\mathbf{r})$ and $\sigma(\mathbf{r})$ are the dielectric constant and conductivity of the material.

FDTD is demonstrated to show the characteristics of wave transmission, scattering, and absorption in X-ray region. The 2-D analysis space is shown in Fig. 2. The incident wave is x-polarized Gaussian beam with angular frequency $\omega = 2\pi f$, beam spot r_0 and beam waist at $z = z_0$. In the simulation model, the electromagnetic fields at point (i, j) at time $n\Delta t$ are calculated by difference equations,

$$\begin{aligned}
&E_x^n(i,j) = C_1 E_x^{n-1}(i,j) - C_2 J_x^{n-1}(i,j) + C_3\left\{H_z^{n-1}(i-1,j) - H_z^{n-1}(i-1,j) - H_y^{n-1}(i,j) + H_y^{n-1}(i,j-1)\right\}\\
&H_y^n(i,j) = H_y^{n-1}(i,j) - C_4\{E_x^n(i,j+1) - E_x^n(i,j)\},\ H_z^n(i,j) = H_z^{n-1}(i,j) + C_4\{E_x^n(i+1,j) - E_x^n(i,j)\}\\
&J_x^{n-1}(i,j)\begin{cases}\neq 0 & (j=1)\\ =0 & (j\neq 1)\end{cases},\ C_1 = \frac{1-\sigma\Delta t/(2\varepsilon)}{1+\sigma\Delta t/(2\varepsilon)},\ C_2 = \frac{\Delta t/\varepsilon}{1+\sigma\Delta t/(2\varepsilon)},\ C_3 = \frac{\Delta t/\varepsilon}{1+\sigma\Delta t/(2\varepsilon)}\frac{1}{\Delta s},\ C_4 = \frac{\Delta t}{\mu_0 \Delta s}
\end{aligned} \tag{2}$$

Here, y, z and t are discretized by $y = i\Delta s$, $z = j\Delta s$ and $t = n\Delta t$ and $0 \leqq i \leqq N_y$, $0 \leqq j \leqq N_z$.

The scatterers are atoms in biological media. In the region of the scatterers, complex dielectric constant ε^* and n_a^* are given by $\varepsilon^* = \varepsilon + \sigma/(j\omega) = n_a^{*2}\varepsilon_0$, $n_a^* = n_1 - jn_2$. In this analysis, the shape of a scatterer is assumed to be a square. In FDTD simulation, the incident wave is generated by current density

$$\begin{aligned}
&J_x^n(i,1) = J_0 \exp\left\{-\left(\frac{(i-i_0)\,\Delta s}{r_0}\right)^2\right\} H(n\Delta t)\sin(2\pi f n\Delta t)\\
&H(n\Delta t) = \begin{cases} 1, & n \le n\Delta t \le \tau \\ 0, & n\Delta t < 0,\ n\Delta t > \tau \end{cases}
\end{aligned} \tag{3}$$

where f is the frequency of the incident wave, k_0 is the wave number, $f = 1/\lambda_0\sqrt{\varepsilon_0\mu_0} = k_0/2\pi\sqrt{\varepsilon_0\mu_0}$, $y_0 = i_0\Delta s$ is the center point of the incident beam, r_0 is the beam waist at $z = z_0$ $(j = 1)$, τ is the transmission time.

3. PARALLEL PROCESSING OF TRANSMISSION AND SCATTERING IN HUMAN BODY USING GRID COMPUTER

In X-ray simulation, the wavelength is very short, so the analysis space much larger than the wavelength is required for FDTD analysis. Therefore, parallel processing of FDTD using grid computer is indispensably important. Grid computer consists of high-speed network with computers those

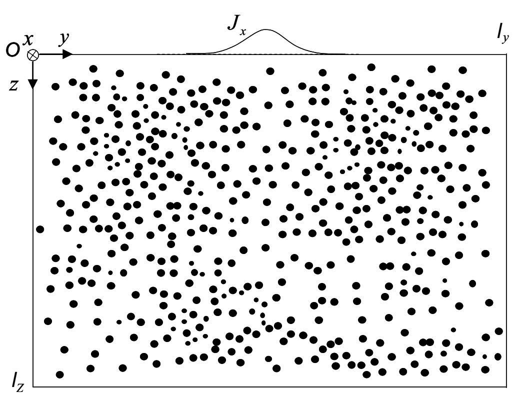


Figure 2: X-ray transmission and scattering in human body.

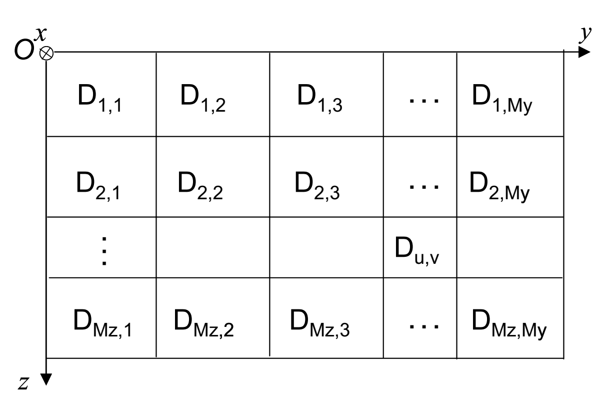


Figure 3: Parallel processing of transmission and scattering in human body, subdomain $D_{u,v}$ $1 \leqq u \leqq M_z$, $1 \leqq v \leqq M_y$.

can work independently. Recently, the performance of PC grows dramatically and its cost is decreasing. Therefore, it is possible to establish the PC grid computer with high performance and low cost. In the parallel processing using grid computer, the high-speed data communication is required among processors. The general way for data communication is the message passing. When one process sends data to another process, data copying from the memory of sending process to the one of receiving process is occurred and this procedure becomes heavy load for processors. Therefore, parallel algorithm with less amount of data communication is desired. In parallel processing using grid computer, the total analysis space is divided into subdomains of Mz×My as shown in Fig. 3 and one divided subdomain $D_{u,v}$ for each u index is assigned to one of My computers. For the calculation of values on the boundary of subdomains, the values in the adjacent subdomains are necessary, according to the FDTD Equation (2). We develop a new approach to perform FDTD parallel processing using grid computer network. In FDTD analysis, the electromagnetic fields are calculated by difference equations derived from Maxwell's equations. We transform the difference equations to discrete linear equations. By the definition of the vector $\mathbf{X}^n_{u,v}$ for the electromagnetic field in subdomain $D_{u,v}$ as $\mathbf{X}^n_{u,v}\left(E^n_x, H^n_y, H^n_z\right)_{u,v}$ and the vector $\mathbf{\Phi}^n_{u,v}$ for the source current as $\mathbf{\Phi}^n_{u,v} = (J^n_x, 0, 0)$, discrete equations for FDTD can be expressed by

$$\mathbf{X}^n_{u,v} = \sum_{u',v'} \left(\mathbf{A}\mathbf{X}^{n-1}_{u',v'} + \mathbf{\Phi}^{\mathbf{n-1}}_{\mathbf{u}',\mathbf{v}'}\right) \tag{4}$$

where $\mathbf{A}$ is the coefficient matrix determined by Equation (2). In parallel processing, $\mathbf{X}^n_{u,v}$ in subdomain $D_{u,v}$ are calculated by the computer assigned to the corresponding subdomain independently, except the value of the vector at the boundary of adjacent subdomains. The flow of parallel processing using grid computer is shown in Fig. 4. For the accuracy of simulation results, the magnitude $\|\mathbf{F}\|$ of vector $\mathbf{F} = \mathbf{X}^n_{u,v} - \mathbf{X}^n_{u,v+1}$ at the boundary surface $S^{(u)}_{v,v+1}$ is evaluated and when $\|\mathbf{F}\| < 10^{-\zeta}$, the processor goes back to the iteration, using $\mathbf{X}^n_{u,v} = \left(\mathbf{X}^n_{u,v} + \mathbf{X}^n_{u,v+1}\right)/2$ at $S^{(u)}_{v,v+1}$.

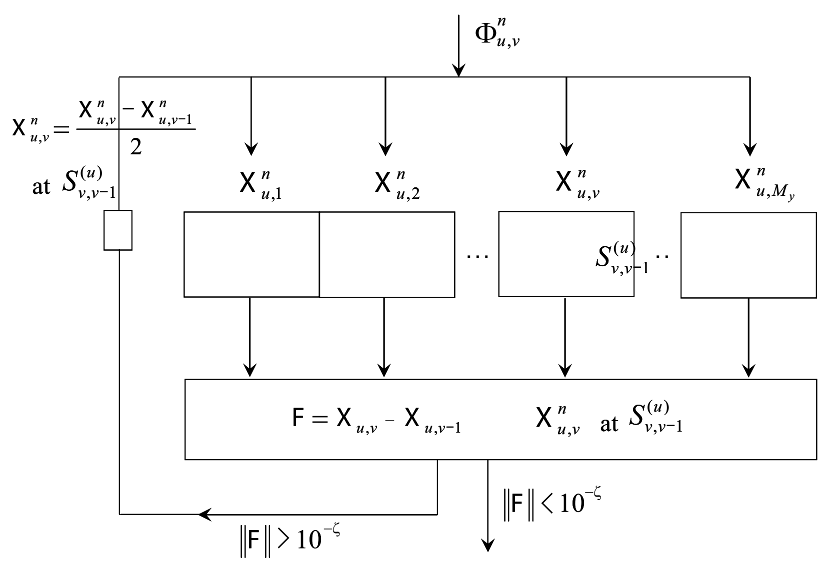


Figure 4: Flow of parallel processing using grid computer.

4. TRANSMISSION AND SCATTERING CHARACTERISTICS OF X-RAY GAUSSIAN BEAM IN BIOLOGICAL TISSUES

Transmission and scattering characteristics of incident X-ray Gaussian beam are evaluated using FDTD method. Simulation model is shown in Fig. 5 and numerical parameters are shown in Table 1. The model in Fig. 5 corresponds to one subdomain in Fig. 3. For the incident wave, $f = 1.5 \times 10^{18}$ Hz (λ_0 =0.2 nm), $r_0 = 15\lambda_0 = 3$ nm, $J_0 = 10^9$ A/m^2 and $\tau = 2 \times 10^{-17}$ s, are used. The size of a subdomain is 30 nm along y-direction and 20 nm along z-direction. Here, the number

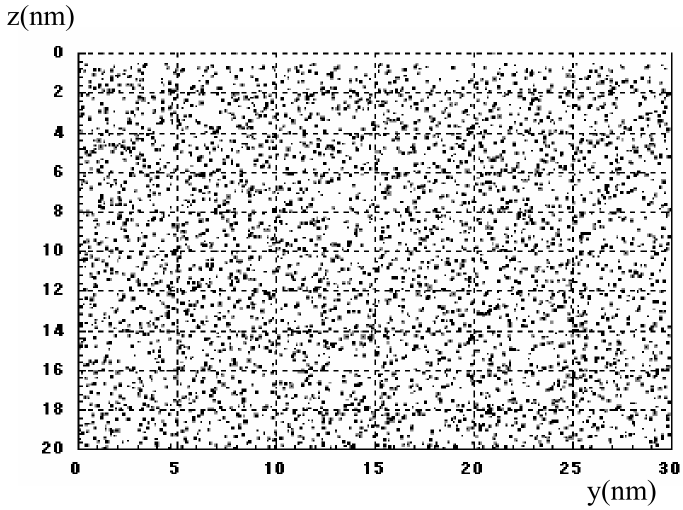


Figure 5: Distribution of dielectric constant of random media $\varepsilon_r^{(1)}$ ($N = 3800$, $a = 0.1 \sim 0.2\,\text{nm}$, $n_a^* = n_1 - jn_2$).

Table 1: Parameters for the simulation of X-ray scattering and attenuation characteristics.

Parameters	Values
f: Frequency of incident wave	1.5×10^{18} Hz
λ_0: Wavelength of incident wave	0.2 nm (c/f)
ℓ_y: Length of an analysis space (y-direction)	30 mn ($150\lambda_0$)
ℓ_z: Length of an analysis space (z-direction)	20 nm ($100\lambda_0$)
y_0: Center point of incident beam	15 nm ($\ell_y/2 = 75\lambda_0$)
r_0: Beam spot at $z = z_0$	3 nm ($r_0/\lambda_0 = 1/20$)
Δs: Length of a cell	0.01 nm ($\Delta s/\lambda_0 = 1/20$)
Δt: Time increment	2.25×10^{-20} s
N: A number of atoms	3800
a: Length of a side of atoms	0.1∼ 0.2 nm ($a/\lambda_0 = 0.5 \sim 1.0$)
n_a^*: Complex refractive index of atoms	0.99 -j0.005 (Case 1)
	0.995 -j0.002 (Case 2)
	0.999 -j0.001 (Case 3)
	0.9995 -j0.0005 (Case 4)

of cells Ny and Nz are 3000 and 2000, respectively. As shown in Table 1, four random media models, Case 1-4 with different complex refractive index of atoms are considered. To evaluate statistical properties of X-ray transmission and scattering, five realizations of random media are analyzed in each case and the average amplitude of transmitted electric fields averaged over five realizations are presented. Fig. 6 shows the average amplitudes of transmitted electric fields $\left|E_x^{(t)}\right|$ at $z = \ell_z = 20\,\text{nm}$. Here,

$$\left|E_0^{(inc)}\right| = \max_{t\in[t',t'+T]} \left|E_0^{(inc)}(y, \ell_z, t)\right|, \quad \left|E_x^{(t)}\right|_{\max} = \frac{1}{S}\sum_{S=1}^{S} \max_{t\in[t',t'+T]} \left|E_x^{(s)}(y, \ell_z, t)\right| \tag{5}$$

t' is the time when the fields assumed to be steady state. FDTD simulation is conducted using five models in each case and the results of electric field amplitudes are averaged to obtain the statistical properties. Simulation results show that qualitatively, as n_a^* becomes smaller than 1.0, the attenuation of the electric field amplitude at the center point of the beam becomes large due to not only absorption effects, but large scattering effects, and the intensity of scattering field becomes relatively strong at the foot of the beam.

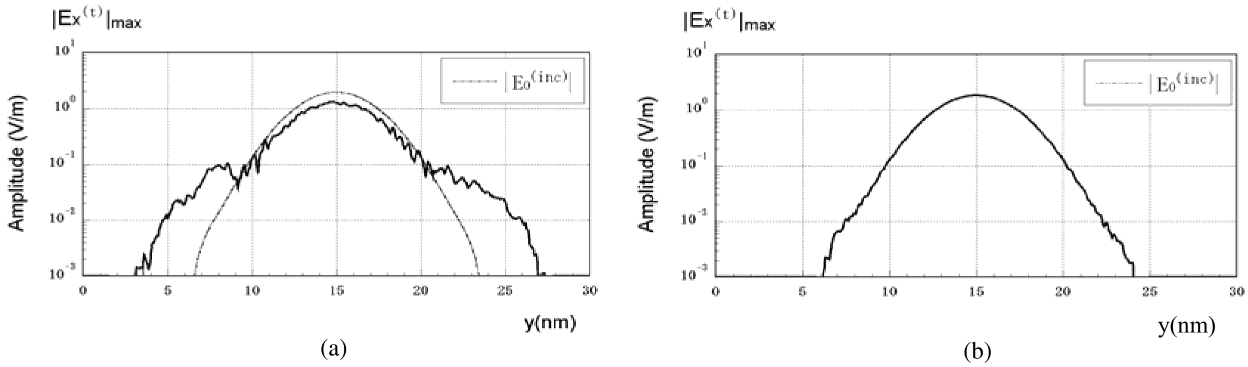


Figure 6: Average amplitude of transmitted and scattered electric field in biological tissues at $z = \ell_z = 20\,\text{nm}$. (a) Case 1, (b) Case 4.

5. X-RAY SCATTERING AND FILTERING CHARACTERISTICS BY WAVEGUIDE-TYPE GRID FILTER IN BIOLOGICAL TISSUES

For X-ray, biological tissues have scattering and absorption characteristics. Particularly, X-ray scattering characteristics are important to improve the accuracy of X-ray image diagnosis. In X-ray image diagnosis, the use of the grid is efficient to remove the scattering wave from the transmitting wave. In this paper, lossy waveguide-type grid is proposed and its filtering characteristics are evaluated using FDTD simulation. The wall of grid is lossy dielectric material with complex refractive index of 0.98-j0.05 and width of the walls d, distance of the walls D, and length of the wall ℓ_g. As shown in Fig. 7, the random media model with grid is considered. As shown in Fig. 8, the total analysis space is divided and will be analyzed using parallel processing with grid computer.

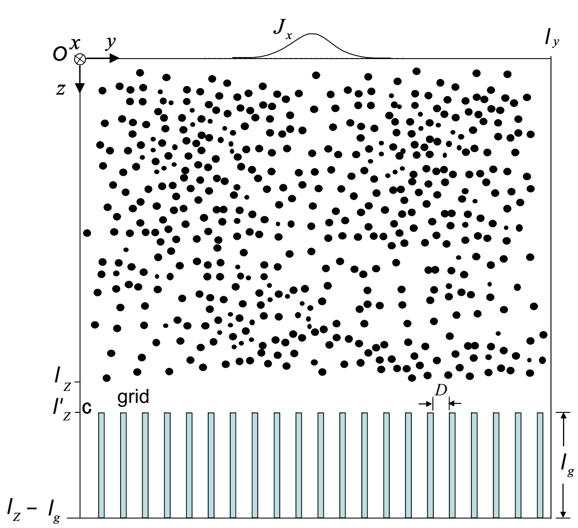


Figure 7: Grid filter of X-ray scattering.

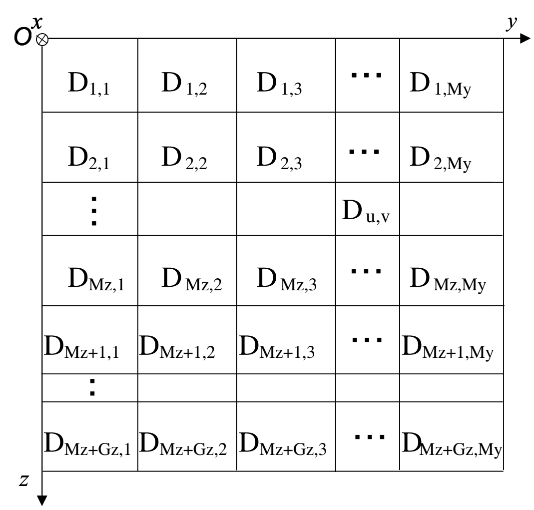


Figure 8: Parallel processing of grid filter analysis for X-ray scattering.

6. CONCLUSIONS

In this paper, the electromagnetic transmission and scattering characteristics, those are important in X-ray image diagnosis are analyzed using FDTD method. To analyze a big space larger than the wavelength, parallel processing using grid computer is considered. Also, to improve the quality of X-ray image, the lossy waveguide-type grid filter is considered and the filtering characteristics will be evaluated. Using grid filter, off-axial scattered waves would be decreased efficiently.

REFERENCES

1. Miyazaki, Y., *Jour. of IEICE*, Vol. 83, No. 2, 132–136, 2000.
2. Aichinger, H., et al., Radiation Exposure and Image Quality in X-Ray Diagnosis Radiology, 2003.
3. Miyazaki, Y., J. Sonoda, and Y. Jyonori, *Trans. IEE Japan*, Vol. 117-C, No. 1, 35–41, 1997.
4. Takahashi, K. and Y. Miyazaki, *Trans. IEE Japan*, Vol. 120-C, No. 12, 1905–1912, 2000.
5. Takahashi, K., H. Xinmin, and Y. Miyazaki, *Digest of ISMOT-2005*, B-16, 109, Fukuoka, Japan, 2005.
6. Rodriguez, G., Y. Miyazaki, and N. Goto, *IEEE Trans. Antennas and Propagation*, Vol. 54, No. 3, 785–796, 2006.
7. Takahashi, K., Y. Miyazaki, and N. Goto, *Proc. of PIERS 2006*, 568, Tokyo, Japan, 2006.

Photonic-crystal Lens Coupler Using Negative Refraction

Pi-Gang Luan and Kao-Der Chang

Wave Engineering Laboratory, Department of Optics and Photonics
Central University, Chungli 320, Taiwan, China

Abstract— A novel photonic crystal lens working in the negative refraction frequency region that can focus a beam of light into a wavelength-sized spot is proposed. The device is formed by drilling air holes of triangular lattice in a portion of a dielectric slab, and the whole structure is designed appropriately such that it can be easily fabricated using today's technology. The finite-difference time-domain simulations confirm that the focusing effect is mainly due to the negative refraction in the photonic crystal region, and the focused light can be efficiently coupled into a dielectric waveguide.

1. INTRODUCTION

Photonic crystals (PC) are artificial periodic dielectric structures made for controlling light [1, 2]. The propagating mode of the electromagnetic (EM) wave in a photonic crystal is the Bloch wave [3], and its dispersion relation is the photonic band structure, containing passbands and bandgaps — the frequency ranges forbidding the propagation of light. Up to now, most applications of the PCs rely on the presence of the bandgaps, and various devices have been designed for reflecting, trapping, and guiding light [4, 5]. Beyond the bandgaps, the photonic passbands are also useful [6]. When operating in the linear region well below the first gap, a PC plays the role of an effective homogeneous medium, and lens-like devices can be constructed from it [7]. However, since such devices must be working at the long-wavelength limit (there the lattice constant is much smaller than the wavelength), it is difficult to fabricate them for applications in optical or telecommunication regime.

Recently, the interesting phenomenon of negative refraction (NR) has received much attention [8–21]. A block of negative refraction medium (NRM) has many fascinating properties. For example, a slab of NRM is a "superlens", which can focus the light emitted from a point source into a subwavelength spot, overcoming the diffraction limit [9, 12, 14–18]. Also, a concave lens made of NRM can focus a beam of light into a sharp image spot [19, 20]. In fact, an NRM can be realized by PC under certain conditions. When an EM wave is injected into a PC from a homogeneous medium, the energy flow and phase propagation directions of the transmitted EM wave are determined by the phase matching condition on the interface together with the equifreqency surfaces (EFS) of the two media [11, 12]. According to this theory, a photon in PC sometimes acquires a negative effective mass, and then the PC becomes an effective NRM [11–13]. More importantly, it has been found that the NR of PC can appear in high frequency bands [11, 13, 16–18]. This fact implies that it is more practical to design devices operating in the NR regime than operating in the long-wavelength regime [21].

In practice, an useful integrated-optical system must have large enough coupling efficiency between different parts of the system. One important step towards this goal is to reduce the insertion loss for the light launched from air and injected into a conventional (dielectric) waveguide (CW) or a photonic crystal waveguide (PCW). Usually the width of the waveguide can be made as small as one wavelength, whereas the incident light has a much larger beam waist. Therefore, a device to focus the incident light before sending it into the waveguides is needed. Since it is difficult to reduce the spot size of the focused light dramatically using conventional methods, we try to use the NR effect of the PC to overcome this problem.

In this letter, we demonstrate numerically that by using a PC lens device working in the NR regime, focusing a beam of light into a wavelength-sized spot is possible. We also show that the focused light can be efficiently coupled into a CW if its width is about the same size as the spot. The shape and the inner structure of the device is appropriately designed so that it can be easily fabricated using today's technology even for using in the optical regime. Besides, to evaluate the size of the focus spot, we propose a method based on the calculation of root-mean-square (RMS) of the field distribution, which gives us more definite results than using other methods.

2. STRUCTURE DESCRIPTION AND SIMULATION RESULTS

The two-dimensional (2D) PC being considered here is a triangular lattice of air holes in a dielectric slab with dielectric constant $\epsilon = 12.96$ (e.g., GaAs). The hole radius is $r = 0.4a$, here a is the lattice constant. The photonic band structure is calculated using the plane wave expansion method [2] for the transverse magnetic (TM) modes (the electric fields are parallel to the axis of the holes). In the calculation, 961 plane waves were used, and the result is shown in Fig. 1. It is observed that, in the second band, the Γ point is a local maximum of the band structure curve. According to Notomi's theory [11], around that point in the **k**-space there is a frequency range the PC has NR property. In the following simulations, the frequency of the incident light will be taken as $f = 0.3c/a$, here c is the speed of light in vacuum.

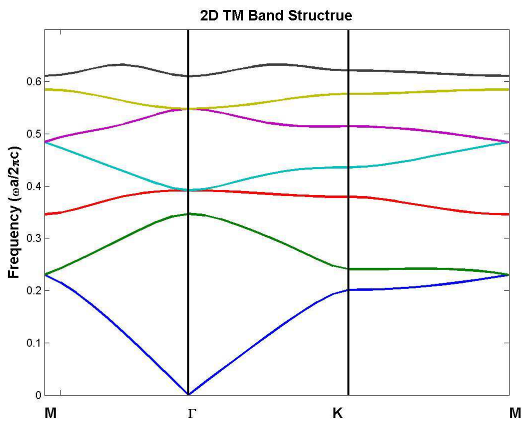


Figure 1: Band structure of the TM (E-polarization) modes of the PC. The radius of the air holes is $r = 0.4$a and the dielectric constant of the background material is $\varepsilon = 12.96$.

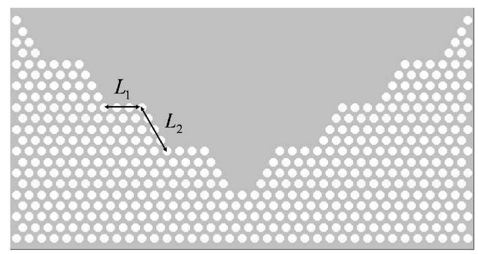


Figure 2: The shape and structure of the PC lens. The two periods of the terraced structure are $L_1 = 3a$ and $L_2 = 4a$, respectively. The size of the lens is $38a \times \sqrt[5]{3}a$.

The structure of the device is shown in Fig. 2. The transverse and longitudinal dimensions of the PC slab are $38a$ and $5\sqrt{3}a$, respectively. The "hole region" and the remaining homogeneous dielectric region are separated by a terraced V-shaped interface, with periods $L_1 = 3a$ and $L_2 = 4a$ along two symmetry directions of the lattice periodicity. The reason for choosing such an interface is because it is easier to align the holes along the symmetry directions of the lattice than to arrange them along an smooth curve [19]. Besides, for the utility in the integrated-optical system, the "hole in dielectric" system is more practical than the "dielectric rods in air" system. If this lens is used for focusing the light of wavelength 1550 nm, the corresponding lattice constant a would be 465 nm, and the diameter of the air holes is $2r = 372$ nm, quite easy to fabricate using today's technology.

The optical behavior of the device can be simulated using the 2D finite-difference time-domain (FDTD) method [22]. The absorbing boundary conditions are the perfectly matched layers [23]. The lattice plane of the PC is taken to be the XZ plane, and the Z-axis is the propagation direction of the incident light beam. The source **E** field is assumed to be a y-polarized EM wave ($\mathbf{E} = E\hat{\mathbf{y}}$) with a Gaussian type modulus along the transverse direction, given by

$$E(\omega, r) = E_0 \exp\left[-\frac{(r - r_0)^2}{W^2}\right] \exp\left(-j\omega t\right), \tag{1}$$

here r is the transverse position, r_0 is the reference center for the source field, and W is the waist of the beam. Throughout the paper, we assume that W is always greater than 3λ, here λ is the wavelength.

In the following simulations, the light beam is incident onto the PC lens from below. After passing through it the light propagates in the air and then converges at the focal point. To evaluate the spot size of the image and locate the position of the focal point, we define the RMS width of the field distribution (at z) as

$$\Delta_{\text{RMS}}(z) = \sqrt{\frac{\sum_r |E(r, z)|^2 r^2}{\sum_r |E(r, z)|^2}} \tag{2}$$

where r is the transverse spatial distance from the distribution center, and $E(r,z)$ is the amplitude of the E field evaluated at (r,z). The $\Delta_{\mathrm{RMS}}(z)$ for incident beams of different waists ($W = 3\lambda, 6\lambda, 9\lambda, 15\lambda$) are shown in Fig. 3. For each case, the location of the focal point is indicated by the z value that corresponds to the smallest Δ_{RMS}, say, z_0, and the spot size is given by $2\Delta_{\mathrm{RMS}}(z_0)$. From these results we find that the spot size is about one wavelength, and the focal point is located at about $z_0 = 7.56\lambda$.

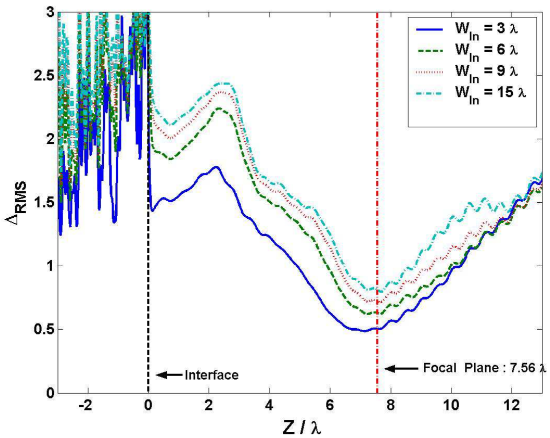


Figure 3: The "root-mean-square"(RMS) widths of the transverse distribution of the E field, evaluated for incident beams of different waists, Win is the waist width. The spot sizes are around one wavelength, and the focal lengths are almost the same in all the cases. The dash line indicates the position of the output interface, whereas the dotted dash line shows the location of the focal plane.

We also compare the result to that obtained by the "maximum to the first minimum" (MTFM) method, as shown in Fig. 4. In the MTFM method, we define the focal point as the location of the peak of the EM field modulus, and the half width of the image as the distance from the peak to the first minimum along the transverse direction. In Fig.4, the range of the beam waist W is chosen between 3λ and 20λ. In this range we find that: first, the RMS method always gives us a beam width larger than that obtained by using the MTFM method; and second, when $4\lambda < W < 9\lambda$, the focal distance obtained by the RMS method is a little smaller than that by the MTFM method. These result seems to imply that the meaning of "subwavelength imaging" or "subwavelength focusing" is a little ambiguous. To avoid the ambiguity, a clear definition about the image size like ours is needed.

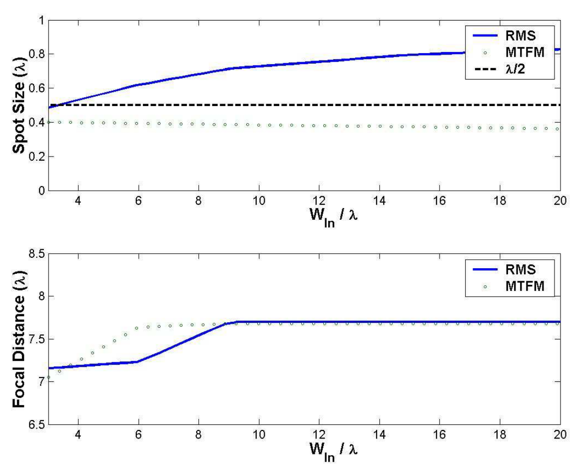


Figure 4: The spot size and the focal distance obtained by using the RMS and MTFM methods. The dash line indicates the half wavelength.

Figure 5 illustrates the modulus of the resulting **E** field for two different situations. In Fig. 5(a), a typical simulation ($W_{\text{in}} = 6\lambda$) of the focusing phenomenon is shown. The light emerging from the output side of the PC lens travels in the air and converges into a focus point, forming an elliptical image. The focusing ability of the PC lens is obvious. However, in order to examine if this focusing effect is indeed caused by the NR effect in the PC area, we do the following test. We replace the "hole region" of the PC lens with a naively averaged medium and then re-simulate the focusing ability of this modified lens. The filling fraction of the PC in the hole region is $f = 2\pi r^2/\sqrt{3}a^2 = 0.5804$, so we have the averaged dielectric constant $\bar{\epsilon} = f + 12.96(1 - f) = 6.0182$, or the averaged refraction index $\bar{n} = \sqrt{\bar{\epsilon}} = 2.4532$. The result of the test is shown in Fig. 5(b). As one can see, the original image disappears, although a much weaker new image near the slab edge has been created. This result seems to imply that both the NR effect in the PC area and the discrete nature of the boundary of the hole region (there the holes are arranged to form a grating-like structure) are important for the formation of the wavelength-sized spot in the far-field region.

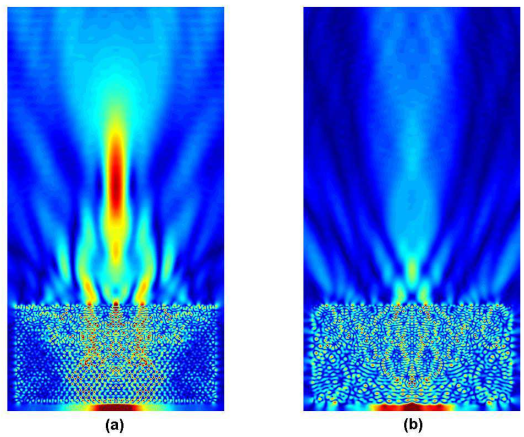


Figure 5: The intensity of the E-field in two different situations. The beam waist is 6λ. (a) The incident beam is focused by the PC lens. A wavelength-sized spot can be easily observed. (b) When the "hole region" in the PC lens is replaced by a naively averaged medium, the spot disappears.

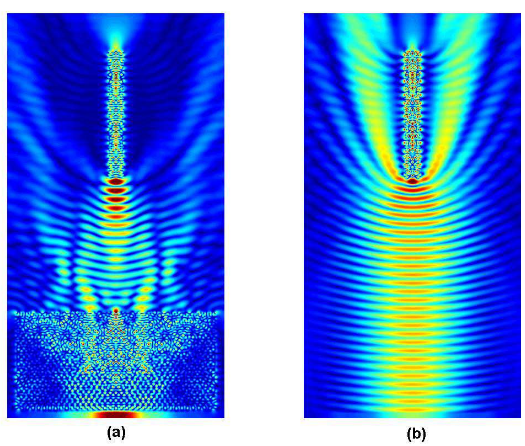


Figure 6: (a) The focused light can be efficiently coupled into a CW of width $3a$. (b) The light cannot be efficiently coupled into the CW without the PC lens.

The PC lens can also be used as an optical coupler. To examine this statement, we add a CW into the original setup, and then recompute the E field. The result is shown in Fig. 6(a). The material of the CW has a dielectric constant $\epsilon = 12.96$, which is the same as the slab. The width of the waveguide is $3a$, about one wavelength. We choose the input edge of the waveguide to be located at the same position of the focal point, *i.e.*, 7.56λ from the output-side edge of the dielectric slab. For comparison, we also plot the field modulus in Fig. 6(b) for the situation that there is no

PC lens between the light source and the waveguide. The simulation result clearly indicates that the light power can be efficiently coupled into the waveguide.

3. CONCLUSION

In conclusion, we have proposed a novel optical component that can focus a beam of light into a small spot of wavelength-sized width. The device is a PC lens working in the NR frequency range, formed by drilling air holes of triangular lattice in a portion of a rectangular dielectric slab background. The "hole region" and the remaining homogeneous dielectric region are separated by a terraced V-curve shaped interface. This structure can be easily fabricated. Based on the band structure calculation of the PC and the FDTD simulations on the system, the optical behavior of the device can be predicted. The spot sizes for incident beams of different waists, evaluated by using the root-mean-square (RMS) method, are all about one wavelength. The image spot vanishes when the hole-region is replaced by an averaged homogeneous medium. This fact indicates that the focusing effect is mainly due to the NR in the PC area. Our simulations also confirm that the device can transmit the focused light efficiently into a dielectric waveguide, thus can be used as an optical coupler.

ACKNOWLEDGMENT

The authors are grateful to financial support from National Science Council (Grant No. NSC95-2221-E-008-114-MY3) and Ministry of Economic Affairs (Grant No. 95-EC-17-A-08-S1-0006) of Taiwan.

REFERENCES

1. Joannopoulos, J. D., R. D. Meade, and J. N. Winn, *Photonic CrystalsCMolding the Flow of Light*, Princeton University Press, 1995.
2. Sakoda, K., *Optical Properties of Photonic Crystals*, Springer-Verlag, 2001.
3. Kittel, C., *Introduction to Solid State Physics*, 7th ed., John Wiley & Sons, Inc., 1996.
4. *Photonic Crystals and Light Localization in the 21st Century*, edited by C. M. Soukoulis, NATO Science Series, Vol. 563, 2001.
5. Chen, C.-C., H. D. Chien, and P. G. Luan, *Appl. Opt.*, Vol. 43, 6187, 2004.
6. Halevi, P., A. A. Krokhina, and J. Arriaga, *Appl. Phys. Lett.*, Vol. 75, 2725, 1999.
7. Kuo, C. H. and Z. Ye, *J. Phys., D*, Vol. 37, 2155, 2004.
8. Veselago, V. G., *Sov. Phys. Usp.*, Vol. 10, 509, 1968.
9. Pendry, J. B., *Phys. Rev. Lett.*, Vol. 85, 3966, 2000.
10. Shelby, R. A., D. R. Smith, and S. Schultz, *Science 292*, Vol. 77, 2001.
11. Notomi, M., *Phys. Rev. B*, Vol. 62, 10696, 2000.
12. Luo, C., S. G. Johnson, D. J. Joannopoulos, and J. B. Pendry, *Phys. Rev. B*, Vol. 65, 201104, 2002.
13. Foteinopoulou, S. and C. M. Soukoulis, *Phys. Rev. B*, Vol. 67, 235107, 2003.
14. Cubukcu, E., K. Aydin, E. Ozbay, S. Foteinopoulou, and C. M. Soukoulis, *Nature*, Vol. 423, No. 604, London, 2003.
15. Cubukcu, E., K. Aydin, E. Ozbay, S. Foteinopoulou, and C. M. Soukoulis, *Phys. Rev. Lett.*, Vol. 91, 207401, 2003.
16. Xiao, S. S., M. Qiu, Z. C. Ruan, and S. He, *Appl. Phys. Lett.*, Vol. 85, 4269, 2004.
17. Guven, K., K. Aydin, K. B. Alici, C. M. Soukoulis, and E. Ozbay, *Phys. Rev. B*, Vol. 70, 205125, 2004.
18. Martnez, A., H. Mguez, A. Griol, and J. Mart, *Phys. Rev. B*, Vol. 69, 165119, 2004.
19. Vodo, P., P. V. Parimi, W. T. Lu, and S. Sridhar, *Appl. Phys. Lett.*, Vol. 86, 201108, 2005.
20. Foca, E., H. Fll, F. Daschner, V. V. Sergentu, J. Carstensen, S. Frey, R. Knchel, and I. M. Tiginyanu, *phys. stat. sol. (a)*, Vol. 202, R35-R37, 2005.
21. Berrier, A., M. Mulot, M. Swillo, M. Qiu, L. Thyle, A. Talneau, and S. Anand, *Phys. Rev. Lett.*, Vol. 93, 073902, 2004.
22. Taflove, A. and S. C. Hagness, *Computational Electrodynamics: The Finite-Difference Time-Domain Method*, Artech House Publishers, Boston, 2000.
23. Berenger, J. P. and J. Comput., *Phys.*, Vol. 114, 185, 1994.

Photonic Crystals with Hexagonal Periodicity for Efficient Light Emission LED

Rafie Mavaddat[1], **Dang Hoang Long**[2], **and Sang-Wan Ryu**[2]

[1] Photonics Group, Engineering Department, Cambridge University, UK
[2] Department of Physics, Chonnam National University, Gwangju, Republic of Korea

Abstract— Application of a slab Photonic Crystal layer could enhances the radiation efficiency of a Light Emitting Diode. In this paper we briefly outline the concept of photonic bandgap introduced by Photonic Crystals with hexagonal periodicity and review the plane wave decomposition method for the analysis and computation of Photonic Crystal band structures. Our computer simulation conforms to results reported in the literature. The mechanism for the enhancement of Light Emitting Diodes is explained.

1. INTRODUCTION

Photonic Crystals (PCs) are assuming important role in many optical devices by virtue of their ability to inhibit the propagation of light in undesirable directions. There are numerous applications of the photonic crystals. An example of this is the application of photonic crystal in improving the efficiency of Light Emitting Diodes (LEDs).

LEDs are small and convenient light sources that are used in backlight display units, sign illuminations, vehicle and traffic signals and even for architectural effects in interior lighting. However, the efficiency of these devices is limited, as the emitted light that is required to propagate outwards along the axis of the device is mostly guided in lateral directions. A layer of the PC, however, substantially modifies the guiding properties of the slab LED waveguide and prevents the lateral propagation of light and enhances the emission properties of the LED.

PCs are periodic dielectric devices that similar to conventional periodic structures exhibit nonlinearity in ω-k dispersion curves, with pass-band and stop-band regions of wave propagation. PC stop-band interval is known as the PC bandgap. LEDs require the PC bandgap at the light emission wavelengths to reduce the waves propagating in lateral directions relative to the LED axis. In this paper we briefly review the analysis of PCs and show the results of our simulation that conforms to results reported in the literature and finally discuss the efficiency enhancement of LED with application of a PC layer.

2. PHOTONIC CRYSTAL ANALYSIS

There are various analytical tools for the bandgap calculation of PCs. Here we use the plane-wave decomposition method as described in references 1–4. Briefly in a non-magnetic, source free region with a medium translationally invariant in z-direction and with the assumption that the wave propagation is confined to the transverse directions, we can write

$$\begin{aligned}
\nabla_t \cdot \mathbf{e}_t &= 0, \quad \nabla_t \cdot \mathbf{h}_t = 0 \\
\nabla_t e_z \times \mathbf{z} &= -j(\omega/c)(\mu_o/\varepsilon_o)^{1/2}\mathbf{h}_t \\
\nabla_t \cdot \mathbf{e}_t &= -j(\omega/c)(\mu_o/\varepsilon_o)^{1/2}h_z\mathbf{z} \\
\nabla_t h_z \times \mathbf{z} &= j(\omega/c)(\mu_o/\varepsilon_o)^{-1/2}\varepsilon_r\mathbf{e}_t \\
\nabla_t \times \mathbf{h}_t &= j(\omega/c)(\mu_o/\varepsilon_o)^{-1/2}\varepsilon_r e_z\mathbf{z}
\end{aligned}$$

where $\mathbf{e}$ and $\mathbf{h}$ are electric and magnetic fields and z and t refer to longitudinal and transverse direction respectively. ε_r is the relative permittivity of the material that varies periodically in transverse directions.

Eliminating electric field, the above equations reduce to two wave equations in the form of

$$\begin{aligned}
\nabla_t \times \eta\nabla_t h_z \times \mathbf{z} &= (\omega/c)^2 h_z\mathbf{z} \\
\nabla_t \times (\eta\nabla_t \times h_t \cdot \mathbf{z}) \times \mathbf{z} &= (\omega/c)^2\mathbf{h}_t
\end{aligned}$$

where $\eta = \varepsilon_r^{-1}$.

The magnetic fields are now expanded as summation of plane waves in the form of

$$h_z(\mathbf{t}) = \sum_{\mathbf{G}} h_{z\mathbf{G}} \exp[j(\mathbf{k}+\mathbf{G})\cdot\mathbf{t}]$$
$$h(\mathbf{t}) = \sum_{\mathbf{G}} h_{t\mathbf{G}} \exp[j(\mathbf{k}+\mathbf{G})\cdot\mathbf{t}]\mathbf{t}_{\mathbf{k}+\mathbf{G}}$$

where $\mathbf{k}$ is a wave vector in the Brillouin zone, $\mathbf{G}$ the reciprocal lattice vector and $\mathbf{t}_{\mathbf{k}+\mathbf{G}}$ a unit vector in $\mathbf{k}+\mathbf{G}$ direction. We can also expand η in the form of

$$\eta = \sum_{\mathbf{G}} \eta_{\mathbf{G}} \exp(j\mathbf{G}\cdot\mathbf{t}) \tag{1}$$

Substituting for h_z, $\mathbf{h}$ and η into the above wave equations we finally find

$$\sum_{\mathbf{G}'} \eta_{\mathbf{G}-\mathbf{G}'}(\mathbf{k}+\mathbf{G})\cdot(\mathbf{k}+\mathbf{G}')h_{z\mathbf{G}'} = (\omega/c)^2 h_{z\mathbf{G}} \tag{2}$$
$$\sum_{\mathbf{G}'} \eta_{\mathbf{G}-\mathbf{G}'} h_{t\mathbf{G}'}|(\mathbf{k}+\mathbf{G})||(\mathbf{k}+\mathbf{G}') = (\omega/c)^2 h_{t\mathbf{G}} \tag{3}$$

for H-mode with $h_t = 0$ and for E-mode with $h_z = 0$, respectively

3. HEXAGONAL LATTICE

In this section we consider the ω-k curve for a special case of hexagonal lattice dielectric structures. These structures consist of three groups of cylindrical rods imbedded in a dielectric substrate producing hexagonal lattice shapes. Each group of rods can be of different size or different substance with dielectric constant greater or less than that of the substrate. The schematic pattern of the lattice and the primitive lattice vectors are shown in Fig. 1.

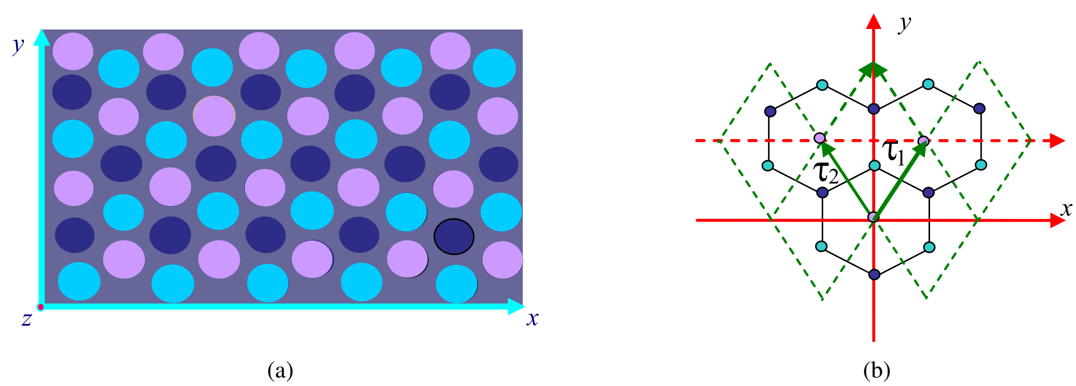
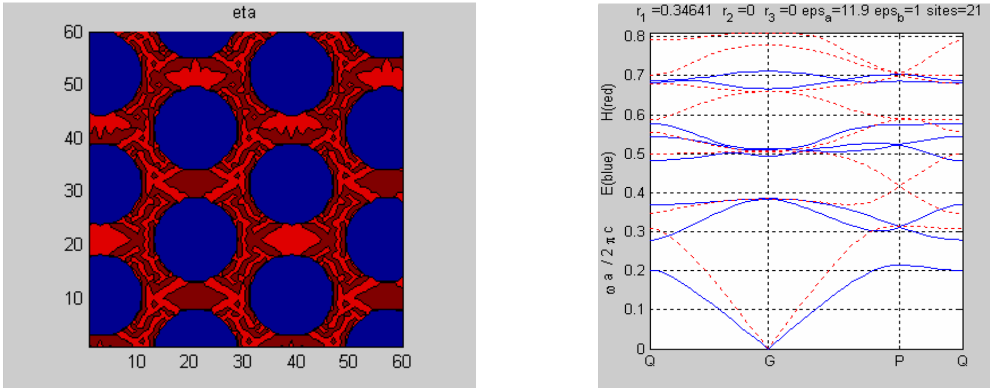


Figure 1: (a) Three different groups of dielectric rods placed in a dielectric substrate, (b) Primitive lattice vectors.

Application of Equations (1) and (2) for the above lattice structures would provide the appropriate eigenvalue equations, with eigenvalues representing the possible frequencies ω for each specific value of $\mathbf{k}$, and hence the required ω-k dispersion curves for both H and E modes.

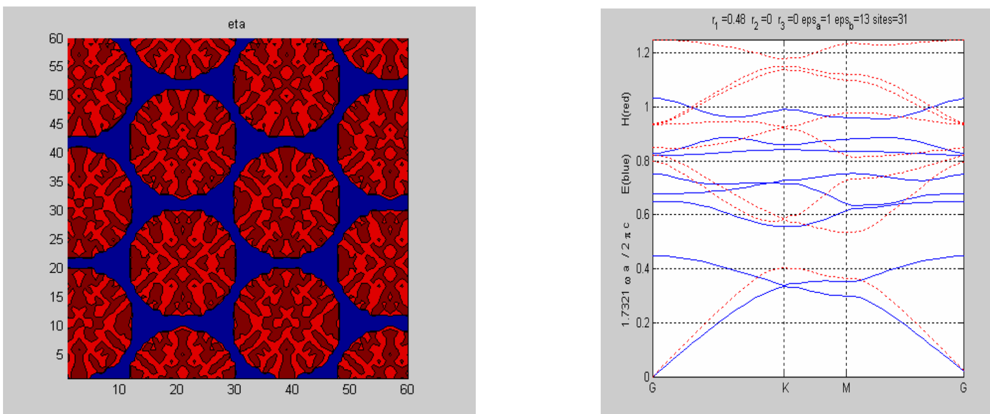
4. RESULTS AND DISCUSSIONS

As an example of identical dielectric rods in air, we assume that the relative permittivity of the dielectric rods to be 11.9 and the radius of the rods to be $0.346a$, where a is the lattice dimension. The result of simulation is shown in Fig. 2 and is well compared to the results given in [1]. In this case the photonic band gaps for H and E modes occur at different frequency intervals. Alternatively we assume air holes of radius $0.48a$ in a dielectric material of relative permittivity 13. The result of simulation is again given below and is very close to the results found in reference [3]. Here we can observe total bandgap for frequency intervals that neither H nor E modes can propagate. We have also considered the so called Graphite and Boron-nitrate lattice structures, with results well compared with that reported in [4].



(a) (b)

Figure 2: (a) Configuration of the PC, (b) the simulation result for identical dielectric rods in air.



(a) (b)

Figure 3: (a) Configuration of the PC, (b) the simulation result for identical air holes in a dielectric substrate.

5. EFFICIENCY ENHANCEMENT IN A LIGHT-EMITTING DIODE

In a standard LED light extraction efficiency is low because of the total internal reflection at semiconductor/air interface. To avoid this problem, two dimensional PC is etched into the upper cladding layer of LED. The structure is shown in Fig. 4 and the nano-pattern could be made by electron-beam lithography (EBL) or Anodic Aluminum Oxide (AAO) methods.

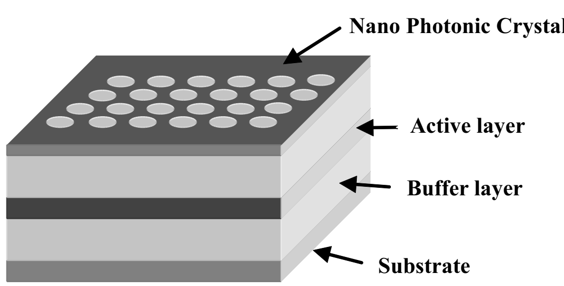


Figure 4: Schematic of LED with PC.

There are two equivalent ways to explain the improved light extraction efficiency of PC LED: one is based on photonic band gaps to prevent emission in guided modes [5]; in the other way, the PC is considered to couple guided modes to radiative modes [6].

In the first case we consider the existence of lateral guided modes in a slab impeding the extraction of light. A uniform dielectric slab that is surrounded by air carries at least one guided mode at all frequencies. However, by introducing two-dimensional PC, there could be a frequency range at which no guided mode would exist. In this frequency range, all the spontaneously emitted power will couple to radiation modes and exit out of the slab, resulting in the increased light extraction efficiency.

In the second way, coupling of guided modes into radiation modes is considered and it reduces the reflection at the PC interfaces. A shallow PC is used so that the vertical profiles of the guided and Bloch modes are matched. For a LED without PC, the rest of the emitted light lies below the light line of air, defined as $\omega = c|\mathbf{k}_0|$, where $\mathbf{k}_0$ is the in-plane wave vector in air and parallel to the surface of the LED structure. Modes below the light line ($|k| > |k_0|$), where $\mathbf{k}$ is the in-plane wave vector of the emitted light propagating in the active region, suffer from total internal reflection at the air/semiconductor interface and cannot couple to the radiation modes ($|k| < |k_0|$). However, the PC diffracts the guided modes at the Brillouin zone boundaries, allowing coupling to the radiation modes that lie above the light line. The diffraction condition is described as $\mathbf{k}_0 = \mathbf{k} \pm \mathbf{G}$, where $\mathbf{G}$ is a reciprocal lattice vector of the photonic crystal. Thus, the photonic crystal Bragg scatters the emitted light out of the active region, leading to higher extraction efficiencies.

Many researches showed that PCs could enhance efficiency of LED. T. Kim et al. compared the performances of the conventional LED and PC LED at 450 nm wavelength. The size of the mesa of a conventional LED is $40 \times 40\,\mu$m and actual lighting area is $20 \times 20\,\mu$m while a PC LED have hexagonal PC layer with 220 nm lattice distance and 110 nm diameter of holes. The external efficiency of the PC LED was increased about 35% compared to the conventional LED at the highest powers. The current voltage characteristic of the PC LED was also improved about 30% compared to the conventional one [7].

Another experiment reported by H. Hong et al. showed that enhanced light output was obtained from GaN-based near-UV light-emitting diodes by using PCs [8]. In the report, three samples were fabricated and compared: standard LED, one dimension (1D) and two dimensions (2D) Photonic Crystal LED at near UV wavelength. The standard LED structure comprised a GaN buffer layers, n-GaN, an InGaN/GaN multi-quantum well active layer and p-GaN. For 1D and 2D PC LED, a Cu-doped indium oxide (ICO) (3 nm) / indium tin oxide (ITO) (400 nm) scheme was used as a top p-contact layer due to its high transmittance and good ohmic behavior. 1D line-shaped and 2D hexagonal pattern were formed on ITO layer. According to the electroluminescence, the peak position remains unchanged and the 2D PC LED produce the highest intensity among three samples. The output-current (L-I) characteristics of the LEDs shows that the light output powers of the LEDs with 1D and 2D PC are much higher than those of the LED without PC. The light out put of 1D and 2D is improved by 33% and 48% compared with standard LEDs.

6. CONCLUSION

We have presented the basic properties of Photonic Crystals and explained the concept of photonic crystal bandgap. The planer decomposition method for the analysis of two-dimensional PC structures was explained and the analysis was applied to Hexagonal structures. Our computer simulations were extensively applied to different situations and the results were compared favorably with the results reported elsewhere in literature. Future work will consist of the calculation of enhanced efficiency of LED with a layer of PC. Experimental investigation of such devices is underway with a novel method of PC realization.

ACKNOWLEDGMENT

This work was supported by R01-2004-000-10830-0 from the Basic Research Program of the Korea Science & Engineering Foundation (KOSEF). The support of IDEC CAD tools and equipments, Dongshin University, Korea, is also gratefully acknowledged. Specially we would like to thank Professor Mike Myung-Ok Lee, School of Information and Communication Engineering, Dongshin University, Naju, Chonnam, Republic of Korea for facilitating this research work and for his continuous encouragement.

REFERENCES

1. Malkova, N., S. Kim, T. DiLazaro, and V Gopalan, "Symmetrical analysis of complex two-dimensional hexagonal photonic crystals," *Physical Review B*, Vol. 67, 2003.
2. Andreani, L. C. and M. Agio, "Photonic bands and gap maps in aphotonic crystal slab," *IEEE Journal of Quantum Electronics*, Vol. 38, No. 7, 2002.
3. Joannopoulos, J. D., R. D. Meade, and J. N. Winn, *Photonic Crystals*, Princeton University Press, 1995.

4. Cassagne, C., C. Jouanin, and D. Bertho, "Hexagonal photonic-band-gap structures," *Physical Review B*, Vol. 53, 1996.
5. Fan, S., P. R. Villeneuve, J. D. Joannopoulos, and E. F. Schubert, "High extraction efficiency of spontaneous emission from slabs of photonic crystals," *Phys. Rev. Lett.*, Vol. 78, 3294, 1997.
6. David, A., C. Meier, R. Sharma, F. S. Dinana, S. P. DenBaars, E. Hu, S. Nakamura, and C. Weisbuch, "Photonic bands in two-dimensionally patterned multimode GaN waveguides for light extraction," *Appl. Phys. Lett.*, Vol. 87, 101107, 2005.
7. Kim, T. A., J. Danner, and K. D. Choquette, "Enhancement in external quantum efficiency of blue light-emitting diode by photonic crystal surface grating," *Electron. Lett.*, Vol. 41, 1138, 2005.
8. Hong, H.-G., S.-S. Kim, D.-Y. Kim, T. Lee, J.-O. Song, J. H. Cho, C. Sone, Y. Park, and T.-Y. Seong, "Enhanced light output of GaN-based near_UV light-emitting diodes by using nanopatterned indium tin oxide electrodes," *Semicond. Sci. Technol.*, Vol. 21, 594, 2006.

Numerical Study on Two-dimensional Magnetic Photonic Crystals Made of Magnetized Ferrites

Jie Xu, Ping Chen, Yue Shi, Xin-Yi Ji, Ai-Min Jiang, and Rui-Xin Wu
Department of Electronic Science and Engineering, Nanjing University, Nanjing 210093, China

Abstract— Transmission characteristics of two-dimensional magnetic photonic crystals with square and hexagonal lattice have been studied by full wave electromagnetic simulation. The magnetic photonic crystals are composed of array of fully magnetized ferrite rods. Using FDTD methods, the transmission characteristics of the magnetic photonic crystals are simulated with different models of permeability. We find that simulation results using tensor permeability of ferrites differ from those of simulation with a scalar effective permeability of $\mathrm{TM_z}$ mode, which demonstrates that for magnetic photonic crystals anisotropy of ferrites cannot be simply modeled by an equivalent isotropic magnetic medium.

During the last few years, photonic crystals (PCs) [1, 2] or photonic band-gap (PBG) materials have been the subject of intensive theoretical and experimental research, due to their promising applications in microwave and optoelectronics [3–7]. Generally, PCs are artificially constructed from dielectric or metallic materials [7]. Recently, some papers are devoted to the investigation of magnetic photonic crystals (MPCs) [8–10]. In all these works, the anisotropy of permeability for the magnetic materials in MPCs is modeled by an effective permeability of $\mathrm{TM_z}$ mode.

Fully magnetized under the external static magnetic field ($\vec{H}_0 = \vec{e}_z H_0$), the ferrites have a tensor relative permeability in the following form [11, 12]:

$$\overleftrightarrow{\mu} = \begin{bmatrix} \mu_1 & j\mu_2 & 0 \\ -j\mu_2 & \mu_1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{1}$$

If a plane wave propagates in magnetized ferrites, the electric and magnetic fields follow the tensor form Maxwell's equation:

$$\nabla \times \vec{E} = -j\frac{\omega}{c} \overleftrightarrow{\mu} \cdot \vec{H} \tag{2}$$

$$\nabla \times \vec{H} = j\frac{\omega}{c}\varepsilon \vec{E} \tag{3}$$

Supposed the ferrites are uniform, for $\mathrm{TM_z}$ mode, the tensor form Maxwell equation can reduce into a scalar form with effective permeability

$$\mu_e = \frac{\mu_1^2 - \mu_2^2}{\mu_1} \tag{4}$$

However, if the ferrites are non-uniform, such as in MPCs, the scalar form Maxwell equation cannot be derived. In this situation, Maxwell equation is in the form of

$$\varepsilon^{-1} \cdot \nabla \times \left[\mu^{-1} \cdot \nabla \times \vec{E}\right] = \frac{\omega^2}{c^2}\vec{E} \tag{5}$$

where μ is a function of position. It is obviously that Eq. (5) cannot be reduced into a scalar form. Therefore, for magnetized MPCs, permeability of ferromagnetic materials cannot be simply modeled by an isotropic medium with the effective permeability μ_e. Because μ in Eq. (5) is a tensor, general plane wave expansion (PWE) method is not suitable. Therefore, in this study full wave analysis method FDTD is used.

We have simulated transmission coefficients of the MPCs by FDTD method in which the anisotropy of ferrites are molded with two different ways. The first is with tensor permeability Eq. (1), and the other is with effective permeability Eq. (4) for $\mathrm{TM_z}$ mode. Simulation results for square lattice are shown in Fig. 1, where we assume the intensity of applied static magnetic field is 120 Oe. Obviously, we can find that the transmission obtained with tensor permeability is

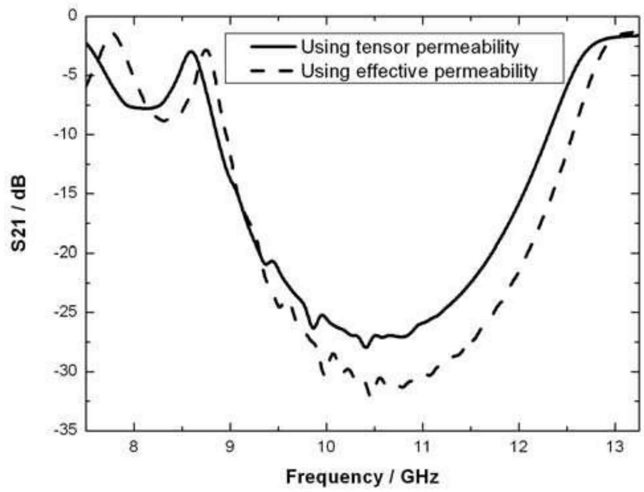


Figure 1: Calculated transmission spectrum of magnetic photonic crystals with square lattice using different models of permeability.

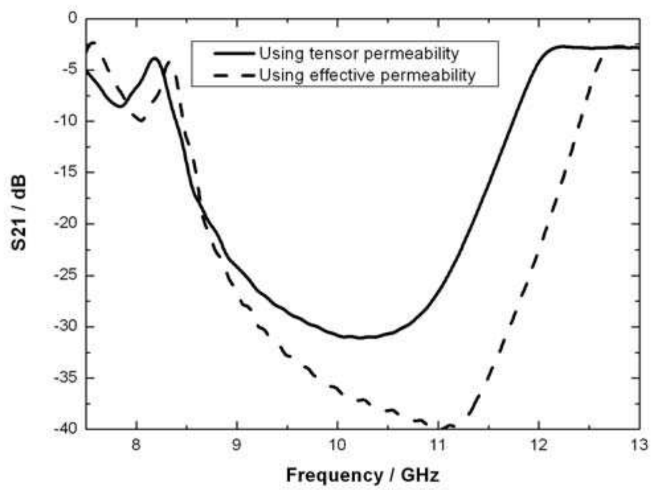


Figure 2: Calculated transmission spectrum of magnetic photonic crystals with hexagonal lattice using different models of permeability.

quite different from that of the anisotropy of permeability is model by an effective permeability. When effective permeability is used in calculation, the mid-gap frequency is higher and the depths of stop-band are deeper than those of using tensor permeability. Similar results are obtained for hexagonal lattice, as plotted in Fig. 2.

In conclusion, for MPCs anisotropy of ferries cannot be simply modeled by an effective permeability μ_e. The tensor permeability should be used in calculations to obtain correct results.

ACKNOWLEDGMENT

This work is supported by National Natural Science Foundation of China (No. 60471020).

REFERENCES

1. Yablonovitch, E., *Phys. Rev. Lett.*, Vol. 58, 2059, 1987.
2. John, S., *Phys. Rev. Lett.*, Vol. 58, 2486, 1987.
3. Radisic, V., I. Qian, et al., *IEEE Microwave Guided Wave Lett.*, Vol. 8, 67, 1998.
4. Yun, T.-Y. and K. Chang, *IEEE Trans. Microwave Theory Tech.*, Vol. 49, 549, 2001.
5. Thevenot, M., C. Cheype, et al., *IEEE Trans Microwave Theory Tech.*, Vol. 47, 2115, 1999.
6. Colburn, J. S. and Y. Rahmat-Samii, *IEEE Trans. Antennas Propagat.*, Vol. 47, 1785, 1999.

7. Lourtioz, J.-M., A. de Lustrac, *C. R. Physique*, Vol. 3, 79, 2002.
8. Sigalas, M. M., C. M. Soukoulis, R. Biswas, and K. M. Ho, *Phys. Rev. B*, Vol. 56, 959, 1997.
9. Kee, C. S., J. E. Kim, and H. Y. Park, *Phys. Rev. B*, Vol. 61, 15523, 2000.
10. Nagesh, E. D. V., V. Subramanian, et al., *Physica B*, Vol. 382, 45, 2006.
11. Liao, C. E., *Fundamentals of Microwave Techniques*, Xidian Press, Xi'an, 1994, (in Chinese).
12. Wu, R. X., X. K. Zhang, Z. F. Lin, S. T. Chui, and J. Q. Xiao, *J. Magn. Magn. Mat.*, Vol. 271, 180, 2004.

Electrically Tunable 2-D Organic Photonic Crystal Lasers

Rachel Jakubiak[1], **Timothy J. Bunning**[1], **Richard A. Vaia**[1], **Pamela Lloyd**[2]
Vincent P. Tondiglia[3], **Lalgudi V. Natarajan**[3], **and Richard L. Sutherland**[3]

[1]Materials & Manufacturing Directorate, Air Force Research Laboratory
Wright-Patterson AFB, OH 45433, USA
[2]UES, Inc., Beavercreek, OH 45432, USA
[3]SAIC, Beavercreek, OH 45432, USA

Abstract— Holographic polymerization of liquid crystal containing photopolymerizable resins enables one-step, rapid formation of multi-phase structures that exhibit partial photonic band gaps. These holographic polymer dispersed liquid crystals (H-PDLCs) provide a versatile platform for diffractive optical elements because the structures are not limited by multi-phase equilibrium but are controlled by the interference of multiple lasers at discrete angles. Incorporation of laser dyes into H-PDLCs form 1-D and 2-D optically pumped distributed feedback lasers. Linewidths as narrow as 1.8 nm are observed with laser thresholds below 1 mJ/cm^2 in 2-D columnar structures compared to 9 nm and 25 mJ/cm^2 exhibited by 1-D H-PDLC Bragg stack lasers. In the 2-D lattices the energy of the laser action can be tuned within the gain spectrum of the lasing medium by an applied electric field.

1. INTRODUCTION

The need for compact dispersive optical components for use in optical networking and other related technologies has been a driver for the development of photonic band gap structures (PBGs). In short, PBGs are structures that have spatially varying dielectric constants with a periodicity on the order of a wavelength of light. The purpose is to control the propagation of photons by the depletion of the number of photonic states for discrete optical frequencies, polarizations and angles[1] (band gaps). Commensurate with the decrease of photonic states is a dramatic increase in the density of photonic states at the band gap edges leading to local field enhancements that are valuable for applications requiring intense electromagnetic fields including nonlinear optics and stimulated emission.

Numerous materials and methods are available to make PBGs, that span the range of frequencies from microwave to ultraviolet. Organic materials lend themselves to facile fabrication and small feature sizes required for PBGs in the visible range, however they lack the high refractive indices of inorganics. Holographic patterning of PBGs, by both lithography [2] and photopolymerization [3] have proven to be viable methods supported by both experimental and theoretical investigations. Recent theoretical work describes laser beam geometries required to produce an infinite number of multi-dimensional structures that have the potential to have a complete band gaps, where no photon propagation is allowed for any angle of incidence or polarization for a given range of frequencies [4].

The lithographic process typically uses a photoresist where the unreacted components are washed away after holographic exposure. Photonic structures fabricated by this method generally have crisp interfaces, but are limited by material removal and capillary forces when small features or lamellar structures are needed. Holographic photopolymerization, namely holographic polymer dispersed liquid crystals (H-PDLCs), have the advantage of combining the desirable processing properties of polymers with an electro-optic medium built-in; albeit the microscopic features are not as clean as the lithographically processed samples. This drawback is not critical because the low index contrast of organics minimizes deleterious effects of rough interfaces when averaged over a large number of structural periods.

The low index contrast of organics does impede the formation of complete band gaps but for applications that require local field enhancement, the formation of a complete band gap may not be important. Field localization is a direct result of slow group velocity that can be achieved by reflections with a band gap, or through the group velocity anomaly where certain electromagnetic eigenmodes are small, particularly in 2-D and 3-D structures [5].

We demonstrate how high local field enhancement is possible in 2-D photonic structures where the index contrast is less than 0.20 by fabricating optically pump organic lasers from H-PDLCs. These results have implications for the future development of organic photonic structures for non-linear optical applications.

2. EXPERIMENTAL

Please refer elsewhere [6] for detailed experimental procedures. Briefly, the H-PDLCs are made from a formulation that contains TL213 liquid crystal, a photo-initiator, co-initiator, chain extender, fatty acid, penta-acrylate monomer and pyrromethene 597 (PM597) (Exciton) (~1 mM). This formulation was placed between ITO (indium-tin-oxide) coated glass slides and then exposed to two pairs of laser beams for the 2-D columnar structures, creating a pattern of laser fringes within the formulation [7]. In regions of constructive interference the penta-acrylate monomer preferentially photopolymerizes leading to phase segregation between the highly cross-linked polymer and liquid crystal droplets. After exposure to the laser beams the sample is cured with a white source creating a floodlit portion of the sample that has undergone photopolymerization without forming a grating. Glass microrods placed in the pre-polymer syrup leads the H-PDLCs to be 10 microns thick.

The diffraction efficiency of the grating was determined through the ratio of the intensity of diffracted light to incident light using a He-Ne (632 nm) laser. SEM micrographs were collected with a Hitachi 900 S operating at 1 keV. The PM597 was excited with the doubled output (532 nm) of a Nd: YAG that had a repetition rate of 10 Hz and pulse duration of 5–8 ns. Photoluminescence (PL) was collected either transmitted through or reflected from the cell with an Ocean Optics CCD/spectrometer that had a resolution of 1 nm.

3. RESULTS AND DISCUSSION

Shown in Figure 1 is an SEM micrograph of a 2-D columnar H-PDLC structure. The 2-D structure consists of essentially two transmission gratings orthogonal to each other that produced columns of LC-rich regions arranged in a square lattice with a lattice constant of 400 nm, with diffraction efficiencies in the meridian and equatorial orientations of 25%. LC columns are perpendicular to the plane of the substrate and have a radius of around 80 nm. The index contrast between the polymer and LC was approximately 0.16.

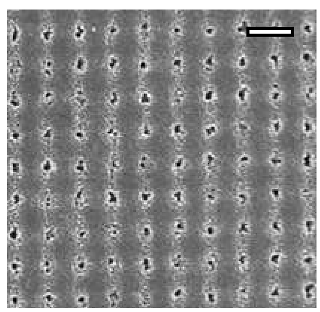


Figure 1: SEM micrographs of a H-PDLC 2-D columnar structure. Dark regions are voids where LC-droplets were prior to removal for imaging. The scale bar represents 500 nm.

The band diagram shown in Figure 2 is derived from the plane wave expansion method where the lattice consists of a square arrangement of rods that have radii 0.11 times the lattice constant and the material has a dielectric contrast of 0.395 ($\varepsilon = n^{1/2}$) corresponding to the H-PDLC. Flattening of the bands as depicted as shaded regions on the diagram, is observed at $\lambda/a = 1.48$ for the direction, corresponding to a laser oscillation of 600 nm for a lattice constant of 400 nm. A dominant cause of localization is a decrease of the group velocity (v_g) of the photons either by reflection within the stop band, such as the case with the 1-D transmission grating, or by band flattening as in the 2-D structures, where $\partial\omega/\partial k = v_g$. The resulting decrease in the group velocity by band flattening has been described in detailed by Sakoda [5] where it is called the group velocity anomaly and is particular to 2-D and 3-D structures, especially those that have low dielectric contrast.

Since the band diagram is generic for any square lattice with the requisite dielectric contrast and lattice constant, a lattice constant of 587 nm should provide lasing at 593 nm in the Γ-X direction. An example of this is shown in Figure 3 where the two PL spectra correspond to two different structures with lattices of 400 and 587 + / − 10 nm. The lasing wavelengths correspond well with what was predicted from the band diagram. Note that since the PM597 gain spectrum only has a bandwidth of 60 nm, it is not possible to observe both lasing modes simultaneously with the lattice constants we used. The broader (8.5 nm), high energy peak at 583 nm corresponds to

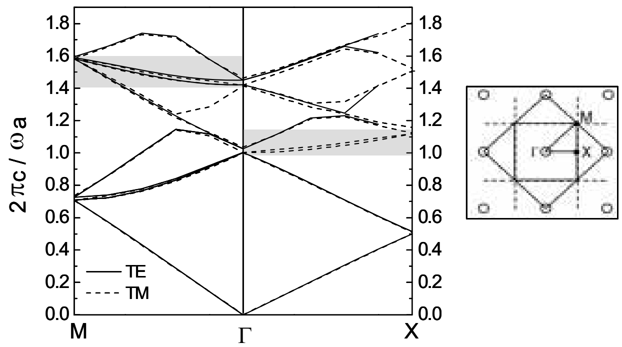


Figure 2: Band diagram for a square lattice of LC rods in a polyacrylate matrix ($\varepsilon = 0.347$, $r = 0.22a$). The frequencies are plotted against the boundaries of the irreducible Brillouin zone shown to the right, with both TE (solid lines) and TM (dashed lines) polarizations shown. Shaded areas on the band diagram highlight frequencies where anomalous group velocity occurs.

amplified spontaneous emission (ASE) that occurs without feedback and has no distinct threshold from the 2-D structure and coincides with the gain maximum of the PM597. At lower energy the resonance narrows to 1.9 nm and is a direct result of coherent feedback provided by the 2-D columnar structure. This laser oscillation has a pump threshold of 0.17 mJ cm^{-2} and occurs off the fluorescence maximum of the dye.

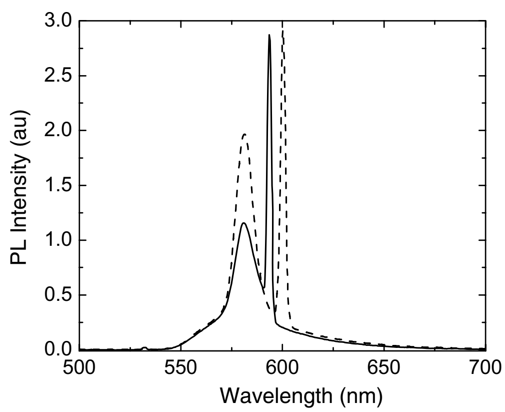


Figure 3: Laser action from PM597 embedded in 2-D columnar structures that had lattice constants of 587 nm (solid line) and 400 nm (dashed line). Pump fluence = 1.6 mJ cm^{-2}.

4. SUMMARY & CONCLUSIONS

We have demonstrated optically pumped laser action from a laser dye embedded in a 2-D H-PDLC photonic structure. The laser oscillations are well off resonance from the gain of the PL spectrum and correspond to frequencies related to the group velocity anomaly. In the future the local field enhancement offered by multidimensional structures with low dielectric contrast will be used in nonlinear optical applications.

ACKNOWLEDGMENT

We gratefully acknowledge financial support from the Air Force Office of Scientific Research, Materials and Manufacturing Directorate of the Air Force Research Laboratory and one of us (RJ) the National Research Council.

REFERENCES

1. Joannopoulos, J. D., P. R. Villeneuve, and S. Fan, "Photonic crystals: Putting a new twist on light," *Nature*, Vol. 386, 143–145, 1997.
2. Campbell, M., D. Sharp, M. T. Harrison, R. D. Denning, and A. J. Tuberfield, "Fabrication of photonic crystals for the visible spectrum by holographic lithography," *Nature*, Vol. 404, 53–56, 2000.

3. Tondiglia, V. P., L. V. Natarajan, R. L. Sutherland, D. Tomlin, and T. J. Bunning, "Holographic formation of electro-optical polymer-liquid crystal photonic crystals," *Adv. Mater.*, Vol. 14, 187–193, 2002.
4. Cai, L. Z., X. L. Yang, and Y. R. Wang, "Holography—all fourteen Bravais lattices can be formed by interference of four noncoplanar beams," *Opt. Lett.*, Vol. 27, 900–902, 2002.
5. Sakoda, K. "Enhanced light amplification due to group-velocity anomaly peculiar to two- and three-dimensional photonic crystals," *Optics Express*, Vol. 4, 167–176, 1999.
6. Bunning, T. J., L. V. Natarajan, V. P. Tondiglia, and R. L. Sutherland, "Holographic polymer dispersed liquid crystals (H-PDLC)," *Annu. Rev. Mater. Sci.*, Vol. 30, 83–115, 2000.
7. Jakubiak, R., L. V. Natarajan, L.,V. P. Tondiglia, P. Lloyd, R. L. Sutherland, T. J. Bunning, and R. A. Vaia, manuscript in preparation.

2D Nonlinear Photonic Crystals Nanocavities in Chalcogenide for All-optical Processing

Christian Grillet[1], **Cameron Smith**[1], **Snjezana Tomljenovic-Hanic**[1], **Eric C. Mägi**[1]
Yinlan Ruan[3], **Darren Freeman**[2], **Barry Luther-Davies**[2], **Steve Madden**[2], **Andrei Rode**[2]
M. Lee[4], **David Moss**[1], **and Benjamin J. Eggleton**[1]

[1]Centre for Ultrahigh bandwidth Devices for Optical Systems, School of Physics
University of Sydney, Sydney, Australia
[2]Centre for Ultrahigh bandwidth Devices for Optical Systems, Laser Physics Centre
Australian National University, Canberra, Australia
[3]Center of Expertise in Photonics, School of Chemistry and Physics
University of Adelaide, Adelaide, Australia
[4]ARC Centre of Excellence, Ultrahigh-bandwidth Devices for Optical Systems (CUDOS)
School of Physics A28, University of Sydney, Sydney, NSW 2006, Australia

Abstract— We demonstrate that high-Q nanocavity mode in 2D photonic crystal made in highly nonlinear chalcogenide glass can be excited using evanescent coupling from a tapered optical fibre. This scheme provides a promising platform to realize low power integrated all-optical switching and logic functions.

The key to all-optical processing is the ability to enhance nonlinear effects. One approach is based on developing optical configurations in which nonlinear effects in a weakly nonlinear material are enhanced by reducing the mode size and increasing the intensity of the light wave. 2D photonic crystal (PhC) membranes are expected to play a key role in this context. It has been shown that ultra small and high quality factor Q can be created by introducing a carefully designed defect in the PhC [1]. Following those advances, there have been predictions of the possibility of light switching light using nonlinear compact PhC microcavities operating at power levels of only a few mW [2, 3].

Chalcogenide glasses are attractive materials for all-optical signal processing. These glasses are composed of heavy elements including the chalcogens: S, Se and Te. The refractive index of chalcogenide is high, typically between 2.4 and 3.0 allowing 2D PhC slab to be created. Absorption losses are low over a wide wavelength range (near- to mid-infrared). Chalcogenide glasses possess a relatively large third-order optical nonlinearity (100–1000 times that of silica), and low two-photon absorption. In addition to reducing the switching power requirements, the pure Kerr-like nonlinearities offer the potential for near instantaneous response times ($< 100\,\mathrm{fs}$) and are only limited by the resonator Q-factor.

Using chalcogenide glasses to fabricate a 2D PhC nanocavity which confines light at the wavelength scale and thus enhances the nonlinear light-material interaction provides the essential ingredients for all-optical ultra-fast switching at low powers. It is well known that such cavities can exhibit optical bistability at incident powers that scale as the inverse square of the cavity's quality factor: $1/\mathrm{Q}^2$. It is also well known that the minimum fractional nonlinear change in the refractive index, $\delta\mathrm{n}/\mathrm{n}$, needed to operate the device has to be greater than the inverse of the cavity's quality factor $1/\mathrm{Q}$. A common characteristic with glasses is that an upper bound exists on the nonlinear index change, $\delta\mathrm{n}$, of a few 10^{-4}. This implies that a minimum Q-factor exists of about 5,000 in order to observe a nonlinear switching. In addition, a strong resonance depth is required to obtain high contrast between the two switching states. Using PhC nanocavities offer the advantage to strongly localize the mode in a small volume while limiting radiation losses through careful design (Fig. 1).

Efficient coupling of light into high-Q PhC cavities is a significant problem. One approach involves coupling light into the cavity via a tapered optical fiber [4–7] (Fig. 1). The challenge is to couple sufficient light into the cavity to induce the bistable behavior. To this end, we use a novel low-loss fiber taper to couple to the PhC nanocavity via evanescent coupling.

Photonic crystal nanocavities either manufactured using FIB milling [8, 9] or e-beam lithography plus chemically assisted ion beam etching (CAIBE) [10] were prepared for optical testing. Coupling to L3 (3 holes removed) resonators with different end-hole shift and diameters has been demonstrated. Predicted Q values obtained from 3D FDTD simulations predict intrinsic Q values greater

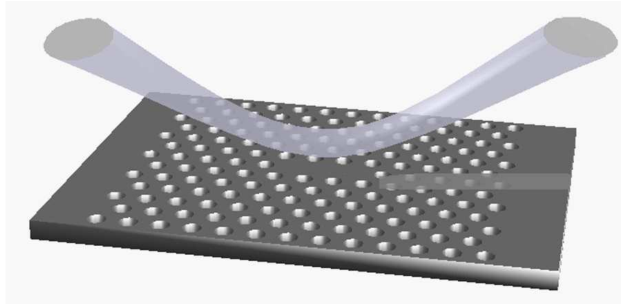


Figure 1: Coupling scheme used: schematic showing the coupling from a tapered fibre to PhC nanocavity.

than 10000 for optimal geometry. Fig. 2 shows experimental measurements performed on a cavity with both a side-hole shift and diameter reduction. Q value as high as 10,000 was measured for a separation of the fibre from the resonator of 800 nm. As this separation decreases, the measured Q factor also decreases down to 2000 and the depth of the transmission dip increases up to 1.5 dB.

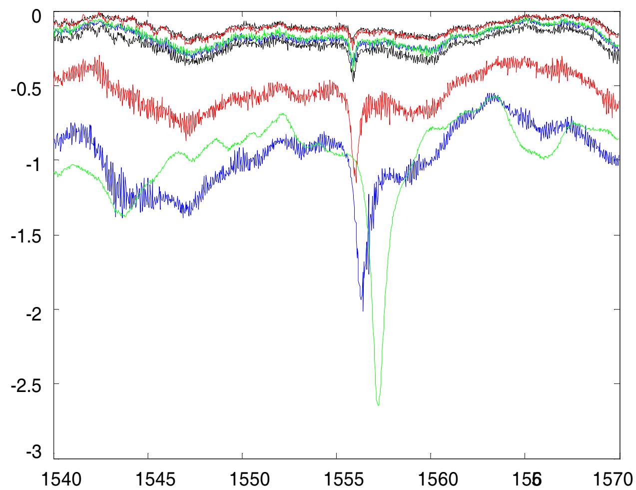


Figure 2: Transmission spectra through the tapered fibre for coupling to a modified L3-type nanocavity as a function of fibre to PhC separation.

These data indicate that, although the transmission depth is in this case restricted to a few dB limiting the contrast ratio of the expected switching device, simple chalcogenide PhC resonators can exhibit sufficiently high Q to make all-optical switching feasible. The main requirement for effective switching will be to increase the inherent resonator Q (before loading via fibre coupling) sufficiently so that the taper-cavity coupling system can exhibit high contrast response while maintaining the desirable funnelling of photons into the cavity. Experiments on all-optical switching are now underway.

ACKNOWLEDGMENT

The support of the Australian Research Council through its Federation Fellow and Centres of Excellence programs is gratefully acknowledged. The assistance of Myung-Ki Kim and Yong-Hee Lee from KAIST in the production of PhC samples via e-beam lithography and CAIBE is gratefully acknowledged.

REFERENCES

1. Akahane, Y., et al., "High-Q photonic nanocavity in a two-dimensional photonic crystal," *Nature*, Vol. 425, 944, 2003.
2. Centeno, E. and D. Felbacq, "Optical bistability in finite-size nonlinear bidimensional photonic crystals doped by a microcavity," *Phys. Rev. B*, Vol. 62, 7683–7686(R), 2000.
3. Soljacic, M., M. Ibanescu, S. G. Johnson, Y. Fink, and J. Joannopoulos, "Optimal bistable switching in nonlinear photonic crystals," *Phys. Rev. E*, Vol. 66, 055601(R), 2002.

4. Srinivasan, K., et al., "Optical-fiber-based measurement of an ultra-small volume high-Q photonic crystal microcavity," *Phys. Rev. B*, Vol. 70, 081306, 2004.
5. Barclay, P. E., et al., "Nonlinear response of silicon photonic crystal microresonators excited via an integrated waveguide and fiber taper," *Opt. Express*, Vol. 13, 801, 2005.
6. Hwang, I.-K., et al., "Optimization of coupling between photonic crystal resonator and curved microfiber," *IEEE Journal of Quantum Electronics*, Vol. 42, 131, 2006.
7. Grillet, C., et al., "Efficient coupling to chalcogenide glass photonic crystal waveguides via silica optical fiber nanowires," *Optics Express*, Vol. 14, 1071, 2006.
8. Freeman, D., et al., "Fabrication of planar photonic crystals in a chalcogenide glass using focused ion beam," *Opt. Express*, Vol. 13, 3079, 2005.
9. Grillet, C., et al., "Characterization and modeling of Fano resonances in chalcogenide photonic crystal membranes," *Optics Express*, Vol. 14, 369, 2006.
10. Ruan, Y., et al., "Fabrication of high-Q chalcogenide photonic crystal resonators by e-beam lithography," Submitted to *Applied Physics Letters*.

Observation of Whispering Gallery Resonances in Circular and Elliptical Semiconductor Pillar Microcavities

V. N. Astratov[1,2], S. Yang[1], S. Lam[2], D. Sanvitto[2], A. Tahraoui[2]
D. M. Whittaker[2], A. M. Fox[2], and M. S. Skolnick[2]

[1]Department of Physics and Optical Science, Center for Optoelectronics and Optical Communications
University of North Carolina at Charlotte, 9201 University City Blvd., Charlotte, NC 28223, USA
[2]Department of Physics and Astronomy, University of Sheffield, Sheffield S3 7RH, UK

Abstract— We observed whispering gallery resonances in semiconductor micropillars by employing geometry in which both excitation and collection of emission is in a direction normal to the sidewall surface of the pillars. The spectral positions of the peaks are found to be in a good agreement with the results of numerical modeling performed by finite difference time domain technique. The quality factors of whispering gallery modes ($Q \sim 20000$ for the 4–5 µm circular pillars) are found to be well in excess of Q-factors for the low k-vector "photonic dot" states measured from the same pillars. Due to high Q-factors and small modal volumes such whispering gallery resonances can be used in cavity quantum electrodynamics experiments.

Several groups demonstrated recently [1–3] strong coupling regime between cavity mode and individual quantum dot resonance in photonic microstructures. The design of the optical cavity for these experiments requires a combination of small modal volumes and ultrahigh quality (Q) factors of resonances. The advantages provided by whispering gallery modes (WGMs) for such experiments have been demonstrated [3] in the case of semiconductor microdisks. In this work we observed for the first time that semiconductor pillar microcavities with well defined "photonic dot" states [4, 5] simultaneously possess high-Q ($> 10^4$) WGM resonances. The combination of optical properties of such pillars is unique in terms of possible cavity quantum electrodynamics experiments since it allows observing coupling with dots with different spatial location inside the pillars. In addition we observed modes with $Q \sim 5000$ in very small pillars with extremely high ellipticity (2.5×1.5 µm size).

Our microcavity structure consists of 27 pairs of alternating AlAs/GaAs layers in the bottom distributed Bragg reflector and 20 repeats on top. The one-wavelength cavity contains one layer of InAs quantum dots of density $\sim 5 \times 10^9$ µm^{-2} positioned at the anti-node of the optical field. The quantum dots serve as an internal light source, enabling observation of the WGMs. The samples were processed into 1–10 µm diameter pillars using a combination of electron beam lithography and inductively coupled plasma etching. A scanning electron micrograph of a 4 µm circular pillar is shown in the inset of Fig. 1. The structure was cleaved to enable easy optical access to the sidewalls of the pillars. The distance between the pillars and the cleaved edge of structure was in the 1–5 µm range. The experiments were performed under high power conditions (~ 1 mW) where emission from single quantum dots is not resolved.

It is well documented that excitation and collection of PL emission in the vertical direction through the Bragg mirrors in such structures result in the observation of "photonic dot" states [4, 5] represented by the peaks in the blue spectra (labeled top to indicted detection through the top mirror) in Fig. 1. The formation of such states with nearly zero in-plane k-vector is a result of the coexistence of the strong vertical confinement produced by the Bragg mirrors with the additional lateral confinement in the pillars. It is also well known that in cylindrical and spherical cavities light can be trapped in high in-plane k-vector WGM states [6] with very high quality (Q) factors due to total internal reflection inside the resonator. In the present work we realised the first observation of such high Q states in semiconductor micropillars by employing geometry in which both excitation and collection of emission is in a direction normal to the sidewall surface, as shown schematically at the top of Fig. 1. In the emission spectra we observed a series of nearly equidistant peaks, a fingerprint of WGMs, illustrated by the red spectra in Fig. 1 and Fig. 2. The separation between the peaks was found to be inversely proportional to the diameter of the pillars as illustrated by the comparison of Fig. 1 and Fig. 2, taken for 4 and 5 µm pillars respectively. The Q-factors of WGMs were investigated using a 0.85 m double spectrometer and gave Q values of 20000 for the 4–5 µm circular pillars, well in excess of Q-factors (6-8000) for the low k-vector "photonic dot" states measured from the same pillars.

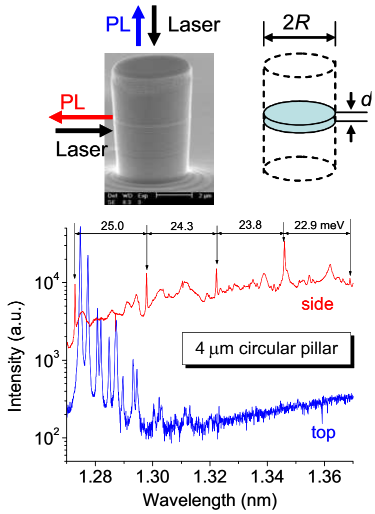


Figure 1: Circular 4 µm pillar microcavity: emission spectra detected from sidewall surface (red) and from the top mirror (blue).

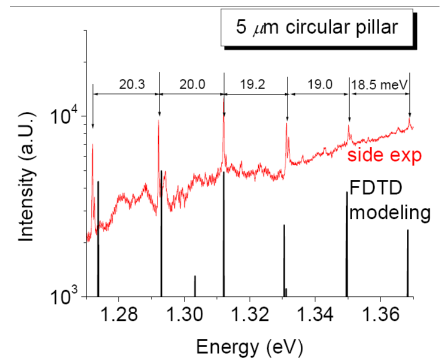


Figure 2: Circular 5 µm pillar microcavity: experimental WGM spectra detected from sidewall surface (red) and FDTD modeling.

High-Q WGMs (Q-values of 12,000) have been observed in semiconductor microdisk laser structures [7–9] in which high sidewall smoothness was obtained by optimisation of the wet etching process. In laser disk structures the sidewall confinement combines with the efficient vertical confinement at the semiconductor/air interfaces to produce the WGMs. In this respect the high Q-values of the WGMs observed in our work in microcavity pillars are surprising since the modulation of index between the one wavelength GaAs cavity regions and the AlAs/GaAs distributed top and bottom Bragg reflectors is much smaller compared to that at the air/semiconductor boundary in microdisks.

Numerical modeling of WGM effects was performed using three-dimensional finite difference time domain (FDTD) $FullWave^{\mathrm{TM}}$ software [10]. We used a simplified model of a disk (index 3.5) with the one-wavelength thickness. The source wave front was generated on the half a wavelength square plane inside the disk placed perpendicular to its surface at small depth to effectively excite WGMs. The discretization grid parameter was equal to 1/16 of the center pulse wavelength $\lambda = 940\,\mathrm{nm}$ in each dimension. We used a femtosecond transverse electric (TE) polarized built-in source to generate a comb of WGM resonances with various radial (n) and angular (l) mode numbers. To calculate the spectra of WGMs we used a Fourier transform of the electrical field averaged over several points inside the disk in the vicinity to its surface.

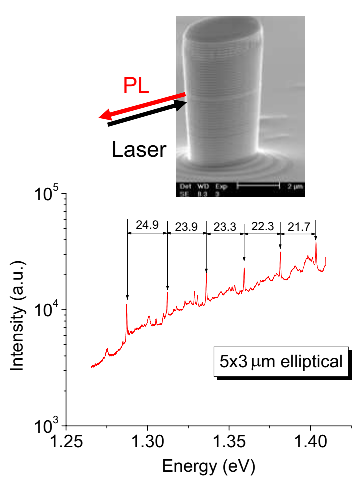


Figure 3: Elliptical $5 \times 3\,\mu\mathrm{m}$ pillar microcavity: emission spectra detected from sidewall surface. Inset shows an SEM image of a pillar.

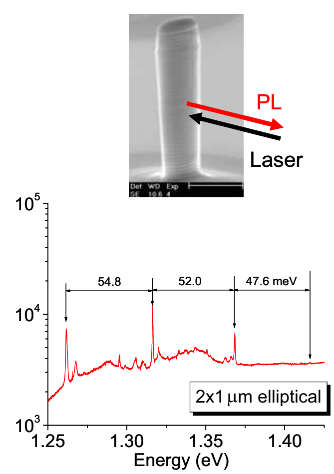


Figure 4: Elliptical $2.5 \times 1.5\,\mu\mathrm{m}$ pillar microcavity: emission detected from sidewall surface. WGM resonances are indicated.

As illustrated in Fig. 2 for 5 μm disk the calculations predict nearly equal separation between sequential modes with the same radial number ($n = 1$) and same (TE) polarization. This separation (18.7 meV) is found to be in very good agreement with the experimentally observed separations between the WGMs peaks. Close inspection of the experimentally observed mode energies in Fig. 1 and Fig. 2 shows however that the separations between WGMs are increased at lower energy. A full theoretical description of the field and energy distribution of the WGMs requires the effects of penetration of the electromagnetic field into the mirrors as well as the effects of absorption introduced by the quantum dots to be taken into account. These points will be addressed by the more accurate modeling in progress.

Asymmetric WG resonators are expected to be paticularly interesting: new phenomena such as chaos-assisted tunneling and dynamical localization [11] have been theoretically predicted. Micrographs of 5×3 μm and 2.5×1.5 μm elliptical pillars are shown in the insets of Fig. 3 and Fig. 4 respectively. Spectra measured from the sidewall surfaces are presented in Fig. 3 and Fig. 4. They illustrate deterministic rather than chaotic WGM resonances.

In conclusion, due to the small modal volume and high Q-factors (up to 20,000 at 4 μm circular and 5000 in e. g., the 2.5×1.5 μm micropillar in Fig. 4), as well as due to their intrinsic interest, the WGMs observed here have potential for future experiments on the observation of strong coupling with individual quantum dots.

ACKNOWLEDGMENT

The authors thank EPSRC (EP/D076145/1 and GR/S76076/01), ARO (W911NF-05-1-0529) and NSF (CCF-0513179) for support. This work was supported in part by funds provided by The University of North Carolina at Charlotte.

REFERENCES

1. Reithmaier, J. P., G. Sek, A. Loffler, C. Hoffman, S. Kuhn, S. Reitzenstein, L. V. Keldysh, V. D. Kulakovskii, and A. Forchel, "Strong coupling in a single quantum dot — semiconductor microcavity system," *Nature*, Vol. 432, 197–200, 11 November, 2004.
2. Yoshie, T., A. Scherer, J. Hendrickson, G. Khitrova, H. M. Gibbs, G. Rupper, C. Ell, O. B. Shchekin, and D. G. Deppe, "Vacuum Rabi splitting with a single quantum dot in a photonic crystal nanocavity," *Nature*, Vol. 432, 200–203, 11 November, 2004.
3. Peter, E., P. Senellart, D. Martrou, A. Lemaitre, J. M. Gerard, and J. Bloch, "Exciton-photon strong-coupling regime for a single quantum dot embedded in a microcavity," *Physical Review Letters*, Vol. 95, 067401, August 2005.
4. Sanvitto, D., A. Daraei, A. Tahraoui, M. Hopkinson, P. W. Fry, D. M. Whittaker, and M. S. Skolnick, "Observation of ultrahigh quality factor in a semiconductor microcavity," *Applied Physics Letters*, Vol. 86, 191109, May 2005.
5. Daraei, A., A. Tahraoui, D. Sanvitto, J. A. Timpson, P. W. Fry, M. Hopkinson, P. S. S. Guimaraes, H. Vinck, D. M. Whittaker, M. S. Skolnick, and A. M. Fox, "Control of polarized single quantum dot emission in high-quality-factor microcavity pillar," *Applied Physics Letters*, Vol. 88, 051113, January 2006.
6. For a recent review see, Matsko, A. B. and V. S. Ilchenko, "Optical resonators with whispering-gallery modes — Part 1: Basics," *IEEE Journal of Selected Topics in Quantum Electronics*, Vol. 12, No. 1, 3–14, January/February 2006.
7. McCall, S. L., A. F. J. Levi, R. E. Slusher, S. J. Pearton, and R. A. Logan, "Whispering-gallery mode microdisk laser," *Applied Physics Letters*, Vol. 60, No. 3, 289–291, January 1992.
8. Gayral, B., J. M. Gerard, A. Lemaitre, C. Dupuis, L. Manin, and J. L. Pelouard, "High-Q wet-etched GaAs microdisks containing InAs quantum boxes," *Applied Physics Letters*, Vol. 75, No. 13, 1908–1910, September 1999.
9. Renner, J., L. Worschech, A. Forchel, S. Mahapatra, and K. Brunner, "Whispering gallery modes in high quality ZnSe/ZnMgSSe microdisks with CdSe quantum dots studied at room temperature," *Applied Physics Letters*, Vol. 89, 091105, August 2006.
10. *FullWave*$^{\mathrm{TM}}$, Rsoft Design Group Inc., U.R.L.: www.rsoftdesign.com.
11. Narimanov, E. E. and V. A. Podolskiy, "Chaos-assisted tunneling and dynamical localization in dielectric microdisk resonators," *IEEE Journal of Selected Topics in Quantum Electronics*, Vol. 12, No. 1, 40–51, January/February 2006.

Electromagnetically Induced Transparency in Photonic Crystal Cavities

Paul M. Alsing, **D. A. Cardimona**, **and Dan H. Huang**
Air Force Research Laboratory, Space Vehicles Directorate, USA

Abstract— We investigate electromagnetically-induced transparency (EIT) in a three-level system in the cavity quantum electrodynamics regime, in which the cavity is a defect mode within a photonic crystal (PhC) material. Using a dressed state formalism, we address the enhancement of the EIT effect in the three-level system in both the V and Λ configurations utilizing both an ordinary cavity and a PhC cavity, thereby demonstrating the similarities and differences between the two situations. We will elucidate the role of the infinite ladder of photon-dependent dark states on the absorption minimum on resonance and the width of the EIT window off resonance. We will discuss the effect of the PhC bandgap on the cavity EIT phenomenon, and compare with recent results for EIT systems embedded in PhC materials.

1. INTRODUCTION

Photonic bandgap (PBG) materials (also called photonic crystals, or PhCs) are to electromagnetic fields as semiconductor materials are to electrons. Just as semiconductor materials have forbidden energy bandgaps due to the periodic arrangement of many atoms, a material with periodic variations in the dielectric constant can lead to a gap in the frequencies of electromagnetic radiation allowed to exist and propagate in the structured material. The principle of the gap formation is the same in both: Bragg-like diffraction of the electron waves off the atoms in the semiconductor lattice and Bragg scattering of optical waves by the dielectric interfaces in the periodically-structured material. Periodic variations in the dielectric constant in one dimension (1D), e.g., planar layers, lead to a PBG for light incident on the material in that dimension. If the variations are in two dimensions (2D), e.g., holes punched in a planar material, then the PBG will manifest for light incident within the plane of the holes. A 3D PhC will consist of dielectric variations in all dimensions, and will produce a PBG for light incident from any direction.

Once a material with a PBG has been created, it is possible to introduce a defect into the material (by subtracting or adding dielectric material at one spot), analogous to acceptor or donor defects in semiconductors. Similar to how these semiconductor defects introduce energy levels into the forbidden bandgap, the defect in a PhC can be designed with any frequency in the PBG, thereby allowing electromagnetic radiation at that frequency to exist and propagate in the crystal. Because the photonic defect is surrounded by a material with a PBG, the field should experience little or no loss, similar to being in a high-Q cavity. Thus, defects within PBG crystals can be considered microcavities. These microcavities can have extremely long lifetimes, at the same time that they have very small modal volumes. These unique features enable a range of novel applications that take advantage of the emerging field of nonlinear photonic crystals [1]. PhC microcavities have properties that greatly enhance optical nonlinear effects, and show great promise for producing devices for all-optical signal processing.

With the advent of PhC microcavities, investigations of cavity quantum electrodynamics (QED) in these novel materials have begun. One of the major issues that needs to be dealt with in this new field of quantum optics in PhCs is the development of alternative models and approaches to describe the new phenomena that are anticipated when two- or three-level atoms or other quantum-confined structures (such as man-made 'atoms' or quantum dots) are placed in PhC microcavities [2].

Two- and three-level systems are of great importance in quantum optics as they represent the simplest systems in which optical transitions can occur — and though they are simplest, they are far from trivial. Three-level systems, in particular, show a plethora of interesting phenomena when in free-space. This makes them very interesting to put into cavities - or into defect-mode microcavities in PhCs. Three-level systems have two possible transition frequencies — each of which can be within or outside the bandgap, or within or outside the cavity mode.

2. QUANTUM INTERFERENCE IN THREE-LEVEL SYSTEMS

In quantum optics, a three-level system in the so-called "V-configuration" (in which there are two closely-spaced excited states each coupled to a common ground state) can exhibit quantum beats when the excited states are in a coherent superposition. It was in this system that a 'steady-state quantum beat,' or a field-induced transparency, was first predicted to occur [3]. The original research performed in atomic physics on this system basically discovered that you could tune an applied pump field in such a way as to trap population in the excited states. The coherent excitation of the two closely-spaced upper states creates a quantum mechanical interference between the 2-to-1 transition and the 3-to-1 transition, resulting in a net zero dipole moment for the system. The population then becomes trapped in the excited states, since with a zero dipole moment, the applied field can no longer interact with the system. The system then becomes "transparent" to the applied field.

A three-level system in the so-called "Λ-configuration" (in which there are two closely-spaced ground states each coupled to a common excited state) has been studied in free-space quantum optics for, among other things, lasing without inversion and a phenomenon similar to that described above, which is now called Electromagnetically-Induced Transparency, or EIT [4].

These systems have been investigated quite extensively in quantum-confined semiconductor heterostructures, both theoretically [5] and experimentally [6].

3. QUANTUM OPTICS IN PHOTONIC CRYSTALS

The use of a cavity to enhance EIT on the probe transition has been proposed in atomic vapor systems [7] with just a few resonator photons or even the vacuum, provided the atom (coupling transition) and the resonator are strongly coupled. Spontaneous and induced atomic decay in photonic band structures was described by Kofman et al., [8], Zhu et al., [9] theoretically investigated the manifestation of dynamic quantum beats in the V-configuration in the presence of a PBG. Bay et al., [10] and Paspalakis et al., [11] studied the Λ-configuration theoretically with the transitions outside of the bandgap of a PhC. Quang et al., [12] assumed the transitions for the Λ-configuration were within a defect mode inside a PhC. None of these studies investigated EIT with one of the transitions within a defect cavity mode and the other either within the PBG or just outside of the PBG. This is what we will present here.

We build on our work with a two-level atom in a cavity [13], in which we found a suppression in the fluorescence of the atom when excited by an external field incident from the side of the cavity, and on our previous work in quantum interference [3, 5]. Using density matrix equations with a quantized cavity field, we find it illuminating to work in the dressed-atom picture, where the dressed states are eigenvectors of the combined atom-field system.

ACKNOWLEDGMENT

This research was supported by the US Air Force Office of Scientific Research (AFOSR).

REFERENCES

1. Soljacic, M. and J. D. Joannopoulos, *Nature Materials*, Vol. 3, 211, 2004; Soljacic, M., E. Lidorikis, J. D. Joannopoulos, and L. V. Hau, *Appl. Phys. Lett.*, Vol. 86, 171101, 2005; Bermel, P., A. Rodriquez, S. G. Johnson, J. D. Joannopoulos, and M. Soljacic, "Single-photon all-optical switching using waveguide-cavity QED," to be published, 2006.
2. Lambropoulous, P., G. M. Nikolopoulos, T. R. Nielsen, and S. Bay, *Rep. Prog. Phys.*, Vol. 63, 455, 2000; Soljacic, M. and J. D. Joannopoulos, *Nature Materials*, Vol. 3, 211, 2004.
3. Cardimona, D. A., M. G. Raymer, and C. R. Stroud, Jr., *J. Phys. B*, Vol. 15, 55, 1982; Cardimona, D. A., M. P. Sharma, and M. A. Ortega, *J. Phys. B*, Vol. 22, 4029, 1989; Cardimona, D. A., *Phys. Rev. A*, Vol. 41, 5016, 1990; Cardimona, D. A. and M. P. Sharma, "Method and apparatus for field-induced transparency using laser radiation," patent number 5, 196, 097, 1993.
4. Harris, S. E., *Phys. Rev. Lett.*, Vol. 62, 1033, 1989; Harris, S. E., J. E. Field, and A. Imamoglu, *Phys. Rev. Lett.*, Vol. 64, 1107, 1990; Harris, S. E., *Phys. Rev. Lett.*, Vol. 72, 52, 1994; Fleischhauer, M., A. Imamoglu, and J. P. Marangos, *Rev. Mod. Phys.*, Vol. 77, 633, 2005.
5. Lee, D. S. and K. J. Malloy, *IEEE J. Quant. Elect.*, Vol. 30, 85, 1994; Zhao, Y., D. Huang, and C. Wu, *Opt. Lett.*, Vol. 19, 816, 1994; Huang, D. and D. A. Cardimona, *J. Opt. Soc. Am B*, Vol. 15, 1578, 1998; Huang, D. and D. A. Cardimona, *Phys. Rev. A*, Vol. 64, 013822, 2001; Cardimona, D. A. and D. Huang, *Phys. Rev. A*, Vol. 65, 033828, 2002; Alsing, P. M., D. Huang, D. A. Cardimona, and T. Apostolova, *Phys. Rev. A*, Vol. 68, 033804, 2003.
6. Serapiglia, G. B. et al., *Phys. Rev. Lett.*, Vol. 84, 1019, 2000; Phillips, M. and H. Wang, *Phys. Rev. Lett.*, Vol. 89, 186401, 2002.

7. Field, J. E., *Phys. Rev. A*, Vol. 47, 5064, 1993.
8. Kofman, A. G., G. Kurizki, and B. Sherman, *J. Mod. Opt.*, Vol. 41, 353, 1994.
9. Zhu, S. Y., H. Chen, and H. Huang, *Phys. Rev. Lett.*, Vol. 79, 205, 1997.
10. Bay, S., P. Lambropoulous, and K. Molmer, *Phys. Rev. A*, Vol. 55, 1485, 1997; *Phys. Rev. Lett.*, Vol. 79, 2654, ibid, 1997.
11. Paspalakis, E., N. J. Kylstra, and P. L. Knight, *Phys. Rev. A*, Vol. 60, R33, 1999.
12. Quang, T., M. Woldeyohannnes, S. John, and G. S. Agarwal, *Phys. Rev. Lett.*, Vol. 79, 5238, 1997.
13. Alsing, P. M., D. A. Cardimona, and H. J. Carmichael, *Phys. Rev. A*, Vol. 45, 1793, 1992.

Design of Photonic Crystal Resonant Cavity Using Overmoded Dielectric Photonic Band Gap Structures

Limei Qi, **Ziqiang Yang**, **Zheng Liang**, **Wenxin Liu**, **Yu Liu**, **and Xi Gao**

Institute of High Energy Electronic, University of Electronic Science and Technology of China
Sichuan, Chengdu 610054, China

Abstract— An overmoded photonic crystal resonant cavity with two dimensional dielectric lattice structures is proposed and simulated. The dominant mode is a higher-order TM_{03}-like at the frequency of 31.14 GHz, the fundamental mode and most other modes are not supported by the cavity. The structure would be potential for application in accelerator, gyrotron and klystron in Ka-band.

1. INTRODUCTION

The photonic crystal is attracting much attention, because it has photonic band gaps (PBG) where no propagating electromagnetic mode exists, and when a defect mode is introduced in a photonic crystal, local electromagnetic modes (defect modes) arise within the forbidden band gap, which acts as a high-Q cavity for such local modes [1]. A photonic crystal cavity utilizing a defect mode (TE_{041}-like mode) in a periodically distributed metal lattice was reported for a 140 GHz gyrotron oscillator [2]. The 17 GHz mental photonic crystal cavity which supports a TE_{010}-like mode and is applied to accelerator was proposed [3], the cold tests for metallic cavities at X- and Ka-bands were analyzed [4, 5], a dielectric photonic crystal cavity where TM_{02}-like mode at 17 GHz frequency is confined has been studied through theory and simulation [6]. To improve the power capacity of high-power microwave devices, an overmoded resonant cavity should be used and the unwanted oscillation would be suppressed [2].

In this paper, we design a potential overmoded photonic crystal resonant cavity with two dimensional dielectric structure instead of metallic lattice one, because the fundamental mode TM_{01} must exist in the photonic band gap (PBG) of mental photonic crystal, a dielectric cavity could avoid this problem. Since the arrays of square dielectric rods provide poor azimuthal symmetry [4], the triangular arrays lattice is considered in the structure.

2. PBG AND TRANSMITTANCE IN TRIANGULAR LATTICE OF DIELECTRIC RODS

Figure 1(a) depicts a two dimensional triangular lattice of a photonic crystal, where the rod radius is r and lattice constant is a. Fig. 1(b) depicts the reciprocal lattice in the wave vector k-plane. The hexagon in the k-plane is the first Brillouin zone of the lattice. The shaded triangle in Fig. 1(b) covers the meaningful values of k-vector at which waves propagating in the lattice should be calculated.

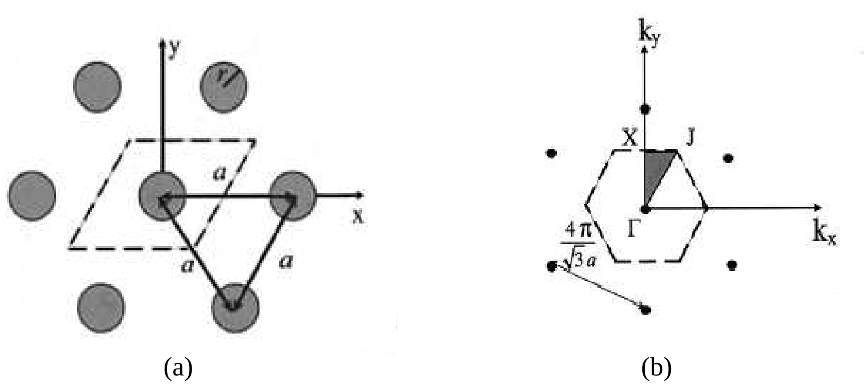


Figure 1: (a) Triangular lattice and (b) reciprocal lattice.

To design a valuable photonic cavity, one of the most important problems is to calculate the PBG. Several methods have been formed. The plane wave expansion method, the finite different time domain scheme, the transfer matrix method, and so on. In this paper, the plane wave expansion method by MATLAB software is used to calculate the PBG (Fig. 2(a)). The transmittance of the photonic crystal is presented in Fig. 2(b), which is simulated by CST software.

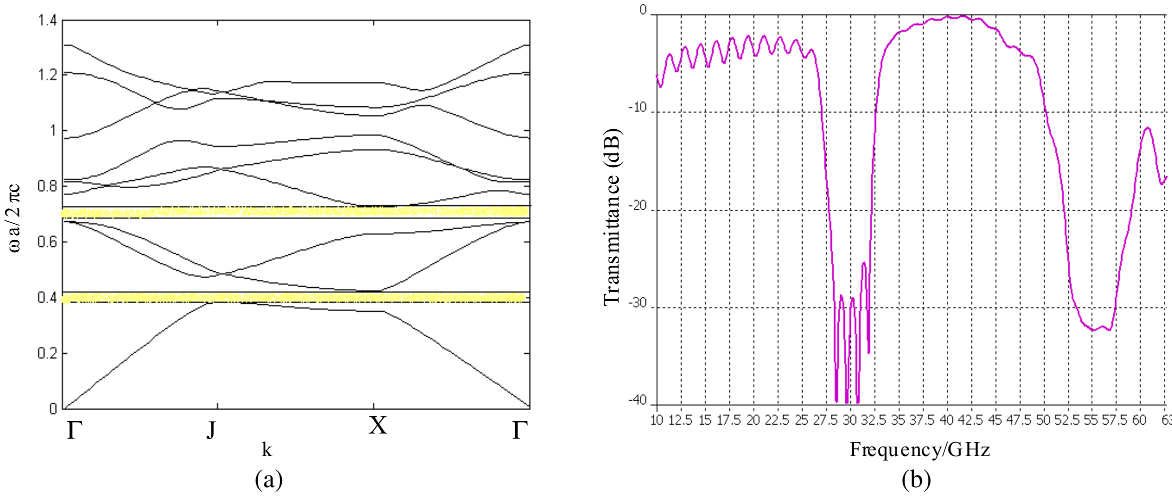


Figure 2: (a) Dispersion diagram for the dielectric triangular lattice rods by the plan wave method, and (b) transmittance of electromagnetic waves obtained by Computer Simulation Technology (CST).

The dispersion diagram indicates the frequencies of the waves propagating in the lattice while the k-vector on the reciprocal lattice varies from Γ-point to J-point, from J-point to X-point, and back to Γ-point. The dispersion diagram plots the normalized frequency $\omega a/2\pi c$, where a is the lattice constant and c is the speed of light. The diagram is calculated when r/a equals to 0.38 and a is 3.8 mm. The rods are the dielectric with dielectric constant $\varepsilon = 3.4$. The first gap and the second gap are marked by yellow shades. The first gap frequency is between 29.26 GHz and 33.53 GHz in which no waves can propagate, and the second PBG frequency is from 53.13 to 57.36 GHz, which is accordant to the transmittance in Fig. 2(b).

3. DESIGN OF PBG CAVITY

The photonic crystal cavity is formed by 180 rods in all with central 37 rods removed. The length of the cavity is chosen to be one second wavelength of the operating frequency (31.14 GHz) in order to optimize the efficiency. The High Frequency Structure Simulator (HFSS), a three dimension electromagnetic code is used to model the cavity. Through simulation, a well-confined TM_{03}-like mode at frequency of 31.14 GHz is found (Fig. 3(a)). The Q-factor calculated for dielectric loss tangent of 10^{-4} is 28499.

In the first PBG, the fundamental mode TM_{01} mode is forbidden, but TM_{21}-like mode and TM_{32}-like mode at 30.9 GHz and 33 GHz respectively appear, but they are less dangerous to the TM_{03} mode than the dipole modes, and their Q-factor are much lower than TM_{03}-like mode. In

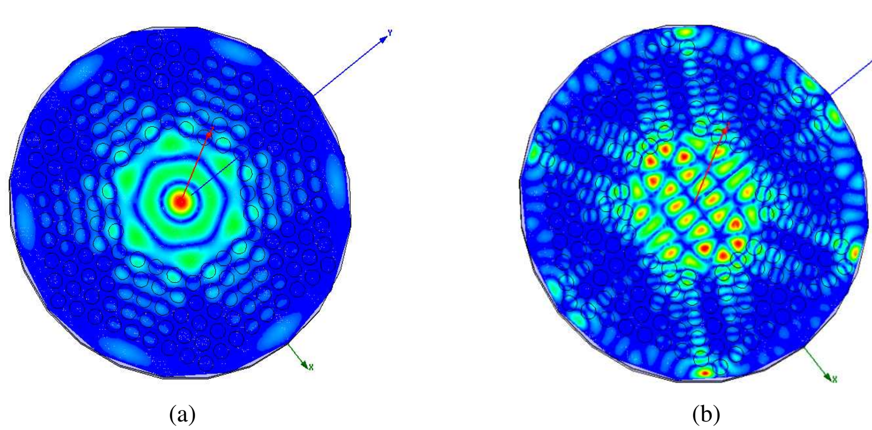


Figure 3: (a) Field pattern of the TM03-like mode at 31.14 GHz confined in the first PBG, and (b) the confined mode at 53 GHz in the second PBG.

the second PBG, a mode at the frequency of 53 GHz exist, it can't be easily excited in actual work state.

4. CONCLUSION

A defect in a photonic crystal would serve as an effective resonant cavity, it would only trap light in a very narrow frequency band and would hardly suffer any losses with high quality factor, In this paper, a 31.14 GHz photonic crystal cavity made of triangular dielectric rods has been designed with HFSS, it is demonstrated that the overmoded structure has a high Q-factor (28449), and the fundamental mode and most of other modes are forbidden, which is advantageous for application in accelerator, gyrotron and klystron in Ka-band.

ACKNOWLEDGMENT

This work is supported by National Nature Science Foundation of P. R. China (60571020).

REFERENCES

1. Joannopoulos, J. D., R. D. Meade, and J. N. Winn, *Photonic Crystals: Molding the Flow of Light*, Princeton Univ. Press, Princeton, 1995.
2. Sirigiri, J. R., K. E. Kreischer, J. Machuzak, and I. Mastovsky, "Photonic-band-gap resonator gyrotron," *Phys. Rew. Lett.*, Vol. 86, No. 24, 5628–5631, 2001.
3. Smirnova, E. I. and A. S. Kesar, "Demonstration of a 17 GHz, high-gradient accelerator with a photonic-band-gap structure," *Phys. Rew. Lett.*, Vol. 95, No. 7, 074801-1-4, 2005.
4. Shapiro, M. A., W. J. Brown, and I. Mastovsky, "17 GHz photonic band gap cavity with improved input coupling," *Phys. Rev. ST Accel. Beams*, Vol. 4, 042001-1-6, 2001.
5. Smirnova, E. I., M. A. Shapiro, C. Chen, and R. J. Temkin, "Photonic band gap structures for accelerator applications," *Advanced Accelerator Concepts, Mandalay Beach, AIP Con. Proc., CA.*, 383–393, December 2002.
6. Smirnova, E. I., C. Chen, and R. J. Temkin, "Theoretical analysis of overmoded dielectric photonic band gap structures for accelerator applications," *Proceeding of the 2003 Particle Acelerator Conference*, 1255–1257, Oregon, Portland, U.S.A., May 2003.

Finite Element Analysis of Photon Density of States for Two-dimensional Photonic Crystals with In-plane Light Propagation

M. C. Lin[1] **and R. F. Jao**[1,2]

[1]NanoScience Simulation Laboratory, Department of Physics, Fu Jen Catholic University
Taipei County 24205, Taiwan, China
[2]Department of Physics, Cheng Kung University, Tainan 701, Taiwan, China

Abstract— In-plane light propagation in two-dimensional (2D) photonic crystals (PCs) has been investigated by using an adaptive finite element method (FEM). Conventionally, the band structures of 2D PCs were calculated by either the plane-wave expansion method (PWEM) or the finite difference time domain method. Here, we solve the eigenvalue equations for the band structures of the 2D PCs using the adaptive FEM in real space. We have carefully examined the convergence of this approach for the desired accuracy and efficiency. The calculated results show some discrepancies when compared to the results calculated by the PWEM. This may due to the accuracy of the PWEM limited by the discontinuous nature of the dielectric functions. After acquiring the whole information of the dispersion relations within the irreducible Brillouin zone of the 2D PCs, the in-plane photon density of states can be calculated, accurately. These results are relevant to the spontaneous emission by an atom, or to dipole radiation in two-dimensional periodic structures.

1. INTRODUCTION

Many numerical methods have been developed and applied to the analysis and investigation of photonic crystals (PCs) including the plane-wave expansion method (PWEM) [1–7], the finite-difference frequency-domain/finite-difference time-domain (FDTD) method [8–10], the multiple-scattering method [11, 12], and the finite-element frequency-domain/finite-element time-domain method [13–20], and others. In spite of successful computations, there are several problems with Fourier-based methods. First of all, the dielectric function is discontinuous, so Fourier-type expansions converge slowly. Many precautions must be taken in order to ensure that the calculated spectra are correct. For instance, it was found that the discontinuous nature of the dielectric function severely limits the accuracy of the PWEM [7]. The other is the field localization due to the complexity of geometry. In addition, the difference relations in the FDTD method generate errors that are due to numerical dispersion, which limits the overall size of the structure. When the shape of the grid does not conform to the shape of a real material such as a circular rod, additional errors can be produced [21]. As experimental techniques to fabricate PCs are improved, a more rigorous method, which is free from these drawbacks, will be required. The finite-element method (FEM) has proved to be a flexible and efficient numerical tool with which to design various types of microwave components with inhomogeneous and complex structures [22].

In this work, in-plane light propagation in a two-dimensional (2D) photonic crystal is investigated by using adaptive finite element method. The polarization characteristics including both the transverse electric (TE) and transverse magnetic (TM) modes are considered in our simulation model. The finite-element method is employed to discretize the dielectric function profile of the PCs and the in-plane band structures are calculated by solving eigenvalue equations with proper periodic boundary conditions following the Bloch theorem [23]. We have carefully examined the convergence of this approach for the desired accuracy and efficiency. The calculated results show some discrepancies when compared to the results by the PWEM. Based on the finite-element analysis of the in-plane band structures of the 2D PCs in the irreducible Brillouin zone, the in-plane photon density of states (PDOS) of the 2D PCs for the TE and TM modes can be calculated accurately.

2. FORMULATION

2.1. In-plane Band Structure

In a 2D periodic system, the refractive index is a periodic function of x and y. We assume that the materials are linear, homogeneous, isotropic, lossless, and nonmagnetic. We have

$$\varepsilon_r(x,y) = \begin{cases} \varepsilon_a, & x,y \in \text{air region} \\ \varepsilon_d, & x,y \in \text{dielectric region} \end{cases} \tag{1}$$

where $\varepsilon_r(x,y)$ is the dielectric function profile, and ε_a and ε_d are the dielectric constants of the air and dielectric regions, respectively. For light propagating in the xy-plane, we can separate the modes into two independent polarizations, TM and TE modes, characterized by the field components parallel to the rods, $E_z(x,y)$ and $H_z(x,y)$, respectively. For in-plane propagation ($k_z = 0$), the Helmholtz's equations for the air (dielectric) region can be rearranged as:

$$\left[\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right] E_z(x,y) + \frac{\omega^2}{c^2}\varepsilon_{a(d)} E_z(x,y) = 0, \tag{2}$$

and

$$\left[\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right] H_z(x,y) + \frac{\omega^2}{c^2}\varepsilon_{a(d)} H_z(x,y) = 0. \tag{3}$$

As the system has discrete translational symmetry in the xy-plane, we have $\varepsilon(\vec{r}_{||}) = \varepsilon(\vec{r}_{||} + \vec{R})$, where $\vec{r}_{||}$ is the in-plane position vector and $\vec{R}$ is the linear combination of the primitive lattice vectors. By applying Bloch theorem, we can focus our attention on the values of the in-plane wave vector, $\vec{k}_{||} = k_x\hat{x} + k_y\hat{y}$, that are in the first Brillouin zone. We can calculate the in-plane band structures of 2D PCs by solving Eqs. (2) and (3) with the following boundary conditions,

$$E_{z,k_{||}}(\vec{r}_{||} + \vec{R}) = E_{z,k_{||}}(\vec{r}_{||}) \exp(i\vec{k}_{||} \cdot \vec{R}), \tag{4}$$

and

$$H_{z,k_{||}}(\vec{r}_{||} + \vec{R}) = H_{z,k_{||}}(\vec{r}_{||}) \exp(i\vec{k}_{||} \cdot \vec{R}). \tag{5}$$

By symmetry, we further restrict the value of k_x and k_y in the irreducible Brillouin zone. For illustrations, we consider a triangular lattice, as shown in Fig. 1. The cross-sectional views of a triangular array of air columns with a radius $r = 0.4297\,a$ drilled in a dielectric substrate of dielectric constant $\varepsilon_d = 11.9$, and the corresponding reciprocal lattice are shown in Fig. 1(a) and Fig. 1(b), respectively. The high-symmetry points at the corners of the irreducible Brillouin zone (1/12 of the first Brillouin zone, in shaded region) are Γ: $(0,\ 0)$, M: $(\pi/a(1,\ -1/\sqrt{3}))$, and K: $(\pi/a(4/3,\ 0))$. For the band diagram of the square lattice, the dimensionless frequency $(\omega/2\pi c)$ is calculated and plotted as a function of Bloch's vector $\vec{k}_{||}$ as the boundary $\Gamma MK\Gamma$ of the irreducible Brillouin zone.

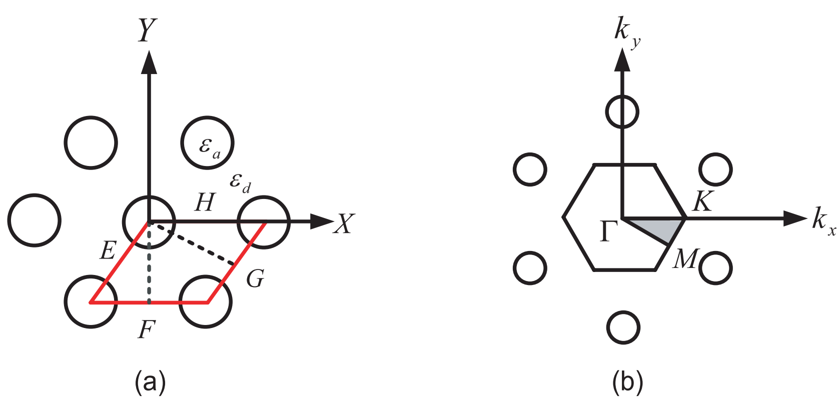


Figure 1: (a) A triangular lattice with unit cell (E, F, G, H) and (b) the corresponding reciprocal lattice with irreducible Brillouin zone (1/12 of the first Brillouin zone, in shaded region).

2.2. Finite Element Method

In any periodic structure only a single cell needs to be considered. The unit cell of the triangular lattice we chose is also shown in Fig. 1(a). With the information of the phase changes, the 2D Helmholtz's equations, Eqs. (2) and (3), and the corresponding equations of boundary conditions, Eqs. (4) and (5), form two sets of eigenvalue problems for the TM and TE polarizations, respectively. To solve the eigenvalue equations, we employ an eigensolver for partial differential equations based on an adaptive finite element method. For the adaptive FEM, a 2D, 6-node triangular finite element with a 5th-order basis function providing continuous derivatives between elements is used and the subdomains are partitioned into triangles, or mesh elements. The spatial domain will be discretized into smaller and smaller triangles for satisfying the Helmholtz's equations and the imposed boundary conditions until the desired accuracy is achieved.

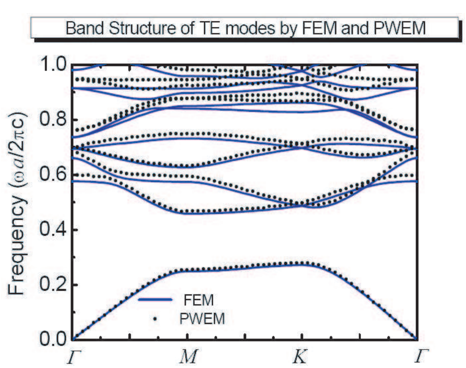
2.3. In-plane Photon Density of States

The dispersion relations of not only the boundaries but the interior of the irreducible Brillouin zone are needed for the calculation of photon density of states. For the dispersion relation of the triangular lattice, the dimensionless frequency is calculated as a function of Bloch's vector as in the interior of the triangle ΓMK of the irreducible Brillouin zone. Then, the dispersion relations can be plotted as 3D $(k_x - k_y - \omega)$ diagrams and it will be a surface for each band. To perform the in-plane PDOS calculation, by definition, we have $dN(\omega) \equiv D(\omega)d\omega$. Therefore, we have the expression for the PDOS as

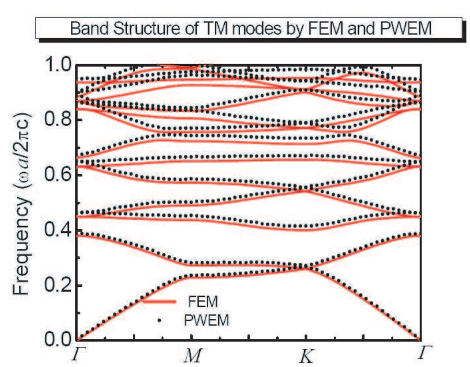
$$D(\omega) = \frac{S}{4\pi^2} \int_{\omega_k} \sqrt{1 + \left(\frac{\partial \omega}{\partial k_x}\right)^2 + \left(\frac{\partial \omega}{\partial k_y}\right)^2} dk_x dk_y \,. \qquad (6)$$

3. RESULTS AND DISCUSSION

We have carefully examined the convergence of the FEM approach for the desired accuracy and efficiency. The convergence tests of the first 15 bands of the triangular array have been done. The fractional errors of all eigenvalues are smaller than 10^{-4} when the mesh number increases to 16,340. Fig. 2 shows the calculated band structures for the TE and TM modes of the triangular lattice. A complete photonic band gap opens at $\omega a/2\pi c \sim 0.4$. The calculated results are quite similar to those in Ref. [6]. However, some discrepancies have been found after we carefully compared the data from our approach with those by the PWEM in Ref. [6]. Fig. 2 also shows the data calculated by the PWEM for comparison. The discrepancy is getting larger as the frequency is going to higher regime. This may be due to the accuracy of the PWEM limited by the discontinuous nature of the dielectric functions [7]. As the contrast of the dielectric constant is high, the step-like dielectric function is usually approximated by limited number of Fourier basis in the PWEM. Therefore, the convergence of the PWEM to realistic cases will be slow and sometimes not so good. After acquiring the whole information of the dispersion relations, the photon density of states can be calculated. The eigenfrequencies for 861 k-points uniformly distributed in the irreducible Brillouin zone have been calculated via the FEM approach. The in-plane PDOS for the TE and TM modes



(a)



(b)

Figure 2: Comparisons of band structure of (a) TE and (b) TM modes for a triangular array by FEM and PWEM [6].

has been calculated and plotted, as shown in Fig. 3. The results calculated by PWEM in Ref. [6] are also given in Fig. 3 for comparison. As one can see in the figure, the complete band gap, i.e., the overlap of the gaps of TE and TM modes, predicted by our method is smaller and shifts to lower frequency for the case.

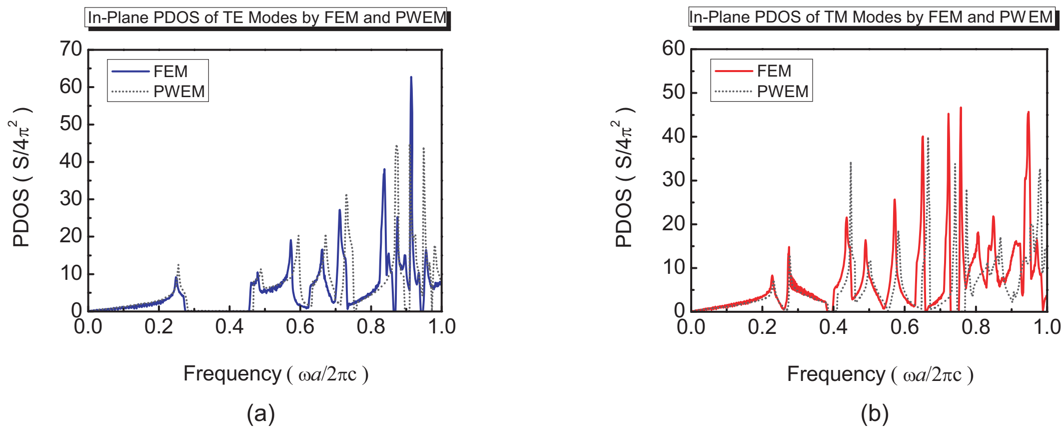


Figure 3: Comparisons of in-plane PDOS for (a) TE and (b) TM modes of a triangular array calculated by FEM and PWEM [6].

4. CONCLUSION

We have investigated the in-plane light propagation in 2D photonic crystals by using the adaptive finite element method in the frequency domain. Conventionally, the band structures of the 2D PCs were calculated by either the plane-wave expansion method or the finite difference time domain method. We have carefully examined the convergence of this approach for the desired accuracy and efficiency. The calculated results show some discrepancies when compared to the results calculated by the PWEM. This may be due to the accuracy of the PWEM limited by the discontinuous nature of the dielectric functions. Based on the finite-element analysis of the in-plane dispersion relations of the 2D PCs in the irreducible Brillouin zone, the in-plane photon density of states for both the TE and TM modes can be calculated accurately. These results are relevant to the spontaneous emission by an atom, or to dipole radiation in two-dimensional periodic structures.

ACKNOWLEDGMENT

This work was partially supported by the National Science Council, Taiwan, China, under Grant No. NSC 95-2112-M-030-002 and the National Center for High-performance Computing, Taiwan, China.

REFERENCES

1. Sakoda, K., *Phys. Rev. B*, Vol. 51, 4672–4675, 1995.
2. Sakoda, K., *Phys. Rev. B*, Vol. 52, 8992–9002, 1995.
3. Hillebrand, R., W. Hergert, and W. Harms, *Phys. Stat. Sol.* (b), Vol. 217, 981–989, 2000.
4. Plihal, M. and A. A. Maradudin, *Phys. Rev. B*, Vol. 44, 8565–8571, 1991.
5. Martin, O. J. F., C. Girard, D. R. Smith, and S. Schultz, *Phys. Rev. Lett.*, Vol. 82, 315–318, 1999.
6. Busch, K. and S. John, *Phys. Rev. E*, Vol. 58, 3896–3908, 1998.
7. Sözüer, H. S., J. W. Haus, and R. Inguva, *Phys. Rev. B*, Vol. 45, 13962–13972, 1992.
8. Ward, A. J. and J. B. Pendry, *Phys. Rev. B*, Vol. 58, 7252–7259, 1998.
9. Yang, H. Y. D., *IEEE Trans. Microwave Theory Tech.*, Vol. 44, No. 12, 2688–2695, 1996.
10. Chan, C. T., Q. L. Yu, and K. M. Ho, *Phys. Rev. B*, Vol. 51, 16635–16642, 1995.
11. Tayeb, G. and D. Maystre, *J. Opt. Soc. Am. A*, Vol. 14, No. 12, 3323–3332, 1997.
12. Zhang, W., C. T. Chan, and P. Sheng, *Opt. Express*, Vol. 8, No. 3, 203–208, 2001.
13. Hwang, J. K., S. B. Hyun, H. Y. Ryu, and Y. H. Lee, *Opt. Soc. Am. B*, Vol. 15, No. 8, 2316–2324, 1998.
14. Axmann, W. and P. Kuchment, *J. Comput. Phys.*, Vol. 150, 468–481, 1999.

15. Koshiba, M., Y. Tsuji, and M. Hikari, *J. Lightwave Tech.*, Vol. 18, No. 1, 102–110, 2000.
16. Pelosi, G., A. Cocchi, and A. Monorchio, *IEEE Trans. Antennas Propag.*, Vol. 48, No. 6, 973–980, 2000.
17. Dobson, D. C., J. Gopalakrishnan, and J. E. Pasciak, *J. Comput. Phys.*, Vol. 161, 668–679, 2000.
18. Dobson, D. C., *J. Comput. Phys.*, Vol. 149, 363–376, 1999.
19. Hiett, B. P., J. M. Generowicz, S. J. Cox, M. Molinari, D. H. Beckett, and K. S. Thomas, *IEE Proc.-Sci. Meas. Technal.*, Vol. 149, No. 5, 293–296, 2002.
20. Kim, W. J. and J. D. O'Brien, *J. Opt. Soc. Am. B*, Vol. 21, No. 2, 289–295, 2004.
21. Taflove, A., *Computational Electrodynamics: The Finite-difference Time-domain Method*, Artech, Boston, Mass., 1995.
22. Jin, J., *The Finite Element Method in Electromagnetics*, 2nd ed., Wiley-Interscience, New York, 2002.
23. Kittel, C., *Introduction to Solid State Physics*, Wiley, New York, 1976.

Combined Depth Migration and Constrained Inversion of Low Frequency Electromagnetic Data

Rune Mittet, **Frank Maaø**, **Odd M. Aakervik**, **and Svein Ellingsrud**
Electromagnetic Geoservices AS, Trondheim, Norway

Abstract— In order to migrate electromagnetic data from a low frequency controlled source, 3D electromagnetic Green functions should be used since the near-field effects may be large. Imaging principles of the correlation type do not have sufficient depth sensitivity to be used in a one-pass migration step. To increase the depth sensitivity, a non-local operator is introduced in the imaging condition. This operator accounts for the lateral propagation of the EM field in the high resistivity reservoir. The non-local operator depends on two parameters related to the resistivity and thickness of an assumed resistivity anomaly. We propose to estimate these parameters from a limited set of forward modeling operations.

1. INTRODUCTION

The Sea Bed Logging (SBL) method is described by Eidesmo et al. [1]. The main idea is to use an active electric source to probe the underground for thin, high resistive, layers. Hydrocarbon filled reservoirs will typically have a resistivity that is one to two orders of magnitude higher than its surroundings or a water filled reservoir. This is sufficient to support a partially guided wave in the reservoir that will leak energy up to receivers placed on the sea bed. The actual experiment is performed by dropping electric and magnetic sensors on the sea bed along a predetermined sail line and thereafter towing a horizontal electric dipole source along the line. It is well known that wave equation prestack depth migration of sea bed seismic data may be successful, given a good migration velocity model. Depth migration of sea bed EM data in a similar fashion is possible if the elastic wave equation is replaced by the Maxwell equations. However, additional problems must be addressed in depth migration of EM data.

First, the intermediate and high offset electromagnetic response from a hydrocarbon reservoir is not dominantly a reflection. The electromagnetic field excited in the reservoir behaves as a partially guided wave, propagating laterally through the reservoir and leaking energy back towards the receivers. The phase behavior as a function of offset is as for a refracted wave. Thus, Claerbouts imaging principle, which amounts to a correlation of up and down going energy in each subsurface location, is not ideal for imaging of hydrocarbon reservoirs with EM data.

Second, absorption and dispersion effects are much larger in EM data than in seismic data. Therefore, only low frequencies are available for imaging.

Third, the phase behavior of the electromagnetic field must be respected. In the far-field the electromagnetic field behaves as a "seismic wave" where phase increase linearly with propagation distance if the velocity is locally constant, however, for a typical overburden formation ($1\,\Omega$m to $3\,\Omega$m) and typical frequencies (0.25 Hz to 2 Hz) the near-field may reach several kilometers into the formation. For the near-field of a horizontal electric dipole in a conducting medium, the phase does not necessarily increase linearly with propagation distance even if the velocity is constant. The near-field is of course causal but appear to be nearly instantaneous for example in the depth direction. It is only in the far-field that the propagation velocity or phase gradient approaches that of a locally plane electromagnetic wave. To get the correct phase behavior of the fields, the Maxwell equations must be solved in 3D.

We recognize that Claerbouts imaging principle is not directly applicable and that the imaging principle should be modified to account for the partially guided wave in the reservoir. We do migration with full electromagnetic 3D Green functions that are calculated with a finite difference algorithm which solve for generally inhomogeneous media and also anisotropy if desired.

2. THEORY

In Mittet et al. [2], the elastic outgoing energy flux density of the misfit field was used as an error functional. The gradient of this error functional with respect to density and the Hooke's tensor could be expressed as correlations of an outgoing field with a reconstructed misfit field. The reconstructed misfit field was given by a Kirchhoff integral. This makes migration and the first iteration in an inversion procedure identical. Zhdanov and Portniaguine [3] have shown that a similar system can be obtained for the electromagnetic field using the electromagnetic energy flux density of the misfit field as the error functional,

$$\varepsilon = \int dt \int dS(\boldsymbol{x}_r) n_i \varepsilon_{ijk} \Delta E_j(\boldsymbol{x}_r, t) \Delta H_k(\boldsymbol{x}_r, t), \tag{1}$$

where n_i is the outward pointing surface normal, ε_{ijk} is the Levi-Civita tensor. The misfit field component $\Delta E_j(\boldsymbol{x}_r, t)$ is the difference between the measured electric field and the electric field predicted by the migration model at the receiver location,$\boldsymbol{x}_r$. The quantity $\Delta H_j(\boldsymbol{x}_r, t)$ is the corresponding magnetic misfit field and $dS(\boldsymbol{x}_r)$ is an infinitesimal receiver surface element.

The gradient for conductivity can be expressed as,

$$g_\sigma(\boldsymbol{x}) = \int dt E_m(\boldsymbol{x}, t) \Delta E_m(\boldsymbol{x}, t), \tag{2}$$

where $E_m(\boldsymbol{x}, t)$ is the outgoing field from the source, calculated in the background migration model. The gradient for resistivity is trivially obtained from the conductivity gradient. The first model update can be approximated to be in the negative gradient direction. In the following, the negative of the resistivity gradient is defined as the migrated image. If the migration results in a positive amplitude value at some location in the image, then a positive resistivity contrast is identified at that location.

Equation (2) is nothing but Claerbouts imaging principle, that is a correlation of an outgoing field with a field reconstructed from recorded boundary conditions. The parameter update in each iteration depends not only on the gradient, but also on the Hessian, which in principle is a non local operator on the gradient. Accounting for the Hessian is a non-trivial task and is not attempted here. The response from a hydrocarbon filled reservoir has a phase behavior with offset that is consistent with a partially guided or refracted event. The given gradient expression is formally correct but numerical tests have shown that it is not very sensitive to the reservoir depth. Thus, this imaging condition may perform poorly in a one-pass migration scheme. One way around this is to modify the imaging principle to include the effect of laterally propagating energy. Here, this is done by transforming the imaging condition in Equation (2) to the frequency domain and introducing a non local operator $\Phi(\boldsymbol{x}|\boldsymbol{x}', \omega)$,

$$I_\rho(\boldsymbol{x}) = \int dV(\boldsymbol{x}') \int d\omega \Phi(\boldsymbol{x}|\boldsymbol{x}', \omega) E_m(\boldsymbol{x}, \omega) \Delta E_m^*(\boldsymbol{x}', \omega). \tag{3}$$

$I_\rho(\boldsymbol{x})$ is the image with respect to resistivity contrasts.

It turns out that a simplified model can explain the main features of the SBL data for intermediate and large source receiver separations. We assume that the field propagate from the source down to the reservoir with a down going Green function, couples with a laterally propagating Green function in the reservoir which again couples with an upgoing Green function that take the response to the receiver. Thus, EM data with small source-receiver offsets are not used in this migration scheme. The laterally propagating Green function, $\Gamma(\boldsymbol{x}-\boldsymbol{x}', \omega)$, can be estimated with a plane layer modeling algorithm where the Green function is excited and recorded at reservoir depth. Thus,

$$\Phi(\boldsymbol{x}|\boldsymbol{x}', \omega) = \Phi(\boldsymbol{x} - \boldsymbol{x}', \omega) = \lambda \Gamma(\boldsymbol{x} - \boldsymbol{x}', \omega), \tag{4}$$

where λ is a (complex) coupling factor describing the field coupling in to, and out of, the thin high resistive layer. The operator $\Phi(\boldsymbol{x} - \boldsymbol{x}', \omega)$ depends on the transverse resistance of the anomaly,

$$R_t = (\rho_r - \rho_f) h_r, \tag{5}$$

where ρ_f is the formation resistivity, ρ_r and h_r the resistivity and thickness of an assumed resistivity anomaly. For a given set of ρ_r and h_r, the width and depth of the anomaly will be determined by the migration scheme.

Both phase and amplitude for the source current and the recorded EM data are used in the migration. Only phases for the Green functions are used. The total phase of the outgoing field from source to image location include the laterally propagating energy. The parameters ρ_r and/or h_r may in principle be unknown. These parameters are estimated by a limited set of forward modeling operations. Based on our experience up to present, we make the assumption that the lateral distance and width of the anomaly is given by the first migration step. Several resistivity models that include a reservoir are generated. The migration resistivity model is used as a basis. Reservoirs that vary in resistivity, thickness and depth are added to the basis model. A forward finite-difference simulation is performed for each of these models. An error based on the difference between the real data and the synthetic datasets is calculated in each case. The difference data with smallest errors point to the most probable models. A final remigration with the most probable ρ_r and h_r values is performed. The migrated depth and the most probable depth from the forward modeling is then compared. An inconsistency may point to an error in the background resistivity model.

3. RESULTS

Several different migrations have been performed with both synthetic and real data. In our real data example, using data from the Troll field, frequencies of 0.25 Hz, 0.75 Hz and 1.25 Hz were used. The reservoir at Troll is known to have a depth of 1400–1550 m. When using real data, ρ_r and h_r are in general not known. When making an initial guess, for instance $\rho_r = 100\,\Omega\text{m}$ and $h_r = 20\,\text{m}$, the reservoir was undermigrated at about 1100 m depth. A series of forward modeling operations with different ρ_r and h_r was then performed. The full set of difference data for all the forward modeling operations give a probability distribution that peak for a reservoir with thickness between 100 m and 200 m and magnitude between 100 Ωm and 200 Ωm. All these models give similar total errors and similar migrated images. By studying the error distribution the lowest error was found to be for the model with $\rho_r = 100\,\Omega\text{m}$, $h_r = 100\,\text{m}$. This model gives a migration result that predicts the depth to top reservoir of 1400 m.

REFERENCES

1. Eidesmo, T., S. Ellingsrud, L. M. MacGregor, S. Constable, M. C. Sinha, S. Johansen, and F. N. Kong, "Sea bed logging (sbl), a new method for remote and direct identification of hydrocarbon filled layers in deepwater areas," *First Break*, Vol. 20, 144–152. 2002.
2. Mittet, R., L. Amundsen, and B. Arntsen, "Iterative inversion/migration with complete boundary conditions for the residual misfit field," *J. Seis. Expl.*, Vol. 3, 141–156, 1994.
3. Zhdanov, M. S., and O. Portniaguine, "Time-domain electromagnetic migration in the solution of inverse problems," *Geophys. J. Int.*, Vol. 131, 293–309, 1997.

Study on Work Parameters of Seafloor Towed Survey Using Transient Electromagnetic Systems

Fengdao Zhou, Changsheng Liu, Jun Lin, Yan Wang, and Hui Li
Key Lab of Geo-Exploration and Instrumentation
Jilin University, Ministry of Education, Changchun, China

Abstract— Unlike the transient electromagnetic exploration on land, when transient electromagnetic system is used for seafloor exploration, the instrument systems including antennas can be towed behind the ship, over seafloor, and measurement may be taken during the sailing. Due to the move of antennas in the course of measure, some special requirements are brought forward for some work parameters. They are the frequency of transmitter wave, the number of stack, the speed of survey ship, and the separation between neighboring stations. The study indicates that the four parameters are associated with each other, and they affect on the survey results together. The value of each parameter can be decided according to the survey purpose and environment. However, because of the inter-limit of these parameters, an optimal scheme for survey task should be based on all of them together, not just on each one solely.

1. INTRODUCTION

Transient electromagnetic method has been widely applied to subsurface exploration on ground, and its theory and technology of field survey have been studied largely. When a station is detected, the antennas are still. Besides, it is also used in airborne geophysical exploration. Unlike exploration on ground, when transient electromagnetic systems are placed in aircraft, the measure is taken during the move of antennas. The difference of airborne and ground electromagnetic exploration has been analyzed by some authors [2, 3]. However, as a pity, the study on optimal work parameters such as the frequency of transmitter wave and the speed of aircraft in airborne exploration is not found. When transient electromagnetic method is applied to seafloor exploration, the measure can also be taken in a moving ship. So a same problem is met like airborne transient electromagnetic exploration, which is the determination of optimal work parameters in moving measure. Likewise, the method theory and systems of seafloor exploration have been studied [4–7], but the research on moving measure is not found. In this paper, we study the influence of four important parameters in towed survey work, which are the frequency of transmitter wave, the number of stack, the speed of survey ship, the separation between neighboring stations.

2. THE FREQUENCY OF TRANSMITTER WAVE

In transient electromagnetic exploration, the frequency of transmitter wave is usually not paid much attention to, because the investigation depth is mainly decided by the observation time and not directly decided by the frequency of transmitter wave in time-domain exploration. However, when the measurement is taken during the move of antennas, it becomes important. The frequency of transmitter wave decides the width of pulse and the stack number in given time. In order to have large number of stack of receiver signal in a certain time, the frequency is expected to be high. The high frequency leads to small pulse width and the influence of up edge become outstanding [1]. Figure 1 shows the influence of up edge on the transient response of conductive sphere. When the pulse width is much larger than the time constant of the sphere, the influence of up edge can be ignored. But with the decrease of pulse width, the influence of up edge becomes severe. Figure 2 is the observed transient response of conductive sphere in model experiments when different frequency of transmitter wave is used. The higher frequency is corresponding to small response due to the influence of up edge. Thus the transmitter frequency should not be too high.

Besides the influence on the transient response of target, the transmitter frequency also influences the noise level. The disturbance of 50 Hz in shallow sea around city is still strong because the electromagnetic waves of this frequency can not be attenuated completely in the seawater of small depth. If the transmitter frequency is not higher than this frequency, the disturbance of 50 Hz can be easily eliminated by the stack of bipolar responses.

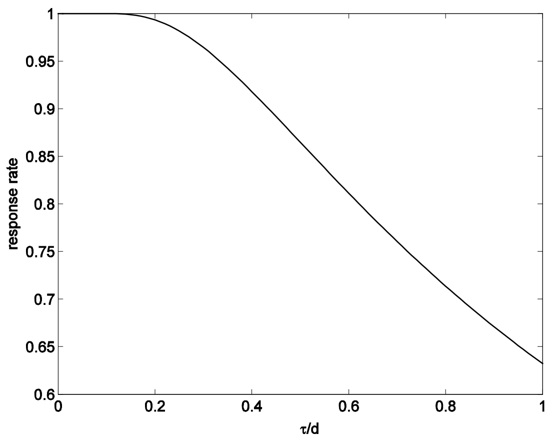


Figure 1: The influence degree curve of pulse width. The ordinate is the ratio of transient response affected by up edge vs that without the effect of up edge. The abscissa is the ratio of time constant of conductive sphere vs the width of pulse.

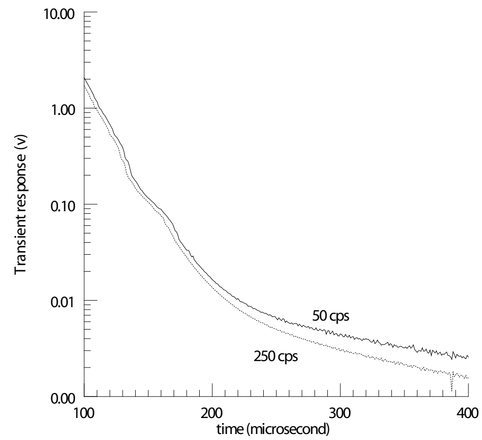


Figure 2: Transient responses of conductive sphere by different transmitter frequency.

3. THE NUMBER OF STACK

The technology of stack is used in transient electromagnetic exploration to suppress noise. With the increase of stack number, the noise in measure data after stack becomes small and the signal-noise ratio increase [8]. When the measure is done using immobile antennas, the stack number can be enough large, which just affects the work efficiency. However, in seafloor exploration, if the measure is done during the sailing of ship, the stack number can not be very large, especially when the speed of ship is high. The large number of stack means long measure time for a station, which will lead to an obvious average effect. Figure 3 shows the average effect of different stack number, in which the antennas are moving slowly during the measure and the diameter of conductive sphere located on 100 m is 15 cm. When increasing the stack number, the amplitude of transient response abnormity reduces. This is because the signal joining stacks on the record station where transient response is most strong is practically not all measured just on the station.

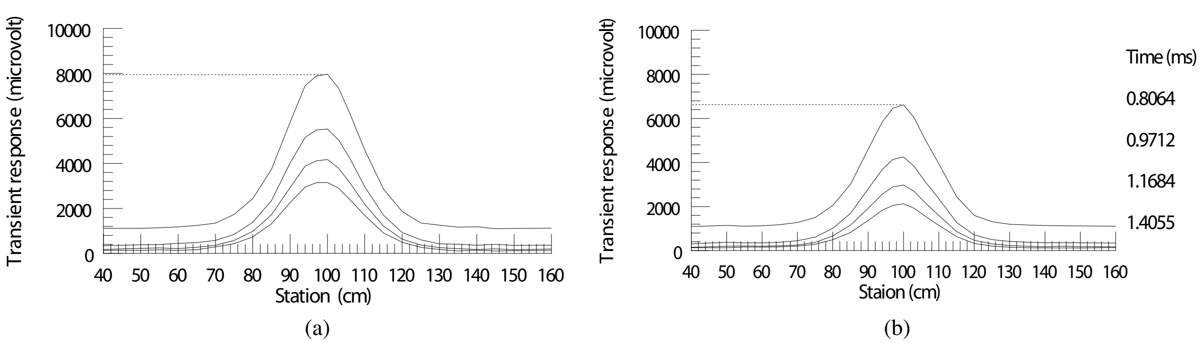


Figure 3: The transient response profile of conductive sphere surveyed by moving antennas. The stack number is respectively 64 in (a) and 512 in (b).

4. THE SPEED OF SURVEY SHIP

Survey in a moving ship will improve the work efficiency, and the higher speed brings the higher efficiency. But this is inconsistent with careful exploration. When small object is surveyed, antennas fast moving are unable to excite the object and receive its signal on the best station for enough time so that one station measure is finished, and the average effect is notable. Compared with Figure 4 which is the transient response measured with still antennas in model experiments, Figure 3 show the effect of moving antennas. In order to reduce the effect, one way is to use small stack number and high transmitter frequency. However, according to above mention the two parameters are restricted by some factors can not be any value. Therefore, the other way that is to limit the speed of antennas should be taken. The fit speed should be determined according to the size of object and considering the restriction of other parameters and work efficiency.

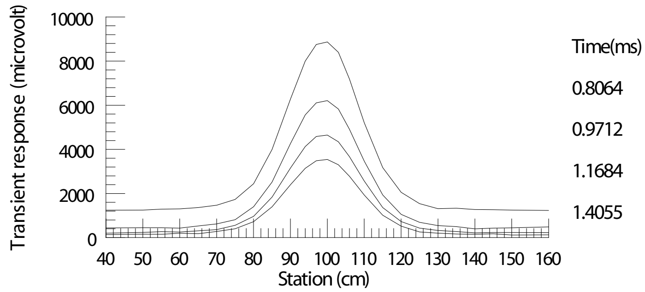


Figure 4: The transient response profile of conductive sphere surveyed by still antennas. The stack number is 64.

5. THE SEPARATION BETWEEN NEIGHBORING STATIONS

In towed survey, the measurement is allowed to go on without interruption and every position in survey line can be detected. But in the data record and interpretation, the measure data in a certain period of time is relegate to one point in survey line which is usually called record station. Since during the measure time of a record station the antennas have gone ahead for some distance, the separation between neighboring record stations can not be smaller than that. The smallest separation between neighboring stations is decided by the speed and the time spending in one station measure. When a seafloor survey is designed, the feasibility of station separation should be considered. If small station separation is required, the small speed of ship and little measure time for one station are necessary.

6. CONCLUSIONS

When towed survey is applied to seafloor exploration, the frequency of transmitter wave, the number of stack, the speed of survey ship, and the separation between neighboring stations are need to be designed carefully before survey. People are unable to achieve high speed of ship and thus high work efficiency for the detection of small body. We are also unable to reach for high signal-noise ratio by high frequency of transmitter wave and large number of stack in the moving measure, because of the influence of up edge. Besides, the average effect in towed survey is inescapable. When a big object in seafloor is detected, relatively high speed of ship, low frequency of transmitter wave, large number of stack, and large separation of neighboring stations can be used. And when a small object in seafloor is under exploration, low ship speed, high transmitter frequency, little stack number, and small stations separation are necessary.

ACKNOWLEDGMENT

This study is funded by 863 Program of China under grant 2002AA615030.

REFERENCES

1. Jiang, B., *Applied Near Zone Magnetic Source Transient EM Exploration*, 73–76, Geological Publishing House, Beijing, 1998.
2. Smith, R. S., A. P. Annan, and P. D. McGowan, "A comparison of data from airborne, semi-airborne, and ground electromagnetic systems," *Geophysics*, Vol. 66, No. 5, 1379–1385, 2001.
3. Christiansen, A. V. and N. B. Christensen, "A quantitative appraisal of airborne and ground-based transient electromagnetic (TEM) measurements in Denmark," *Geophysics*, Vol. 68, No. 2, 523–534, 2003.
4. Cheesman, S. J., R. N. Edwards, and A. D. Chave, "On the theory of sea-floor conductivity mapping using transient electromagnetic systems," *Geophysics*, Vol. 52, No. 2, 204–217, 1987.
5. Edwards, R. N., "Two-dimensional modeling of a towed in-line electric dipole-dipole sea-floor electromagnetic system: The optimum time delay or frequency for target resolution," *Geophysics*, Vol. 53, No. 6, 846–853, 1988.

6. Wang, Y., J. Wang, J. Wang, et al., "The research of conductivity on sea-floor by transient electromagnetic system (in Chinese with English abstract)," *Chinese Journal of Geophysics*, Vol. 41, No. 6, 841–847, 1998.
7. Yang, J. and R. N. Edwards, "Controlled source time-domain electromagnetic methods for sea-floor conductivity mapping (in Chinese with English abstract)," *The Chinese Journal of Nonferrous Metals*, Vol. 8, No. 4, 705–713, 1998.
8. Macnae, J. C., Y. Lamontagne, and G. F. West, "Noise processing techniques for time-domain EM systems," *Geophysics*, Vol. 49, No. 7, 934–948, 1984.

TDEM by FDEM

Wim A. Mulder[1,2] **and Evert C. Slob**[2]

[1] Shell International Exploration and Production B.V., P.O. Box 60, 2280AB Rijswijk, The Netherlands
[2] Delft University of Technology, Faculty of Civil Engineering and Geosciences
Department of Geotechnology, P.O. Box 5028, 2600GA Delft, The Netherlands

Abstract— Time-domain electromagnetic measurements of induction currents are useful for geophysical prospecting in shallow sea water and on land. We review the complexity of several numerical modelling schemes. A multigrid solver makes frequency-domain modelling followed by a Fourier transform an appealing choice. Examples are included.

Controlled-source electromagnetic measurements of induction currents in the earth can provide resistivity maps for geophysical prospecting. In marine environments, the current source often employs one or a few frequencies. In shallow sea water or on land, the response of air is dominant and time-domain measurements are more appropriate. Because electromagnetic signals in the earth are strongly diffusive, direct interpretation of measured data can be difficult. Inversion of the data for a resistivity model may provide better results. Therefore, an efficient modelling and inversion algorithm is required. In the frequency domain, the multigrid method allows for a reasonably fast solution of the discretized equations [1, 2]. On equidistant or mildly stretched grids, the number of iterations required to solve the equations at a given frequency is about 10, independent of the number of unknowns. Only with stronger stretching does the number of iterations increase. A more powerful method based on semi-coarsening and line-relaxation [3] is less sensitive to grid stretching but the required computer time per iteration is much larger. For time-domain modelling, there are a number of methods. The simplest is explicit time stepping, but this is rather costly. The Du Fort-Frankel method [4] is more efficient, but involves an artificial light speed term. Implicit methods are only efficient if a fast solver is available. Drushkin and Knizhnerman [5, 6] proposed a technique based on Lanczos reduction and matrix exponentials. Time-domain solutions can also be obtained from a frequency-domain code after a Fourier transform. An example for horizontally layered media can be found in, for instance, [7].

Here, the computational cost of these methods is compared by complexity analysis. This provides an estimate of the cost as a function of the number of unknowns, but without the actual constants. The next section compares the various methods. The frequency-domain method appears to be attractive. Examples that highlight some of the issues when using a frequency-domain method are included.

1. COMPLEXITY

Numerical modelling of transient EM signals can be performed by various methods. Here we consider an explicit time-stepping scheme, the Du Fort-Frankel method, implicit schemes, matrix exponentials and Lanczos reduction, and Fourier transformed frequency-domain solutions. Complexity analysis provides cost estimates in terms of the number of unknowns, without the constants that determine the actual computer run-time. The latter strongly depend on implementation and hardware details.

The method that is the simplest to implement is the explicit scheme. Unfortunately, this is only stable if the time step $\Delta t \leq ch^2$, where c is a constant depending on the material properties and the discretization, and h is the smallest grid spacing used in the problem. For a three-dimensional problem with n the number of grid points in each coordinate, the spacing $h = O(1/n)$ and the cost of a single time step is $O(n^3)$, so the overall complexity for computing the solution over a given time span T is $O(n^3T/\Delta t) = O(n^5)$. In practice, this is too slow for practical purposes, except perhaps on massively parallel computers.

The Du Fort-Frankel method [4] offers one way to get around the restrictive stability limit. An artificial light-speed is introduced with size $h/(\Delta t\sqrt{2})$ that allows the time step to grow with the square root of time, without doing too much harm to the accuracy of the solution. Geophysical applications of this method to TDEM problems can be found in, for instance, Oristaglio and Hohmann [8] for the 2D case and Commer and Newman [9] for 3D problems. The cost of the method is of $O(n^4)$.

An implicit scheme can avoid the $O(h^2)$ stability limit as well. The price paid is the solution of a large sparse linear system, which may be costly. An efficient iterative solution method for the frequency-domain equations [1–3] can also be used to solve the time-domain equations at an $O(n^3)$ cost per time step if the grid stretching is sufficiently mild. Together with a time step that scales with the square root of time, this method has the same complexity as the Du Fort-Frankel scheme, although the cost per step will be larger by at least an order of magnitude because of the work required by iterative solver. The method does not require an artificial light-speed term, which may allow for larger time steps without ruining the accuracy.

Drushkin and Knizhnerman [5, 6] proposed a technique that appears to be attractive for 3D applications. The Lanczos method was applied to reduce the original sparse matrix that describes the linear problem to a dense but much smaller one. This small matrix was used to quickly compute the time evolution using matrix exponentials.

The Lanczos method constructs the small matrix iteratively. Drushkin and Knizhnerman [5] show that accurate results can be obtained by performing m iterations, where $m = O(n\sqrt{T \log n})$. Here T is the length of time for which the solution needs to be computed, and n is the number of grid points in one of the spatial coordinates. Because the number of non-zero elements of A for a 3D problem is $O(n^3)$, the cost of the Lanczos decomposition will be of order $n^4\sqrt{\log n}$ for a given length of time T.

Time-domain solutions can be computed by first selecting a number of frequencies, then solving the frequency-domain problem at those frequencies, and finally performing an inverse Fourier transform to the time domain. For n_f frequencies and with an efficient solver that requires $O(1)$ iterations, the complexity is $O(n_f n^3)$, which can be favourable if n_f is small relative to n.

Comparison of the above methods shows that three of them have an asymptotic complexity of $O(n^4)$: the method based on Lanczos reduction, ignoring a logarithmic factor, the Du Fort-Frankel method, and an implicit scheme using an optimal solver with $O(1)$ complexity. The application of a frequency-domain method with the same solver results in an $O(n_f n^3)$ complexity, which may be better if n_f can be small relative to n.

These are only asymptotic results. In practice, the performance will depend on the details of the implementation and the actual constants in the complexity estimates.

The choice of grid is another topic. Diffusion problems have length scaling with the square-root of time. This implies that accurate modelling of a problem with a point-like source in space and time requires an initial grid that is very fine close to the source and gradually becomes less fine. Adaptive grid refinement will accomplish this, but leads to complicated software. Also, the Lanczos decomposition cannot be easily used with dynamic adaptive grid refinement. In the Fourier domain, the computational grid should depend on the skin depth and therefore on the frequency. Each frequency requires a different grid, but that is easier accomplished than time-dependent adaptive local grid refinement.

Although it remains to be seen which of the four methods requires the least computer time for a given accuracy, the frequency-domain approach appears to be quite attractive. Examples that highlight some of the issues in that approach will be presented next.

2. EXAMPLES

The Maxwell equations and Ohm's law for conducting media in the frequency domain can be written as

$$\imath\omega\mu_0\tilde{\sigma}\hat{\mathbf{E}} - \nabla \times \mu_r^{-1}\nabla \times \hat{\mathbf{E}} = -\imath\omega\mu_0\hat{\mathbf{J}}_s.$$

The vector $\hat{\mathbf{E}}(\omega, \mathbf{x})$ represents the electric field components as a function of angular frequency ω and position $\mathbf{x}$. The current source is $\hat{\mathbf{J}}_s(\omega, \mathbf{x})$. The quantity $\tilde{\sigma}(\mathbf{x}) = \sigma - \imath\omega\epsilon_0\epsilon_r$, with $\sigma(\mathbf{x})$ the conductivity, $\epsilon_r(\mathbf{x})$ the relative permittivity, $\mu_r(\mathbf{x})$ the relative permeability, and ϵ_0 and μ_0 their vacuum values. We use the Fourier convention $\mathbf{E}(t, \mathbf{x}) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\hat{\mathbf{E}}(\omega, \mathbf{x})e^{-\imath\omega t}d\omega$.

The first example is a point current source $J_s = \mathbf{j}_s\delta(\mathbf{x})\delta(t)$, $\mathbf{j}_s = (1, 0, 0)^{\mathsf{T}}$ A m, in a homogeneous formation with a conductivity of $\sigma = 1$ S/m. The frequency-domain solutions were computed on a grid which was adapted to the skin depth and finest near the source. For each frequency, the grid was different. Fig. 1 shows the real and imaginary parts of E_1, the x-component of the electric field. Its modulus is included. Frequencies were initially chosen as 10^q Hz, with q ranging from -2 to 2.25 with an increment of 0.25. To capture the variations, more frequencies were inserted between $q = -1$ and 1 at an increment of 0.125.

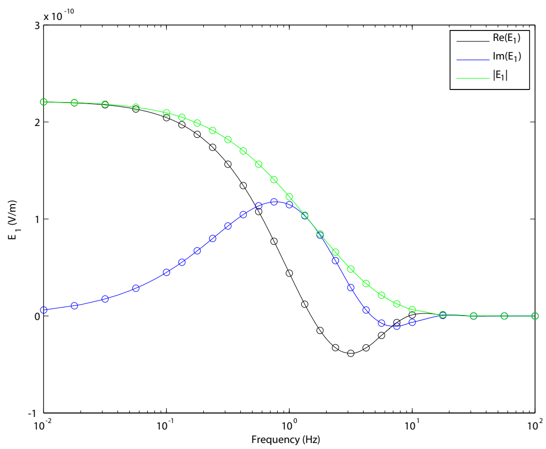


Figure 1: Real and imaginary part and the modulus of E_1 for various frequencies. The circles indicate the computed values, the lines were determined by shape-preserving piecewise cubic interpolation.

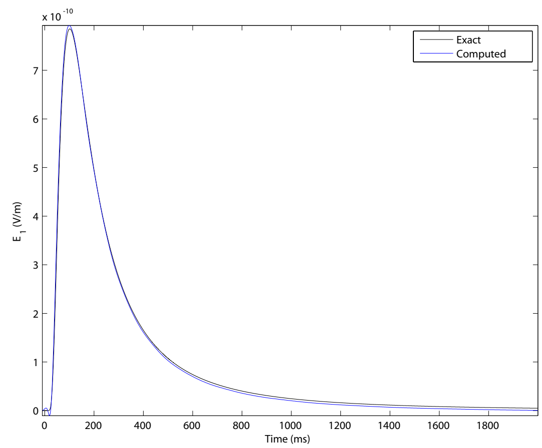


Figure 2: Time-domain solution for the homogeneous problem. Shown in blue is the horizontal component of the electric field. The exact solution is drawn in black.

The data points were interpolated by piecewise cubic Hermite interpolation [11] to an equidistant grid of frequencies and transformed to time. A comparison to the exact solution (see, e.g., [10]) is shown in Figs. 2 and 3. The error are largest at early and late times, due to lack of the lowest and highest frequencies.

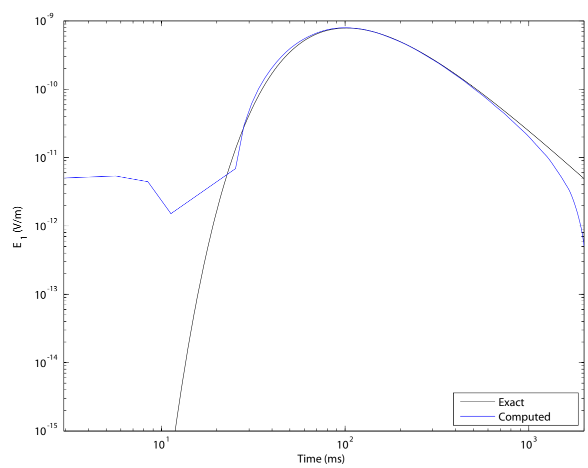


Figure 3: The same time-domain solution for the homogeneous problem as in the previous figure, but now on a logarithmic scale. Only the positive values are displayed.

The second example is a conductive scatterer in a homogeneous background with a conductivity of 1 S/m. The rectangular scatterer with $x \in (-300, 300)$ m, $y \in (-200, 200)$ m, and $z \in (400, 600)$ m has a conductivity of 10 S/m. The source is the same as in the previous example. In this case, we used a primary/secondary formulation in which the homogeneous response is subtracted so that the source term and its singular response is replaced by a source term that involves the exact solution [10].

The frequency-domain solution for a source at the origin and a receiver located at $(900, 0, 0)$ m and computed on grids with 128^3 cells is displayed in the left panel of Fig. 4. For comparison, we have computed the full electric field for the homogeneous medium with the scatterer and subtracted the numerical solution for the homogeneous medium without the scatterer. The difference is shown in the right panel of Fig. 4.

The time-domain response is shown in Fig. 5 next to the exact homogeneous solution.

In the examples, frequencies were selected on a logarithmic scale and more were added where needed. The grid was based on the skin depth and was different for each frequency. Solutions were obtained in 8 iterations for the higher frequencies. However, several hundreds of iterations were required at 0.01 Hz, where strong grid stretching was required to include the small scatterer and have an outer boundary sufficiently far away. This calls for a better solver, for instance, the one described in [3].

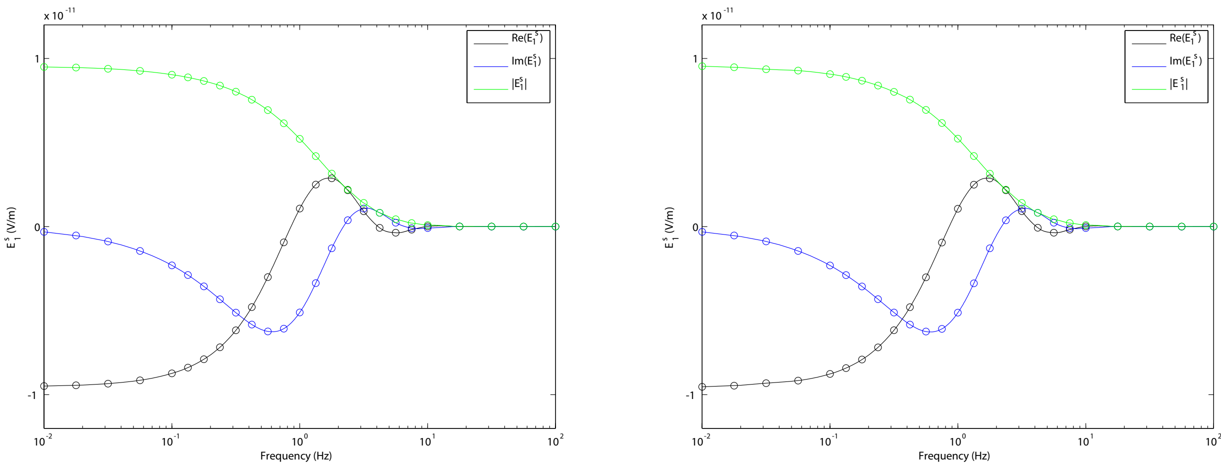


Figure 4: The left figure shows the secondary solution in the frequency domain computed for grids with 128^3 cells. At the right, solutions obtained by taking the difference between the full solutions with and without scatterer is plotted.

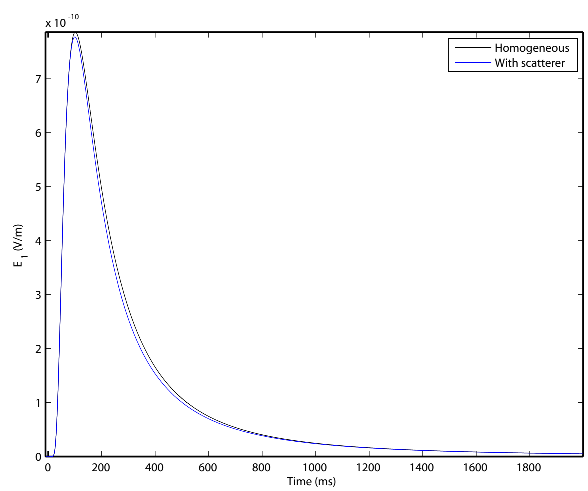


Figure 5: The time-domain solution for the scatterer computed with the primary/secondary formulation on a grid with 128^3 cells. The exact solution for the homogeneous case is shown in black for comparison.

A realistic geophysical model was considered in [2]. Here we use the same model, but without adding an extra 500 m of sea water, so the sea is fairly shallow. Fig. 6 displays the resistivity on a logarithmic scale. Initial solutions for computed at frequencies 10^q Hz, with q between -2.75 and 2.5 at a 0.25 increment. Next, frequencies were added by comparing receiver values at a given frequency to results obtained by interpolation without including that frequency. If the difference between the interpolated and actual value exceeded a tolerance, extra frequencies were added. This was repeated a few times. Cubic interpolation or extrapolation of solutions for other frequencies was use to obtain an initial guess for the iterative solution method. The spatial grid was again based on a balance between the skin depth at the given frequency and the details of the model. One of the time-domain solutions is shown in Fig. 7. The air-wave shows up as an early peak. The anti-causal part must be caused by missing high frequencies and numerical errors.

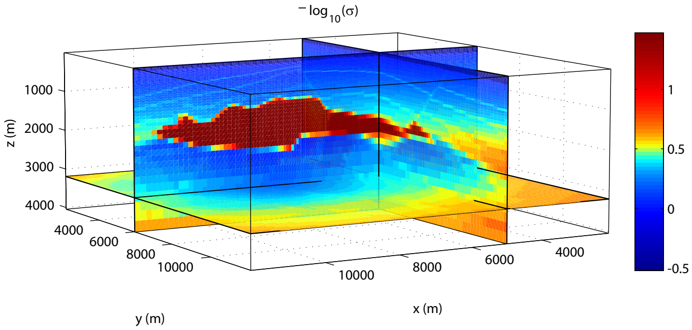


Figure 6: Resistivity model with a highly resistive salt body.

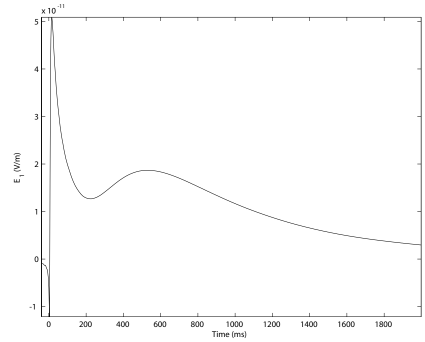


Figure 7: Time response for E_1, for a source at (6500,6500,50) m and a receiver at (9000,6500,100) m at the sea bottom.

3. CONCLUSIONS

Complexity analysis of time-domain methods for modelling electromagnetic diffusion shows that some common methods have an $O(n^4)$ complexity, where n is the number of points per spatial coordinate. Synthesizing time-domain solutions by using a frequency-domain method has a complexity of $O(n_f n^3)$, with n_f the number of frequencies, if the solver convergences in a fixed number of iterations. This can be accomplished by multigrid on uniform or mildly stretched grids. When n_f is small relative to n, this frequency-domain method appears to be the most appealing. However, as our complexity analysis only provides estimates in terms of the number of unknowns and the actual required computer time will also depend on the constants in the estimates, a true comparison of methods should involve the operation count or the cpu-time measured for an actual implementation.

Examples were included to show how frequencies can be selected and how time-domain solutions can be obtained by monotone piecewise cubic interpolation and Fourier transforms. A remaining problem is the large number of iterations that the multigrid solver needs when strong grid stretching is required at the very low frequencies.

REFERENCES

1. Mulder, W. A., "A multigrid solver for 3D electromagnetic diffusion," *Geophys. Prosp.*, Vol. 54, 663–649, 2006.
2. Mulder, W. A., "Geophysical modelling of 3D electromagnetic diffusion with multigrid," *Computing and Visualization in Science*, in press, 2006.
3. Jönsthövel, T. B., C. W. Oosterlee, and W. A. Mulder, "Improving multigrid for 3-D electromagnetic diffusion on stretched grids," *Proceedings, ECCOMAS European Conference on Computational Fluid Dynamics*, Egmond aan Zee, The Netherlands, September 2006.
4. Du Fort, E. C. and S. P. Frankel, "Stability conditions on the numerical treatment of parabolic differential equations," *Math. Tables and Other Aids to Comput. (Math. Comp.)*, Vol. 7, 135–152, 1953.
5. Drushkin, V. L. and L. A. Knizhnerman, "Spectral approach to solving three-dimensional Maxwell's diffusion equations in the time and frequency domains," *Radio Science*, Vol. 29, 937–953, 1994.
6. Drushkin, V. L., L. A. Knizhnerman, and P. Lee, "New spectral Lanczos decomposition method for induction modeling in arbitrary 3-D geometry," *Geophysics*, Vol. 64, 701–706, 1999.
7. Newman, G. A., G. W. Hohmann, and W. L. Anderson, "Transient electromagnetic response of a three-dimensional body in a layered earth," *Geophysics*, Vol. 51, 1608–1627, 1986.
8. Oristaglio, M. L. and G. W. Hohmann, "Diffusion of electromagnetic fields into a two-dimensional earth: A finite-difference approach," *Geophysics*, Vol. 49, 870–894, 1984.
9. Commer, M. A. and G. A. Newman, "A parallel finite-difference approach for 3D transient electromagnetic modeling with galvanic sources," *Geophysics*, Vol. 69, 1192–1202, 2004.
10. Ward, S. A. and G. W. Hohmann, "Electromagnetic theory for geophysical applications," *Electromagnetic Methods in Applied Geophysics — Theory*, Vol. 1, 1987.
11. Fritsch, F. N. and R. E. Carlson, "Monotone piecewise cubic interpolation," *SIAM J. Numer. Anal.*, Vol. 17, 238–246, 1980.

Green's Function Retrieval by Crossconvolutions

Evert Slob
Department of Geotechnology, Delft University of Technology, The Netherlands

Abstract— Several formulations exist for retrieving the Green's function from cross correlation of (passive) recordings at two locations. Usually these formulations retrieve Green's functions from sources on a closed boundary. Then they apply to media without losses inside the domain spanned by the sources. Until recent, these formulations were only developed for wave phenomena. Now Green's function representations for Green's functions for electromagnetic fields in matter exist. When they exploit cross correlations, the sources must lie on a boundary of a lossless medium and outside this boundary the medium can be lossy. Then such methods can be exploited for passive applications using transient or ambient noise sources, either natural or man-made. For seabed logging methods this is not a realistic scenario because possible are necessarily located in the conductive medium. A second formulation employs cross convolutions to retrieve the Green's function and this formulation remains valid when the media in the domain spanned by the sources is conductive. We derive here general exact electromagnetic Green's function retrieval and simplify the obtained results for practical applications.

1. INTRODUCTION

Since the early theoretical work of Clearbout [3] and Cox [4], and the experimental work of Weaver and Lobkis [8, 21], many others have contributed to our understanding of Green's function retrieval from cross-correlating two recordings in a noise field [2, 6, 9, 11, 15, 17–19, 22]. From one-dimensional and pulse-echo experiments the subject has evolved to arbitrary three-dimensional media, ranging from having statistical properties to being fully deterministic.

Recently, based on the principle of reciprocity representations have been derived for electromagnetic waves and fields in media with non-zero conductivity values or other relaxation mechanisms, using transient or uncorrelated noise sources [13, 14]. Here we derive representations of electromagnetic Green's functions for conductive media. When the sources lie on a closed boundary, crosscorrelation type techniques cannot be used for recordings of diffusive electromagnetic fields. An example of under which conditions a correlation type technique can be used for recordings of diffusive fields can be found in [16], who assumes that sources are distributed in a finite volume with a particular strength that is related to the local loss factor. When the sources are on a boundary and the medium inside the volume spanned by this boundary is lossless, then also cross correlation techniques can be used for exact retrieval of the full heterogeneous medium Green's function. In a seabed logging configuration this is not realistic, because if sources would exist in the air then still almost no signal would penetrate the sea layer. Therefor here we exploit the possibility of starting from the reciprocity theorem of the time-convolution type [12]. We investigate sources located on the boundary of a finite domain. We show here that for seabed logging applications an exact Green's function representation can be obtained, using sources on the boundary, by convolving two recordings at two different locations using the reciprocity theorem of the time-convolution type. The retrieved Green's function corresponds to the actual heterogeneous and anisotropic medium. This requires either electric and magnetic current sources or, in case only electric sources are used, dipole and quadrupole sources. We discuss the effects of the simplifying assumption, where the quadrupole source is approximated by an equivalent dipole source, which is necessary for practical applications in a measurement situation.

2. CAUSAL FIELD EQUATIONS

All representations are valid in the time domain for transient or noise signals, but we develop our theory in the frequency domain. To this end, we define the time-Fourier transform of a space-time dependent vector-quantity as

$$\hat{\boldsymbol{u}}(\boldsymbol{x}, \omega) = \int_{t=0}^{\infty} \exp(-\mathbf{j}\omega t)\boldsymbol{u}(\boldsymbol{x}, t)\mathrm{d}t, \tag{1}$$

where j is the imaginary unit and ω denotes angular frequency.

In the space-frequency domain Maxwell's equations in matter are given in matrix-vector form [20] by

$$\boldsymbol{D}_x\hat{\boldsymbol{u}} + \left[\hat{\boldsymbol{B}} + \mathrm{j}\omega\boldsymbol{A}\right]\hat{\boldsymbol{u}} = \hat{\boldsymbol{s}}, \tag{2}$$

where the field vector $\hat{\boldsymbol{u}}^T(\boldsymbol{x},\omega) = (\hat{\boldsymbol{E}}^T, \hat{\boldsymbol{H}}^T)$, $\hat{\boldsymbol{E}}$ and $\hat{\boldsymbol{H}}$ being the electric and magnetic field vectors and the superscript T denotes transposition, $\hat{\boldsymbol{s}}^T(\boldsymbol{x},\omega) = -(\{\hat{\boldsymbol{J}}^e\}^T, \{\hat{\boldsymbol{J}}^m\}^T)$ is the source vector, with $\hat{\boldsymbol{J}}^e$ and $\hat{\boldsymbol{J}}^m$ the external electric and magnetic current density vectors, while $\boldsymbol{D}_x$ is the matrix of spatial differential operators given by

$$\boldsymbol{D}_x = \begin{pmatrix} \boldsymbol{O} & \boldsymbol{D}_0^T \\ \boldsymbol{D}_0 & \boldsymbol{O} \end{pmatrix}, \qquad \boldsymbol{D}_0 = \begin{pmatrix} 0 & -\partial_3 & \partial_2 \\ \partial_3 & 0 & -\partial_1 \\ -\partial_2 & \partial_1 & 0 \end{pmatrix}. \tag{3}$$

The material matrices are defined as $\boldsymbol{A} = \text{blockdiag}\,(\boldsymbol{\varepsilon}, \boldsymbol{\mu})$, with $\boldsymbol{\varepsilon}$ and $\boldsymbol{\mu}$ the electric permittivity and magnetic permeability tensors and $\hat{\boldsymbol{B}} = \text{blockdiag}\,(\hat{\boldsymbol{\sigma}}^e, \hat{\boldsymbol{\sigma}}^m)$, with $\boldsymbol{\sigma}^e$ and $\boldsymbol{\sigma}^m$ the electric and magnetic conductivity tensors. Notice that we have defined the electric permittivity and magnetic permeability as frequency independent functions. This presents no loss of generality because all possible relaxation mechanisms are incorporated in the frequency dependent conductivity tensors. Further a real-valued diagonal matrix $\boldsymbol{K} = \boldsymbol{K}^{-1}$ is introduced as $\boldsymbol{K} = \text{diag}(-1, -1, -1, 1, 1, 1)$, such that $\boldsymbol{K}\boldsymbol{D}_x\boldsymbol{K} = -\boldsymbol{D}_x = -\boldsymbol{D}_x^T$, $\boldsymbol{K}\boldsymbol{A}\boldsymbol{K} = \boldsymbol{A} = \boldsymbol{A}^T$ and $\boldsymbol{K}\hat{\boldsymbol{B}}\boldsymbol{K} = \hat{\boldsymbol{B}}^T$, where the latter two definitions represent the non-negative definiteness of the material tensors. Such media are called self-adjoint or reciprocal [5].

In the next section we use the causal fields in the time-convolution type reciprocity relations. A reciprocity theorem in general interrelates two independent states, labeled A and B, in one and the same domain, but the fields, sources and the medium parameters in the two states need not be the same [1, 5, 7, 10]. In our derivations here we assume all medium parameters to be the same in both states ($\boldsymbol{A}_A = \boldsymbol{A}_B = \boldsymbol{A}$ and $\hat{\boldsymbol{B}}_A = \hat{\boldsymbol{B}}_B = \hat{\boldsymbol{B}}$). First we establish an expression for source receiver reciprocity and then we formulate the integral representation for Green's function retrieval. The Green's function corresponds to the actual heterogeneous and anisotropic medium.

3. CONVOLUTION-TYPE ELECTROMAGNETIC GREEN'S FUNCTION REPRESENTATIONS

To allow for relaxation phenomena and non-zero electric and magnetic conduction currents we now consider the reciprocity theorem of the time-convolution type. We use the interaction quantity

$$\boldsymbol{u}_A^T\boldsymbol{K}\boldsymbol{D}_x\boldsymbol{u}_B + \boldsymbol{u}_A^T\boldsymbol{K}\underleftarrow{\boldsymbol{D}}_x\boldsymbol{u}_B. \tag{4}$$

Substituting Equation (2) for the two states in this interaction quantity, integrating the result over the domain $\mathbb{D}$ and applying Gauss' divergence theorem to the interaction quantity, we find the global form of the reciprocity theorem of time-convolution type as [20]

$$\int_{\mathbb{D}} \left[\hat{\boldsymbol{u}}_A^T\boldsymbol{K}\hat{\boldsymbol{s}}_B - \hat{\boldsymbol{s}}_A^T\boldsymbol{K}\hat{\boldsymbol{u}}_B\right]\mathrm{d}^3\boldsymbol{x} = \oint_{\partial\mathbb{D}} \hat{\boldsymbol{u}}_A^T\boldsymbol{K}\boldsymbol{N}_x\hat{\boldsymbol{u}}_B\mathrm{d}^2\boldsymbol{x}, \tag{5}$$

where the minus sign in the left-hand side arises because use has been made of $\boldsymbol{D}_x^T\boldsymbol{K} = -\boldsymbol{K}\boldsymbol{D}_x$. Notice that in the convolution type representations the relaxation and loss mechanisms do not occur in the expression for reciprocal media and hence we do not have to assume that the medium is lossless.

To localize the electric field receiver locations at $\boldsymbol{x}_A$ and $\boldsymbol{x}_B$ we specify the artificial point sources by replacing the space and frequency dependent 6×1 vector $\hat{\boldsymbol{s}}_A$ by the 6×6 matrix $\boldsymbol{I}\delta\,(\boldsymbol{x}-\boldsymbol{x}_A)$, $\boldsymbol{I}$ being the identity matrix. The corresponding 6×1 field vector $\hat{\boldsymbol{u}}_A$ is replaced by the 6×6 Green's matrix $\hat{\boldsymbol{G}}\,(\boldsymbol{x},\boldsymbol{x}_A,\omega)$, given by

$$\hat{\boldsymbol{G}}\,(\boldsymbol{x},\boldsymbol{x}_A,\omega) = \begin{pmatrix} \hat{\boldsymbol{G}}^{Ee} & \hat{\boldsymbol{G}}^{Em} \\ \hat{\boldsymbol{G}}^{He} & \hat{\boldsymbol{G}}^{Hm} \end{pmatrix}(\boldsymbol{x},\boldsymbol{x}_A,\omega), \tag{6}$$

where the superscripts $\{E,H\}$ denote the observed fieeld type at $\boldsymbol{x}$ and the superscripts $\{e,m\}$ denote the source type at $\boldsymbol{x}_A$. In the submatrices each Green's tensor denotes one 3×3 Green's tensor. Each column of $\hat{\boldsymbol{G}}$ represents a field vector at $\boldsymbol{x}$ due one particular source type and component at $\boldsymbol{x}_A$. For state B we make similar choices, replacing $\hat{\boldsymbol{s}}_B$ by $\boldsymbol{I}\delta(\boldsymbol{x}-\boldsymbol{x}_B)$ and $\hat{\boldsymbol{u}}_B$ by $\hat{\boldsymbol{G}}(\boldsymbol{x},\boldsymbol{x}_B,\omega)$. In case both source locations are outside $\mathbb{D}\cup\partial\mathbb{D}$ or inside $\mathbb{D}$ the boundary integral vanishes [1]. When both $\boldsymbol{x}_A$ and $\boldsymbol{x}_B$ are inside $\mathbb{D}$ we find from substituting these choices for the sources and the fields in Equation (5) the source-receiver reciprocity relation as,

$$\hat{\boldsymbol{G}}^T(\boldsymbol{x}_B,\boldsymbol{x}_A,\omega)\boldsymbol{K}-\boldsymbol{K}\hat{\boldsymbol{G}}(\boldsymbol{x}_A,\boldsymbol{x}_B,\omega)=\boldsymbol{O}. \tag{7}$$

When we transpose Equation (7) and use $\boldsymbol{K}\boldsymbol{K}=\boldsymbol{I}$ we find the alternative expression as

$$\hat{\boldsymbol{G}}(\boldsymbol{x}_B,\boldsymbol{x}_A,\omega)\boldsymbol{K}=\boldsymbol{K}\hat{\boldsymbol{G}}^T(\boldsymbol{x}_A,\boldsymbol{x}_B,\omega). \tag{8}$$

Making general replacements for sources and fields Equation (5) is replaced by

$$\hat{\boldsymbol{G}}(\boldsymbol{x}_B,\boldsymbol{x}_A,\omega)\boldsymbol{K}\left[\chi_{\mathbb{D}}(\boldsymbol{x}_A)-\chi_{\mathbb{D}}(\boldsymbol{x}_B)\right]=\oint_{\partial\mathbb{D}}\hat{\boldsymbol{G}}(\boldsymbol{x}_B,\boldsymbol{x},\omega)\boldsymbol{N}_x\boldsymbol{K}\hat{\boldsymbol{G}}^T(\boldsymbol{x}_A,\boldsymbol{x},\omega)\,\mathrm{d}^2\boldsymbol{x}, \tag{9}$$

where Equation (8) has been used also for the Green's functions in the integrand in the right-hand side of Equation (9). Equation (9) is an exact representation for the electromagnetic Green's function between $\boldsymbol{x}_A$ and $\boldsymbol{x}_B$ in terms of cross-convolutions of impulsive field responses observed at the observation points $\boldsymbol{x}_A$ and $\boldsymbol{x}_B$ due to tangential electric and magnetic point sources on the boundary $\partial\mathbb{D}$ and integrating over all source locations on the closed boundary surface $\partial\mathbb{D}$. Possible applications of Equation (9) for electromagnetic interferometry will be investigated in the next section.

4. MODIFICATIONS FOR EM INTERFEROMETRY

In the present form, Equation (9) contains the matrix $\boldsymbol{N}_x\boldsymbol{K}$ in the cross-convolution expressions in the surface integral. For a direct application in terms of convolutions of observed wave fields due to uncontrolled sources the matrix $\boldsymbol{N}_x\boldsymbol{K}$ should be diagonalized, in which process a source decomposition is necessary into sources for inward and outward traveling waves and fields. We first diagonalize the representations by rewriting them in terms of observations of the electric field due to electric current sources on the boundary only. In a second step we make simplifying assumptions for the inward and outward traveling waves. These are necessary for practical transient sources.

Now we reduce the field vector to the electric field and reduce the full Green's matrix to the electric field Green's tensor for an electric source. Then we find [14]

$$\begin{aligned}\hat{\boldsymbol{G}}^{Ee}(\boldsymbol{x}_B,\boldsymbol{x}_A,\omega)\left[\chi_{\mathbb{D}}(\boldsymbol{x}_A)-\chi_{\mathbb{D}}(\boldsymbol{x}_B)\right]=\;&\frac{1}{\mathbf{j}\omega\mu}\oint_{\partial\mathbb{D}}\hat{\boldsymbol{G}}^{Ee}(\boldsymbol{x}_B,\boldsymbol{x},\omega)\left\{\boldsymbol{n}\cdot\nabla\hat{\boldsymbol{G}}^{Ee}(\boldsymbol{x}_A,\boldsymbol{x},\omega)\right\}^T\mathrm{d}^2\boldsymbol{x}\\&-\frac{1}{\mathbf{j}\omega\mu}\oint_{\partial\mathbb{D}}\left\{\boldsymbol{n}\cdot\nabla\hat{\boldsymbol{G}}^{Ee}(\boldsymbol{x}_B,\boldsymbol{x},\omega)\right\}\left\{\hat{\boldsymbol{G}}^{Ee}(\boldsymbol{x}_A,\boldsymbol{x},\omega)\right\}^T\mathrm{d}^2\boldsymbol{x},\end{aligned} \tag{10}$$

which is still an exact representation under the assumption that the medium in the neighborhood of the boundary is homogeneous. If we assume the points $\boldsymbol{x}_A$ and $\boldsymbol{x}_B$ lie in the far field of the boundary, then we can approximate the normal derivative and replace it by a multiplicative factor of $\pm\sqrt{\mathbf{j}\omega\sigma\mu-\omega^2\varepsilon\mu}$, where the plus sign applies to outward traveling waves and the minus sign to inward traveling waves. This multiplicative factor also assumes the contribution from fields diffusing away from the boundary in the normal direction have the major contribution in the final result. Substituting this in Equation (10) yields

$$\hat{\boldsymbol{G}}^{Ee}(\boldsymbol{x}_B,\boldsymbol{x}_A,\omega)\left[\chi_{\mathbb{D}}(\boldsymbol{x}_A)-\chi_{\mathbb{D}}(\boldsymbol{x}_B)\right]+\text{“ghost”}=-2Y\oint_{\partial\mathbb{D}}\hat{\boldsymbol{G}}^{Ee}(\boldsymbol{x}_B,\boldsymbol{x},\omega)\left\{\hat{\boldsymbol{G}}^{Ee}(\boldsymbol{x}_A,\boldsymbol{x},\omega)\right\}^T\mathrm{d}^2\boldsymbol{x}, \tag{11}$$

where $Y=\sqrt{(\sigma+\mathbf{j}\omega\varepsilon)/(\mathbf{j}\omega\mu)}\approx\sqrt{\sigma/(\mathbf{j}\omega\mu)}$ denotes the complex admittance and the “ghost” term represents spurious events that are suppressed when the boundary is irregular.

We define the matrix of measured electric fields generated by transient electric current sources as

$$\hat{\varepsilon}^{\text{obs}}_{A,B}(\boldsymbol{x})=\hat{\boldsymbol{G}}^{Ee}(\boldsymbol{x}_{A,B},\boldsymbol{x},\omega)\hat{\boldsymbol{S}}(\boldsymbol{x},\omega), \tag{12}$$

where $\hat{\boldsymbol{S}}(\boldsymbol{x},\omega) = \mathrm{diag}[s_1(\boldsymbol{x},\omega),\, s_2(\boldsymbol{x},\omega),\, s_3(\boldsymbol{x},\omega)]$ denotes the source frequency spectrum matrix at position $\boldsymbol{x}$, which can be different for each direction and for each source position. The power spectrum matrix of the sources is defined as

$$\hat{\boldsymbol{S}}^P(\boldsymbol{x},\omega) = \mathrm{diag}\left(|s_1(\boldsymbol{x},\omega)|^2,\, |s_2(\boldsymbol{x},\omega)|^2,\, |s_3(\boldsymbol{x},\omega)|^2\right). \tag{13}$$

Using these definitions in Equation (11) we find

$$\hat{\boldsymbol{G}}^{Ee}(\boldsymbol{x}_B,\boldsymbol{x}_A,\omega)\left[\chi_{\mathbb{D}}(\boldsymbol{x}_A) - \chi_{\mathbb{D}}(\boldsymbol{x}_B)\right] \approx -2Y \int_{\boldsymbol{x}\in\partial\mathbb{D}} \hat{\boldsymbol{\varepsilon}}_B^{\mathrm{obs}}(\boldsymbol{x}) \left(\hat{\boldsymbol{S}}^P\right)^{-1} \left\{\hat{\boldsymbol{\varepsilon}}_A^{\mathrm{obs}}(\boldsymbol{x})\right\}^T \mathrm{d}^2\boldsymbol{x}, \tag{14}$$

where the approximate sign has replaced the equality sign because we have omitted the explicit mention of the "ghost" term. The fact that the inverse of the power spectrum matrix is required indicates that it should be known to use this method for transient sources.

ACKNOWLEDGMENT

This work is part of the research program of the Netherlands research center for Integrated Solid Earth Science (ISES).

REFERENCES

1. Bojarski, N., "Generalized reaction principles and reciprocity theorems for the wave equations, and the relationships between time-advanced and time-retarded fields," *Journal of the Acoustical Society of America*, Vol. 74, 281–285, 1983.
2. Campillo, M. and A. Paul, "Long-range correlations in the diffuse seismic coda waves," *Science*, Vol. 299, 547–549, 2003.
3. Claerbout, J., "Synthesis of a layered medium from its acoustic transmission response," *Geophysics*, Vol. 33, 264–269, 1968.
4. Cox, H., "Spatial correlation in arbitrary noise fields with application to ambient sea noise," *J. Acoust. Soc. Am.*, Vol. 54, 1289–1301, 1973.
5. De Hoop, A., *Handbook of Radiation and Scattering of Waves*, Academic Press, Amsterdam, 1995.
6. Derode, A., E. Larose, M. Tanter, J. de Rosny, A. Tourin, M. Campillo, and M. Fink, "Recovering the Green's function from field-field correlations in an open scattering medium (L)," *JASA*, Vol. 113, No. 6, 2973–2976, doi:10.1121/1.1 570 436, 2003.
7. Lindell, I., A. Sihvola, S. Tretyakov, and A. Viitanen, *Electromagnetic Waves in Chiral and Bi-isotropic Media*, Artech House, London, 1992.
8. Lobkis, O. and R. Weaver, "On the emergence of the Green's function in the correlations of a diffuse field," *JASA*, Vol. 110, 3011–3017, 2001.
9. Roux, P., K. Sabra, W. Kuperman, and A. Roux, "Ambient noise cross correlation in free space: Theoretical approach," *JASA*, Vol. 117, No. 1, 79–84, 2005.
10. Cheo, B. R.-S., "A reciprocity theorem for electromagnetic fields with general time dependence," *IEEE Trans. Ant. Prop.*, Vol. 13, 278–284, 1965.
11. Shapiro, N., M. Campillo, L. Stehly, and M. Ritzwoller, "High-resolution surface-wave tomography from ambient seismic noise," *Science*, Vol. 307, 1615–1618, 2005.
12. Slob, E. and K. Wapenaar, "Electromagnetic Green's functions retrieval by cross-correlation and cross-convolution in media with losses," *Phys. Rev. Lett.*, submitted, 2006.
13. Slob, E., D. Draganov, and K. Wapenaar, "Gpr without a source," in *Proceedings of the 11th International Conf. on GPR*, Ohio State University, Columbus Ohio, ANT.6, 8pp, 2006.
14. Slob, E., D. Draganov, and K. Wapenaar, "Interferometric electromagnetic Green's functions representations using propagation invariants," *Geoph. J. Int.*, accepted, 2006.
15. Snieder, R., "Extracting the Green's function from the correlation of coda waves: A derivation based on stationary phase," *Phys. Rev. E*, Vol. 69, 046 610-1-8, doi:10.1103/PhysRevE.69.046 610, 2004.
16. Snieder, R., "Extracting the Greens function of attenuating media from uncorrelated waves," *JASA*, submitted, 2006.
17. Van Tiggelen, B., "Green function retrieval and time reversal in a disordered world," *Phys. Rev. Lett.*, Vol. 91, 243 904-1-4, doi:10.1103/PhysRevLett.91.243 904, 2003.

18. Wapenaar, K., "Retrieving the elastodynamic Green's function of an arbitrary inhomogeneous medium by cross correlation," *Phys. Rev. Lett.*, Vol. 93, 254 301-1-4, doi:10.1103/PhysRevLett.93.254 301, 2004.
19. Wapenaar, K., D. Draganov, J. Thorbecke, and J. Fokkema, "Theory of acoustic daylight imaging revisited," in *Expanded Abstracts of the 72nd Annual Internat. Mtg. Soc. Expl. Geophys.*, 2269–2272, 2002.
20. Wapenaar, K. and J. Fokkema, "Reciprocity theorems for diffusion, flow and waves," *A.S.M.E. Journal of Applied Mechanics*, Vol. 71, 145–150, 2004.
21. Weaver, R. and O. Lobkis, "Ultrasonics without a source: Thermal fluctuation correlations at MHz frequencies," *Phys. Rev. Lett.*, Vol. 87, 134 301-1-4, doi:10.1103/PhysRevLett.87.134 301, 2001.
22. Weaver, R. and O. Lobkis, "Diffuse fields in open systems and the emergence of the Green's function (L)," *JASA*, Vol. 116, No. 5, 2731–2734, doi:10.1121/1.1 810 232, 2004.

Solving Electromagnetic Inverse Scattering Problems by SVRMs: a Case of Study Towards Georadar Applications

G. Angiulli, V. Barrile, and M. Cacciola

Universitá *Mediterranea* degli Studi di Reggio Calabria, DIMET
Via Graziella, loc. Feo di Vito-89100 Reggio Calabria, Italy

Abstract— In this paper, an heuristic approach based on Support Vector Regression Machines (SVRMs) is presented in order to solve a simple inverse scattering problem. Interesting results have been obtained, with a remarkable reduction of computational time. Future development of this works will interest the evaluation of the performances of SVRMs for detection of buried objects in stratified media. This is the starting point to develop models for typical Georadar applications.

1. INTRODUCTION

Inverse electromagnetic scattering by objects that lie in free space or in layered media plays an increasing role in a wide range of technological applications. It is for this reason that, during the years, many methodological approach have been developed for a variety of problems involving e.g., one-dimensional and high dimensional unknowns in a homogeneous space with linear scattering approximations, and higher dimensional unknowns in a homogeneous space considering multiple scattering mechanism (see [1] and references within). Among the numerous technological applications of the electromagnetic inverse scattering, Ground-Penetrating Radar (GPR) also known as Georadar, is one of the most important. GPR is a near-surface remote sensing tool for detecting buried targets (see [2] and references within). Interesting applicative fields of GPR are measurements of object location into the subsoil (i. e., pipings, electric or telephonic cables, and so on) or soil characterization. In all these applications, it is very important to quickly obtain measures with an high level of precision in terms of location and dimensions of buried objects. However, the accurate modeling of a GPR is a complex task. Nevertheless, in order to obtain a quick GPR data processing, it is necessary to develop suitable models able to face the inverse problems.

In the last years, Soft Computing techniques, such as Neural Networks, Neuro Fuzzy Networks have been introduced in order to provide a fast treatment of the direct and inverse scattering problems (see [3, 4] and references within). Ability and adaptability to learn and generalize, fast real-time operation, and ease of implementation have made these techniques very popular. Very recently, another Soft Computing technique named as Support Vector Machines (SVMs), developed by Vapnik [5], has gained popularity due to many attractive features capable to overcome the limitations connected to Neural Networks. This is due to the Structural Risk Minimisation principle embodied by SVMs, which has been demonstrated to be more effective than the traditional Empirical Risk Minimisation principle employed by Neural Networks [5]. This different philosophy provides Support Vector Machines with a greater ability to generalise, if compared with Neural Networks. In this paper we investigate the performances of SVMs in the field of the inverse scattering. To this aim, our attention is focused on a simple electromagnetic inverse problem: the location of a perfect conducting thin metal strip immersed in the free space starting from the scattered field evaluated at a suitable number of measure points. This is a simple inverse problem, which can be solved exploiting the direct one (see [6, 7] and references within).

The paper is organized as follows: Section 2 gives the basics of SVRMs. Section 3 gives a brief account of the direct electromagnetic scattering problem exploited to collect data for SVRMs-based experimentations. Next, Section 4 describes the characteristics of collected dataset and hosts some discussions about retrieved preliminary results. Finally, Section 5 draws up our conclusions. All the computer codes exploited in this work have been implemented in Matlab®, using also a freeware toolbox for SVRMs [8].

2. A QUICK OVERVIEW OF SVRMS

SVRMs are learning machines that can be applied to regression problems (see [5] and references within). Their operation principle can be summarize as follow: like in standard regression problems, it is supposed that the relationship between the independent and dependent variables is given by a

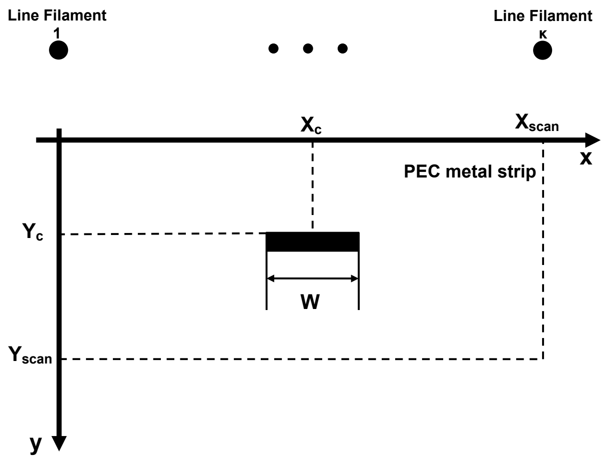


Figure 1: Pictorial representation of the direct scattering problem discuss in Sec. 3. Please note an interesting X_{scan}/W ratio, which is very important in GPR applications to detect and characterize small buried objects, according to the definition abilities of used Georadar apparatus.

function f plus the addition of some additive noise. The task is to find a functional form for f which can correctly predict new cases that the SVRM has not been presented with before. Considering the problem of approximating the set of data $D = \{(x^1, y^1), \ldots, (x^l, y^l)\}, \mathbf{x} \in \Re^n, y \in \Re$ with a linear function $f(\mathbf{x}) = \langle \mathbf{w}, \mathbf{x} \rangle + b$, the optimal regression function is given by the minimum of the functional

$$\Phi(\mathbf{w}, \xi) = \frac{1}{2}\|\mathbf{w}\|^2 + C \sum_i \left(\xi_i^- + \xi_i^+\right) \tag{1}$$

where C is a user defined value, and ξ_i^-, ξ_i^+ are slack variable representing upper and lower constraints on the output of the system. If a linear regression is not possible due to non-linear relationships between data, SVRM non-linearly map the n-dimensional input space into a high dimensional space where a linear separation can be performed (see [9] and references within). In this case, a non-linear function, also known as kernel function (see [10] and references within), is used. This can be achieved by training the SVRM on a suitable training set. This process involves sequential optimization of an error (or loss) function (see [11] and references within). Depending on its loss function definition, two kinds of SVRMs can be recognized: ε-SVRM and ν-SVRM. In this research, ν-SVRM has been used (see [12] and references within).

3. CREATION OF THE TRAINING SET: DESCRIPTION OF THE DIRECT SCATTERING PROBLEM

In this section a brief account of the direct electromagnetic scattering problem, exploited to collect data for SVRM-based experimentations, is given. In Fig. 1 is shown a perfect conducting two dimensional metal strip, located in free-space within a box defined by coordinates $0 < x < x_{scan}$, $y_{scan} < y < 0$. $(x_c,\, y_c)$ are the coordinates of the strip centre. A number of κ line filaments, working as either a source or a receiver, are located at an coordinate $y = y_Q$, in equally spaced points such that the x-coordinates of the first and last filament are 0 and x_{scan}, respectively. Each line filament illuminates the metal strip by a TM_z cylindrical electromagnetic wave. As well known, the incident fields induce electric current on the strip, which radiates a scattered TM_z wave, and this back-scattering field is detected by the same line filament in the receiving phase. Applied the boundary condition for the tangential electric field on the the strip, it is possible to write [13]:

$$E_z^{scatt}(x) = \frac{k_0 \eta_0}{4} \int_{-\frac{W}{2}}^{\frac{W}{2}} \mathsf{H}_0^{(2)} \left(k_0 |x - x'|\right) J_z(x')\, dx' \tag{2}$$

where $E_z^{scatt}(x)$ is the tangential scattered electric field, $J_z(x')$ is the induced unknown current density flowing on the metal strip, k_0 is the free-space propagation constant, η_0 is the free-space intrinsic impedance, $\mathsf{H}_0^{(2)}$ is the Hankel function of the second kind and zero order and x and x'

represent the x-coordinates of the observation and source points, respectively. Solving Eq. (2) for $J_z(x')$ by MoM, the back-scattered field can be expressed as [13]:

$$E_z^{scatt}(\bar{\rho}) = -\left(\frac{2}{\pi W}\right)\frac{k_0\eta_0}{4}\int_{-\frac{W}{2}}^{\frac{W}{2}} \mathsf{H}_0^{(2)}\left(k_0|\bar{\rho}-\bar{\rho}'|\right) J_z(x')\,dx' \tag{3}$$

where $\bar{\rho}$ and $\bar{\rho'}$ are the magnitude of vectors related to observation and source points respectively in a polar reference system.

4. DATA COLLECTION AND PRELIMINARY RESULTS

By using the described formulation and considering suitable numeric values for above described quantities, it is possible to solve the direct problem (2) by MoM and so collect a dataset. The values exploited in our numerical experimentation are reported in Table 1 (these data are the same of those employed in [7]). A number of 676 patterns have been collected in order to carry out training and test of a suitable SVRM (called MSLSVRM). It is possible to discriminate training and test elements by considering Fig. 2. In order to train MSLSVRM, it is necessary to adequately choose the kernel function. This operation has been carried out by fixing the other SVRM parameters, that is $C = 1$ and $\xi_i^- = \xi_i^+ = \xi = 0.1 \quad \forall i$. In this way, an RBF kernel has been selected as the kernel having the best regression performances. Subsequently, in order to improve MSLSVRM performances, C and ξ values have been tuned to the following values: $C' = 10$ and $\xi' = \xi$. Time elapsed to train MSLSVRM for estimation of (x, y) coordinates of metal strips' centres as a function of the backscattered field is 0.876 s. Test results are instantaneously communicated by MSLSVRM.

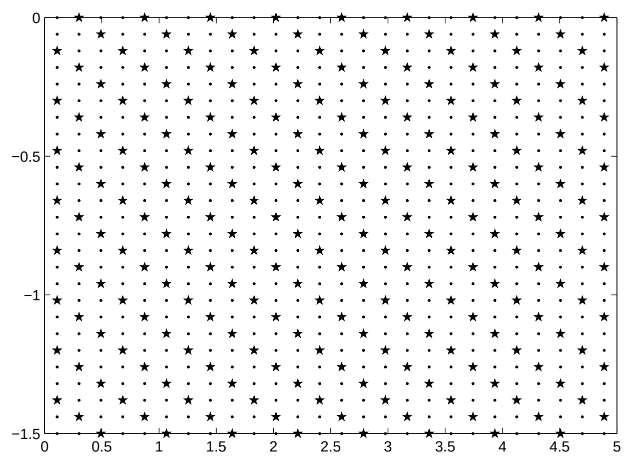


Figure 2: Considered centres of metal strips for training (dots) and testing (stars) elements.

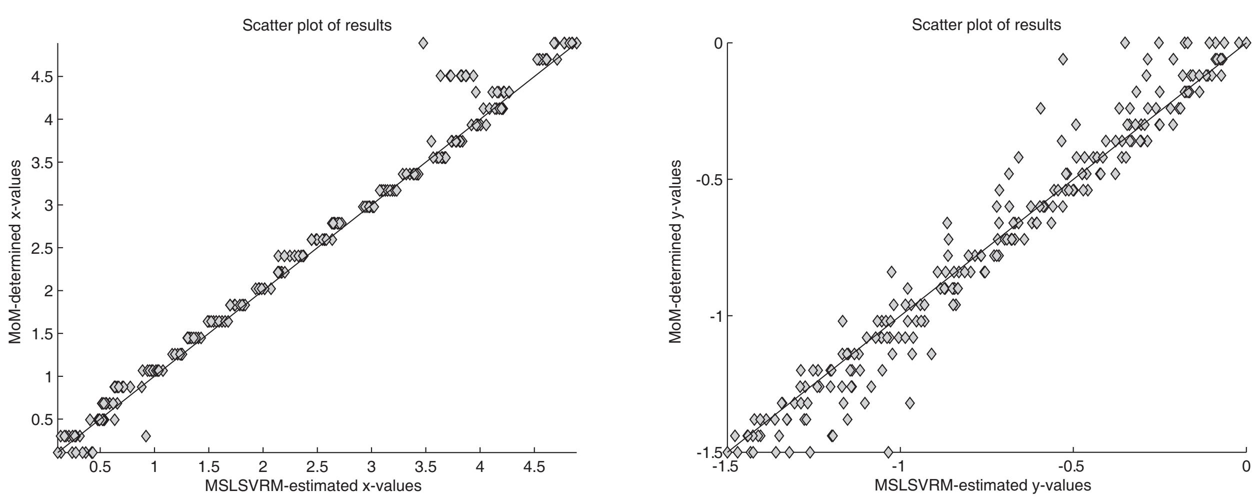


Figure 3: Scatter plots between MoM-determined and MSLSVRM-estimated values of x-coordinate (left) and y-coordinate (right) of metal strips' centres. Plotted lines are best-fitting lines.

Table 1: Numerical quantities used to build the dataset.

Quantity	f [MHz]	W [m]	x_{scan} [m]	y_{scan} [m]	N	y_Q [m]	κ
Value	100	0.2	5	-1.5	5	5	10

Table 2: Statistics about distances between observed metal strip's centers and relative values estimated by MSLSVRM.

Minimal error [m]	**Maximal error [m]**	**Average error [m]**	**Standard deviation [m]**	**Root mean squared error (RMSE) [m]**	**RMSE percent** $\left(\frac{RMSE}{D} * 100\right)$
0.004	1.292	0.163	0.168	0.234	4.667%

Preliminary results obtained by our experimentation are very satisfying (Fig. 3), above all for the estimation of x-coordinates of metal strips' center. The average distance between observed and simulated values of metal strips' centres is equal to 0.163 m, the 3.25% on the greatest possible distance for the case of study, i. e., $D = \|(0.11, 0) - (4.89, -1.5)\|$. For details on estimative errors, see Table 2.

5. CONCLUSIONS

In this paper a new Soft Computing approach for solving inverse scattering problems has been investigated. It is based on the use of heuristic SVRMs. In order to test the performances of this approach a simple case of study has been selected: the localization of a perfect conducting metal strip lying in a free-space environment. The obtained results shown that the proposed approach is accurate and fast, consequently very promising for real time applications. Future development of this work will interest the evaluation of the performances of SVRMs on real experimental data for detection of buried objects in stratified media. Therefore, presented method is a preliminary starting point to develop suitable models for typical Georadar applications.

REFERENCES

1. Chew, W. C., J. Jin, Michielssen, and J. Song, *Fast and Efficient Alghorithms in Computational Electromagnetics*, Artech House, Boston, 2001.
2. Angiulli, G., V. Barrile, and M. Cacciola, "The GPR technology on the seisimic damageability assessment of reinforced concrete building," *PIERS Online*, Vol. 1, No. 3, 303–307, 2005.
3. Rekanos, I. T., "Inverse scattering of dielectric cylinders by using radial basis function neural networks," *Radio Science*, No. 36, 841–849, 2001.
4. Angiulli, G. and M. Versaci, "Resonant frequency evaluation of microstrip antennas using a neural fuzzy approach," *IEEE Trans. on Magnetics*, Vol. 39, No. 3, 1333–1336, 2003.
5. Cortes, C. and V. Vapnik, "Support vector network," *Machine Learning*, Vol. 20, 273–297, 1997.
6. Joachimowicz, N., C. Pichot, and J.-P. Hugonin, "Inverse scattering: An iterative numerical method for electromagnetic imaging," *IEEE Trans. Antennas Propagat.*, No. 39, 1742–1752, 1991.
7. Soliman, E. A., A. K. Abdelmageed, and M. A. El-Gamal, "Estimating the location of a metal strip using radial basis function neural networks," *Electromagnetics*, No. 23, 431–443, 2003.
8. "The Spider Matlab Toolbox," freeware Matlab® toolbox available at http://www.kyb.tuebingen.mpg.de/bs/people/spider/main.html (August 30, 2006).
9. Gunn, S. R., "Support vector machines for classification and regression," *ISIS Technical Report*, 1998.
10. Boser, B. E., I. M. Guyon, and V. Vapnik, "A training algorithm for optimal margin classifiers," *Proceedings of the 5th Annual ACM Workshop on Computational Learning Theory*, ACM Press, 1992.
11. Smola, A. J., "Regression estimation with support vector learning machines," Master's thesis, Technische Universitat Munchen, 1996.

12. Chen, P.-H., C.-J. Lin, and B. Schölkopf, "A tutorial on ν-support vector machines," available at www.csie.ntu.edu.tw/~cjlin/papers/nusvmtutorial.pdf (August 30, 2006).
13. Balanis, C., *Advanced Engineering Electromagnetics*, John Wiley and Sons, 1989.

A Ground-wave Technique for Pavement Permittivity and Thickness Estimation from GPR Data

Jing Li, Huichun Xing, Ying Wang, Chienping Kao, and Richard Liu

Department of Electrical and Computer Engineering, University of Houston, Houston, TX, USA

Abstract— A Ground-wave technique is introduced in this paper to directly extract pavement permittivity and thickness information from measured data of groundcoupled ground-penetrating radar (GPR). This technique enables a bistatic radar to obtain both thickness and permittivity by just one measurement, which effectively reduces measurement and computation time for GPR applications.

1. INTRODUCTION

Permittivity is one of the most important electrical parameters of a material. In the subsurface sensing industry, it is related to the physical properties of the subsurface media, like the material types, the moisture content in the material, and the wave travel speed through the material. When electromagnetic wave is employed to measure the thickness of subsurface layers, the permittivity must be determined in advance or simultaneously with the thickness. Hence measuring the permittivity of materials has interested researchers and engineers for decades, and many measuring methods have been developed. Typical lab methods include transmission-line Technique [1–4] and cavity technique [5]. In transmission-line technique, a material sample of certain shape and size is inserted in a waveguide or coaxial transmission line. By measuring the scattering parameters or the propagation constants of the sample-loaded transmission line, the permittivity of the material sample can be calculated by various inversion algorithms. In the cavity technique, the material sample is inserted in a calibrated cavity. The calibration is usually conducted by using a known sample of the same shape and size as the sample to be measured. By measuring the reflections of the sample-loaded cavity, the permittivity of the sample can be determined. Though lab methods can produce accurate results, they are only applicable to the shaped material samples; it is not a solution to in-situ measurements. The open-ended waveguide method [6] for the permittivity determination sets a step forward to the in-situ applications. This method uses an open-ended waveguide to radiate into the layered materials and measures the reflection parameters to determine the permittivity of the layers. But this method needs expensive device such as network analyzers to obtain accurate measurements, which is very inconvenient to the GPR users. A GPR user is always expecting to derive both the permittivity and thickness of the subsurface layers directly from the measured GPR data. For an air-coupled GPR that is usually setup a few feet above the pavement surface, the surface reflection method [7] can be used to determine the permittivity of the layer by:

$$\varepsilon = \left(\frac{1 + A_{\text{pave}}/A_{\text{metal}}}{1 - A_{\text{pave}}/A_{\text{metal}}} \right)^2 \tag{1}$$

where, ε is the layer permittivity; A_{pave} is the surface reflection amplitude of the pavement to be measured, and A_{metal} is the metal plate reflection amplitude when a big metal plate is placed on the pavement surface. However, this method only applies to the air-coupled GPR. For the ground-coupled GPR, the GPR is setup very close to or simply sits on the surface of the pavement; the metal plate calibration technique is no longer available. Furthermore, the surface reflected wave merges into the direct wave that propagates directly from the transmitting antenna to the receiving antenna through the air such that is surface reflection amplitude is not identifiable. Hence the surface reflection method represented by Equation (1) is not applicable to the groundcoupled GPR. Fortunately, when the GPR antennas are very close to or in direct contact with the ground and the transmitting and receiving antennas separation is fixed, the pulse echoes from a subsurface "point target" sampled at equally spaced along the x-direction produce a hyperbola in space-time, as shown in Figure 1.

Theoretically if two or more measurements are made, the permittivity and the depth of the target can be determined. Several algorithms have been successfully developed based on this phenomenon [8–10]. Hyperbolic feature of a subsurface target provides a reliable way for the ground permittivity

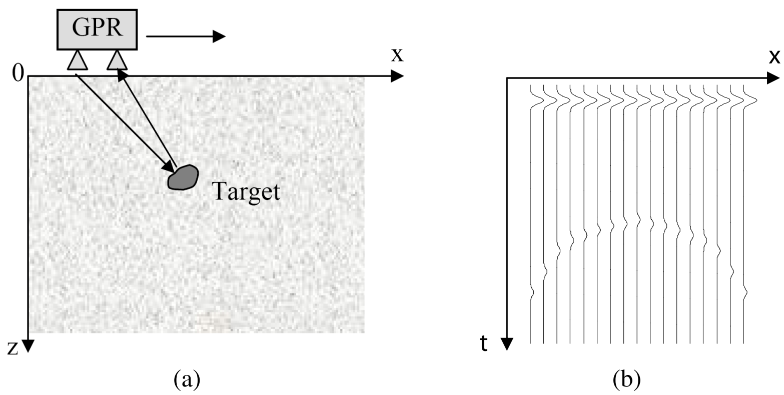


Figure 1: (a) GPR configuration, and (b) pulse echoes from the target.

determination. But in the applications such as pavement thickness measurement (except the steel bar reinforced case), subsurface point targets may not exist; hence no hyperbolic feature to be employed. Under this condition, the improved common middle point method [11] is the conventional way to solve the permittivity and thickness of the pavement. The common middle point method requires two or more measurements with different transmitting-to-receiving antenna separations to establish a set of related equations at a measuring location. The drawbacks of this method includes: 1) it consumes time to adjust antenna separation for each measurement, and 2) the obtained related equations are usually ill-conditioned.

In this paper, authors will introduce a ground-wave technique to extract the ground permittivity directly from the measured GPR data using a very simple algorithm. This method is especially suitable for vehicle-mounted fast speed GPR measurements.

2. GROUND-WAVE TECHNIQUE

The ground-wave technique is proposed based on the assumption of existence of an evanescent ground wave that propagates underground from the location under the transmitting antenna to the location under the receiving antenna, as the path 2 illustrated in Figure 2. Path 2 indicates the propagation direction of the ground wave, and physically this wave decays exponentially if it leaves the ground and propagates into air (in the negative z-direction). Only when the GPR antennas are close enough to the ground, this wave can be detected. Path 1 in Figure 2 denotes the direct coupling between the transmitting and receiving antennas as well as the ground surface reflections. Paths 3 and 4 denote the subsurface interface reflections. The path 4 has been introduced and employed for the determination of permittivity and thickness of the layered media [11].

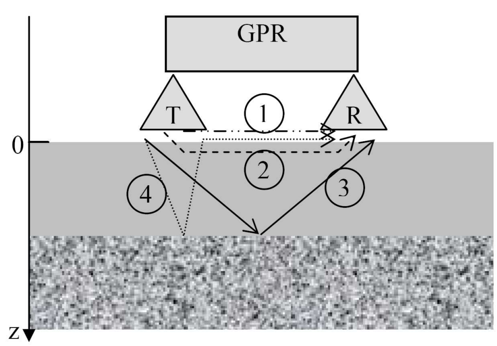


Figure 2: Wave paths in ground-coupled GPR measurements.

In order to verify the existence of path 2 in real GPR applications and search for new methods for dielectric constant estimation, numerical simulations of ground-coupled GPR are conducted using TLM method.

3. NUMERICAL SIMULATION OF THE RECEIVED GPR ECHOES

Figure 3 illustrates the geometry of GPR setup. The transmitter is at a fixed position, and the receiver is placed at 20 cm, 30 cm, 40 cm, 50 cm, and 60 cm away, respectively, from the transmitter. The simulated waveforms in different transmitter-receiver offsets are given in Figure 4.

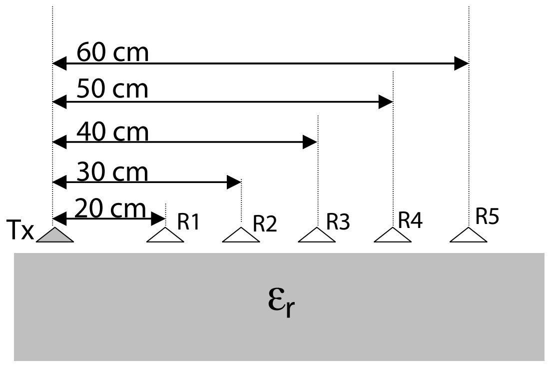


Figure 3: Geometry for numerical simulation.

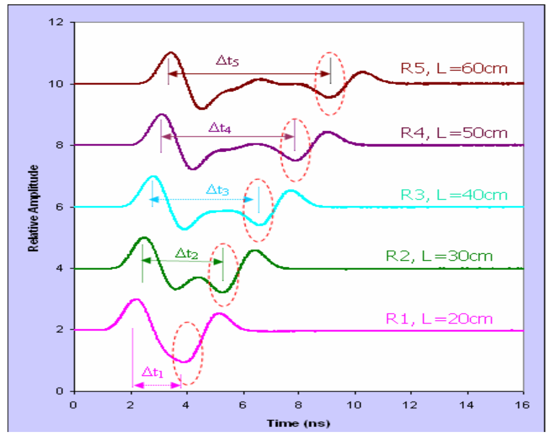


Figure 4: Simulated waveforms at positions R1, R2, R3, R4, and R5.

It is illustrated by Figure 4 that the signals in the dashed circles delay with respect to the direct wave as the transmitter-receiver offset increases. The correspondence of the time delay to the transmitter-receiver offset is plotted in Figure 5, in blue line. On the other hand, if we assume the wave take the path 2 in Figure 2 and estimate the time delay Δt by,

$$\Delta t = (\sqrt{\varepsilon_r} - 1)\, L/c \tag{2}$$

where L is the transmitter-receiver offset and c is the light velocity in free space, then it is found that the estimated time delay, see the pink line in Figure 5, is identical to the TLM simulated time delay. This phenomenon implies that the signals in the dashed circles in Figure 4 are indeed the proposed ground wave. By measuring the time delay of the ground wave with respect to the direct wave, the dielectric constant of the pavement can be determined,

$$\varepsilon_r = \left(\frac{c\Delta t}{L} + 1\right)^2 \tag{3}$$

With the dielectric constant predetermined, the thickness of the pavement can be directly solved by analytic formula. This procedure greatly increases the in-situ processing time, realizing real time measurements.

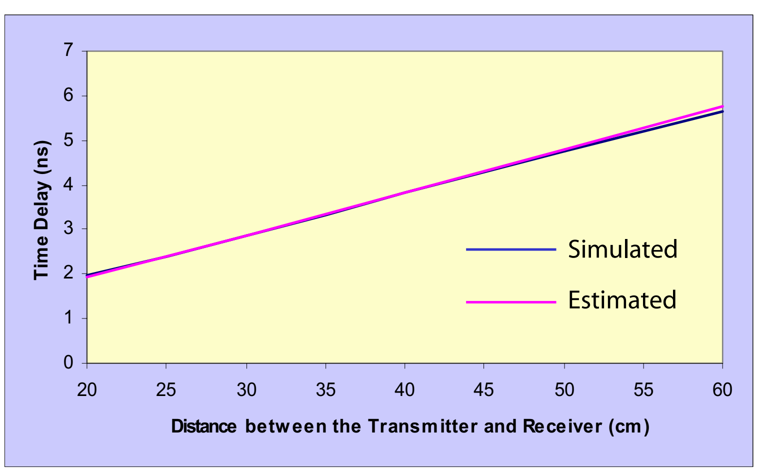


Figure 5: Comparison of simulated time and modeled time.

4. CONCLUSIONS

This paper proposed a ground-wave technique for extracting the permittivity of pavement from measured ground-coupled GPR data. This method has been verified by numerical simulations. With this technique, the permittivity of the pavement can be analytically derived from the measured GPR data and no iteration and other complicated algorithms are needed, which greatly facilitate the in-situ real time measurements.

REFERENCES

1. Deshpande, M. D., C. J. Reddy, P. I. Tiemsin, and R. Cravey, "A new appoach to estimate complex permittivity of dielectric materials at microwave frequencies using waveguide measurements," *IEEE Trans. MTT*, Vol. 45, No. 3, 1997.
2. Sarabandi, K., "A technique for dielectric measurement of cylindrical objects in a rectangular waveguide," *IEEE Trans. IM*, Vol. 43, No. 6, 1994.
3. Janezic, M. D. and J. A. Jargon, "Complex permittivity determination from propagation constant measurements," *IEEE Trans. Microwave and Guided Wave Letter*, Vol. 9, No. 2, 1999.
4. Nikawa, Y., M. Chino, and K. Kikuchi, "Measurement of complex permittivity of dielectric material using partially loaded waveguide," *1997 Topical Symposium on Millimeter Waves*, 63–66, 7–8 July, 1997.
5. Zhao, X., C. Liu, and L. C. Shen, "Numerical analysis of a TM010 cavity for dielectric measurement," *IEEE Trans. MTT*, Vol. 40, No. 10, 1951–1959, 1992.
6. Tantot, O., M. C. Moulin, and P. Guillon, "Measurement of complex permittivity and permeability and thickness of multilayered medium by an open-ended waveguide method," *IEEE Trans. IM*, Vol. 46, No. 2, 519–522, 1997.
7. Maser, K. R. and T. Scullion, "Automated detection of pavement layer thickness and subsurface moisture using ground-penetrating radar," TRB paper, 1991.
8. Walker, P. D. and M. R. Bell, "Subsurface permittivity estimation from ground-penetrating radar measurements," *The IEEE 2000 International Radar Conference*, 341–346, 7–12 May, 2000.
9. Osumi, N. and K. Ueno, "Microwave holographic imaging of underground objects," *IEEE Trans. Ant. Prop.*, Vol. AP-33, No. 2, 152–159, Feb. 1985.
10. Caffey, T. W. "A range algorithm for ground penetrating radar," *IGARSS'96, Remote Sensing for a Sustainable Future*, 2023–2026, May 1996.
11. Liu, R., J. Li, et al, "A new model for estimating the thickness and permittivity of subsurface layers from GPR data," *IEE Proceedings on Radar Sonar and Navigation*, Vol. 149, No. 6, 2002.

Comparison between Impulse and Holographic Subsurface Radar for NDT of Space Vehicle Structural Materials

S. I. Ivashov[1], **I. A. Vasiliev**[1], **T. D. Bechtel**[2], **and C. Snapp**[3]

[1]Remote Sensing Laboratory, Bauman Moscow State Technical University, Russia
[2]Enviroscan, Inc./University of Pennsylvania, USA
[3]NASA Johnson Space Center, USA

Abstract— A subsurface holographic radar using a multi-frequency signal has been developed for inspecting dielectric construction materials. The characteristic feature of this device is the ability to obtain one-sided radar soundings/images with a high sensitivity and high resolution (2 cm) in the frequency band of 3.6–4.0 GHz. One promising application of the device is non-destructive evaluation of the heat protection system and other materials on the U.S. Space Shuttle, and proposed crewed exploration vehicle (CEV). The advantages of this continuous-wave holographic radar over traditional impulse subsurface radars are discussed and illustrated by experimental results.

The disastrous loss of the space shuttle Columbia, as well as even more recent dangerous incidents that were thankfully resolved, have aroused interest in possible new methods and devices for non-destructive testing and evaluation of the Space Shuttle Thermal Protection System, the external fuel tank insulating foam, and other materials and structures on the shuttle (see Figure 1), proposed CEV, and other space vehicles. Voids in or under the external tank insulating foam are considered potential sites for "cryopumping" where water seeps in and then evaporates explosively at altitude, pulling the foam from the tank (Figure 2).

One of the possible means for non-destructive testing (NDT) and evaluation of structural materials is subsurface radar. This method is based on the propensity of electromagnetic waves to be reflected at permittivity contrasts. Up to now, the use of radar for NDT has been hindered by

Figure 1: Suspect flaws on the external hydrogen tank coating [1].

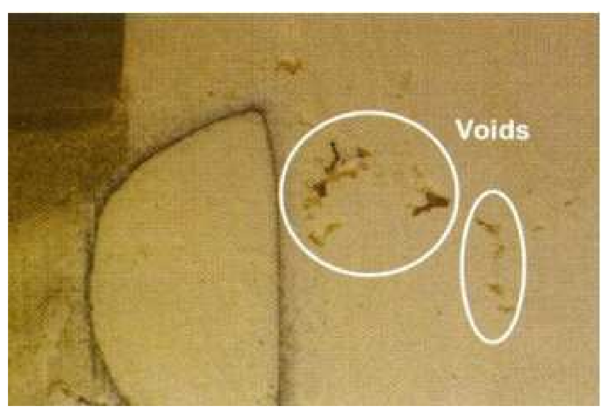


Figure 2: Close-up of voids in the external tank insulating foam [2].

insufficient resolution of available subsurface radars. For many important practical NDT problems, it is sufficient to have a sounding depth in the range of 10 to 20 centimeters (cm), but the spatial resolution should be not less than 2 to 3 cm — beyond the abilities of conventional impulse radar systems. However, taking into account the small required sounding depths, it is possible to use a continuous wave (as opposed to impulse), multifrequency radar signal. This device is reflection mode radar (i. e., the transmitting and receiving parts of the antenna are located on the same side of the sounded surface), however the subsurface images are generated as a plan-view holograph rather than as travel time cross-sections as in impulse radar (Figure 3). The holographic radar system operates in frequency range of 3.6 through 4.0 GHz, and uses 5 working frequencies at two receiver polarizations each [3, 4].

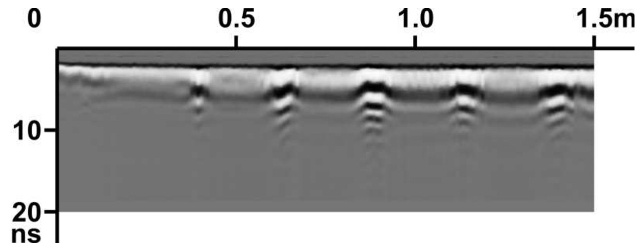


Figure 3: Impulse radar image of spaced metal rebar in a concrete airstrip.

At shallow depths, holographic radar has a distinct advantage in resolution over impulse radar because the radar frequency range can be easily adapted to the demands of a particular NDT task. Another extremely important advantage of this holographic radar technology is the possibility that it can image, without reverberation, dielectric materials that lie above a metal surface. Such materials cannot currently be inspected non-destructively with traditional time-domain impulse radar technology.

For example, Figure 3 depicts an impulse radar cross-section of a reinforced concrete airstrip [5]. Reverberation of pulses between the radar antenna and shallow metal objects obscures the actual location and shape of the reinforcing bars (rebar) which are seen as multiple reflections (often called ghosts or phantoms) of the transmitted impulse signal on the relatively uniform background. Furthermore, the reflected pulses characteristically form images that are segments of arcs or parabolas because of a relatively wide antenna-beam pattern (up to 60° opening angle) of impulse radars.

In this experiment impulse subsurface radar OKO (Logis Ltd, Russia) with central frequency of 1.2 GHz had been used. In Figure 3, the horizontal axis at the image corresponds to distance along the airstrip surface, and the vertical axis corresponds to a time axis, which is related to delay-time

Figure 4: Plastic foam over an aluminum sheet as used for experiments.

of radar signal.

Laboratory experiments with the holographic radar have been conducted to investigate various types of model flaws or anomalies within a rigid 5 centimeter thick foam plastic board overlying an aluminum sheet (Figure 4). The anomalies consisted of voids in the foam, some of which were water-filled. These known defects were imaged using the holographic radar.

The results of the experiments are presented in Figures 5 and 6. Note that the holographic radar produces plan-view or footprint images of the anomalies. The model defects in and under the foam were wetted to varying degrees. Since water has a permittivity of about 80, even small amounts of moisture produce dramatic defect images. The visible "waviness" in these radar images is due to the holographic nature of the radar.

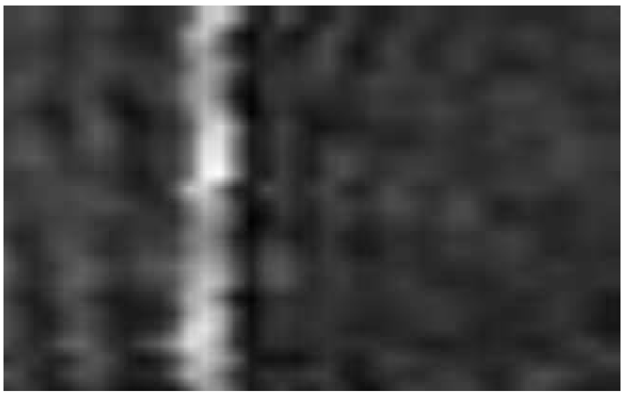


Figure 5: Holographic radar image of a wetted seam in the foam.

Figure 6: Holographic radar image of a damp cloth between the plastic foam and aluminum sheet.

In addition, preliminary experiments with actual Space Shuttle heat protection tiles and adhesive were performed at the NASA Jet Propulsion Laboratory (Figure 7). Dimensions of the tiles were 6 by 6 inches in plan, with thicknesses of 1 to 2 inches. The purpose of the experiment was to determine whether it is possible to detect different types of de-bonding of the tiles from the underlying aluminum sheet. To accomplish this, parts of the aluminum sheet surface were intentionally left without glue (see Figure 8). In this Figure, two round holes (or holidays in the common parlance of technicians) are evident; #1 dry, and #2 with several drops of water. A tile with 2 inch thickness

Figure 7: Different types of Space Shuttle tiles that were used in experiments.

was pressed into the glue, and allowed to cure overnight. Results of subsequent radar scanning of this tile are presented in Figure 9. The water-filled holiday has a dramatic contrast, while the air-filled holiday is more subtle, but still clearly visible. The originally round shapes of the holidays (see Figure 8) have been distorted by pressing the tile into the glue, and the water-filled holiday may appear larger due to seepage of water beyond the original holiday boundary. Although more detailed experiments are needed to understand all of the capabilities of holographic radar for NDT of space vehicle heat protection systems, these initial experiments are extremely promising.

Figure 8: Part of aluminum sheet was covered with glue. Two spots inside were without glue. One of the spots (2) was filled with water.

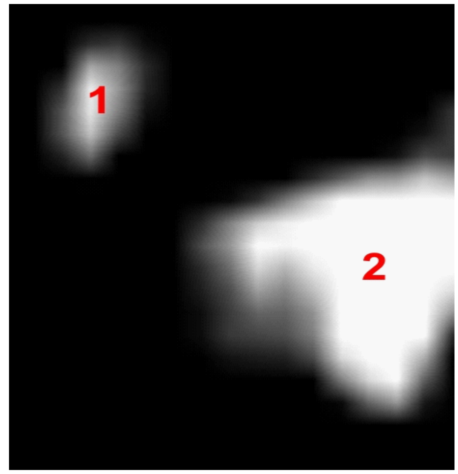


Figure 9: Result of radar scanning of the tile. This tile is visible in right upper corner of Figure 7.

In addition, it is shown that radio-frequency hologram reconstruction allows estimation of the depth of shallowly-buried objects and improves the lateral resolution of images with increasing target depths [6]. Furthermore, for even better resolution of the microwave images, the design of a holographic radar operating at a higher frequency (up to 20 to 30 GHz) is discussed.

ACKNOWLEDGMENT

Experiments with Space Shuttle tiles were conducted with the help and active participation of Dr. T. Lu, NASA Jet Propulsion Laboratory, USA.

Support for this work was provided by the International Science and Technology Center.

REFERENCES

1. *Aviation Week* & *Space Technology*, 28, February 17, 2003.
2. *Aviation Week* & *Space Technology*, 31, April 7, 2003.
3. Vasiliev, I. A., S. I. Ivashov, V. I. Makarenkov, V. N. Sablin, and A. P. Sheyko, "RF band high resolution sounding of building structures and works," *IEEE Aerospace & Electronic Systems Magazine*, Vol. 14, No. 5, 25–28, May 1999.
4. Ivashov, S. I., V. V. Razevig, A. P. Sheyko, I. A. Vasilyev, and T. D. Bechtel, "Holographic radar as a tool for non-destructive evaluation of structural materials," *Proceedings of the 2005 SEM Annual Conference & Exposition on Experimental and Applied Mechanics*, Portland, Oregon, USA, Tuesday, June 7–Thursday, June 9, 2005.
5. Chapursky, V. V., S. I. Ivashov, V. V. Razevig, A. P. Sheyko, I. A. Vasilyev, V. V. Pomozov, N. P. Semeikin, and D. J. Desmond, "Subsurface radar examination of an airstrip," *Proceedings of the 2002 IEEE Conference on Ultra Wideband Systems and Technologies, UWBST'2002*, 181–186, Baltimore, Maryland USA, May 20-23, 2002.
6. Chapursky, V. V., S. I. Ivashov, V. V. Razevig, A. P. Sheyko, and I. A. Vasilyev, "Microwave hologram reconstruction for the RASCAN type subsurface radar," *Proceedings of the Ninth International Conference on Ground Penetrating Radar, GPR'2002*, 520–526, Santa Barbara, California USA, April 29–May 2, 2002.

Analysis of Time Domain Ultra-Wide-Band Radar Signals Reflected by Buried Objects

P. Falorni[1], L. Capineri[1], L. Masotti[1], and C. G. Windsor[2]

[1]Laboratorio Ultrasuoni e Controlli Non Distruttivi, Dipartimento Elettronica e Telecomunicazioni
Università di Firenze, 50139 Firenze, Italy
[2]116, New Road, East Hagbourne, OX11 9LD, UK

Abstract— The aim of this work is the analysis of the signal composition observed in a single radar sweep during an underground investigation with an ultra-wide-band (UWB) radar. The electromagnetic (EM) response of a buried object, the radar pulse spectrum and the antenna set-up, all strongly influence the accuracy of the time of flight estimate. The analysis of the time domain signal will discuss the effects of the antenna coupling with the ground (first arrival pulse from air-soil interface) and the interference of overlapping pulses due to multiple interfaces and multiple reflections. The results of this analysis are based on simulations with parameters characteristic of an investigation of layered medium and signal processing schemes to extract information about soil and buried objects composition will be addressed.

1. INTRODUCTION

The aim of this work is the analysis of the signal composition observed in a single radar sweep during an underground investigation with an UWB radar. The electromagnetic response of a buried object, the radar pulse spectrum and the antenna set-up, all strongly influence the accuracy of the time of flight estimate. The analysis of the time domain signal will discuss the effects of the antenna coupling with the ground (first arrival pulse from air-soil interface) and the interference of overlapping pulses due to multiple interfaces and multiple reflections. The electromagnetic response of this experimental configuration has been already described in previous works (e. g., Dai and Young in [1]). In many practical cases the experimental conditions can be assumed for a linear response and therefore the time domain signals are generated by time convolution between the transmitted current pulse and the characteristic response of the layered medium [7].

In this work the effects described above have been simulated and illustrated by computer modeling. The assumed model considers the propagation in a layered soil and transmitting-receiving antenna placed at different positions above the ground surface. Losses in the medium have been also considered by the complex dielectric constant and multiple reflections in each layer are calculated recursively.

2. ELECTROMAGNETIC MODELING OF LAYERED MEDIA

The adopted model considers a layered media where the layers are defined by their electromagnetic properties — i. e., relative dielectric constant ε_R, magnetic permittivity μ_R conductivity σ — and thickness d. The simple mono-dimensional model assumes a source generating a plane wave with assigned spectrum (E) and placed in a given layer and a receiving antenna placed in a layer that can also be the same of the transmitter.

In each layer it is assumed that the transversal component of the electric field is due to the contribution of the components from the two adjacent layers plus the eventually present transmitting source. Outer layers of the model should be defined as semi-infinite. The calculated solution for the received signal is obtained by a recursive process that returns the EM field spectrum at the receiving antenna position. Inverse Fourier transform is then applied to the received spectrum for obtaining the time domain signal.

At each run, the recursive function propagates the EM field into current layer, than recall itself to propagate the EM field through next layer and into current layer but in the opposite direction. The returned EM field at the antenna position is summed to its current value and returned to the caller function.

The recursion stops if executed into a semi-infinite layer — because no further back-propagation can occur — or if the energy carried by the EM field is lower than a predefined signal-to-noise ratio of the receiving antenna. If the current layer contains the receiving antenna, the recursive function evaluates the EM field at the antenna position and uses it as return value for the caller.

With reference to Figure 1, let assume that the EM field is propagating into layer K in the FORW direction; the recursive function propagates the EM field through layer K, than:

- Evaluates the EM field transmitted into layer $K+1$ (EM-FORW) than recall itself to process layer $K+1$ in the FORW direction using EM-FORW as starting value.
- Evaluates the EM field reflected into layer K (EM-BACKW) than recall itself to process layer K in the BACKW direction using EM-BACKW as starting value.

At each run the recursive function checks the stop conditions and, if necessary, calculates the EM field at the receiving antenna position.

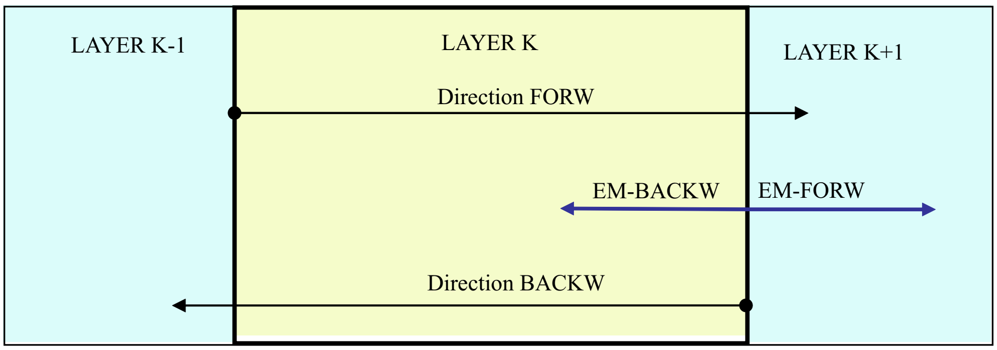


Figure 1: Electromagnetic model of layered media based on the recursive calculation of the propagating electric field.

The electromagnetic modeling has been used to generate the signal shown in Figure 2. A propagation medium composed of a 0.26 m thick layer of sand in air, monostatic antenna placed in air at 0.5 m from the sand ($\varepsilon_R = 3$, $\mu_R = 1$, $\sigma = 7 \times 10^{-3}$ $(\Omega\text{m})^{-1}$) layer. The dashed line is the transmitted pulse with central frequency $f_{\text{central}} = 550\,\text{MHz}$ and $-3\,\text{dB}$ bandwidth of 650 MHz. The solid line is the time domain received signal. The simulation is carried out without superimposed noise on amplitude samples. The sampling frequency used is 6 GHz and the number of time samples is 121; the transmitted pulse has been delayed by 4 ns and the radar acquisition system is configured with a signal to noise ratio of 100 dB. The aliasing in the time domain has been removed setting to zero all the frequency samples having a total delay greater than the simulation time window.

3. ANALYSIS OF PULSE RESPONSE FOR TIME-OF-FLIGHT ESTIMATION

Recalling that the time-of-flight (tof) for an homogenous layer with propagation velocity V and thickness d is defined by: $tof = 2\,\text{d/V}$, the main issue for tof estimation is the finite duration of the transmitted pulse.

The finite duration of the probing pulse introduces an uncertainty because the direct estimate (time differences) deals with wavelets instead of delta functions.

The estimation of tof could be also carried out by using correlation techniques operating on the "mainbang" (first large amplitude reflection from air-soil interface) and the target signal; these methods do not give accurate results mainly because the two signals have been differently modified during propagation. Several works have been published in order to get an accurate estimation of the time domain response by EM modelling of the GPR experiments [1, 6]. The phenomena that modify the transmitted pulse are the propagation characteristics of the layer and overlapping wavelets due to close (comparable to wavelength) interfaces. Furthermore, in the case of a bistatic antenna in contact with soil, the "mainbang" signal is the summation of two signals, one propagating in air and the other propagating into soil [1].

This situation, in general leads to a different shape for the "mainbang" with respect to the signals reflected by a planar target. Hence the "mainbang" is not a good template for accurate tof estimation with correlation methods. A possible approach investigated here is the signal homomorphic deconvolution [5] applied to the summation between a reference signal $r(t)$ and the received signal $s(t)$. Using a bi-static antenna, the measurement of the reference signal can be obtained with the free space response, taking care to avoid saturation phenomena during the analog to digital conversion.

In this work we study the possibility of using the reference signal $r(t)$ to overcome the problems due to the ill-conditioned features in the transformed space, named *cepstral* domain or *quefrency*

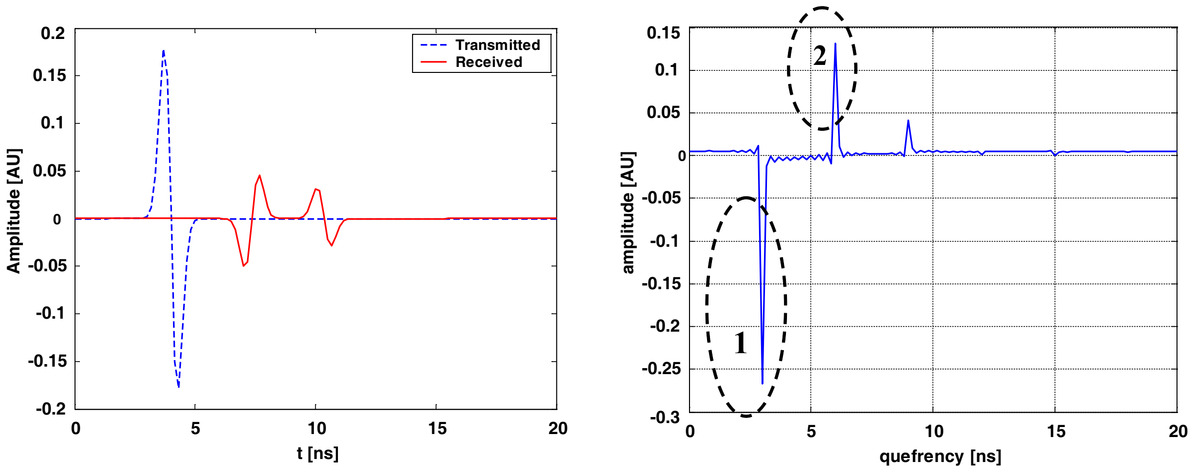


Figure 2: (LEFT) Simulation of the received signals for a propagation medium composed by a 0.26 m thick layer of sand in air, monostatic antenna placed in air at 0.5 m from the sand ($\varepsilon_r = 3$, $\mu_r = 1$, $\sigma = 7\times10^{-3}(\Omega\text{m})^{-1}$) layer. Dashed line: the transmitted Gaussian pulse with central frequency equal to 550 MHz and -3 dB bandwidth of 650 MHz. Solid line: time domain received signal. The simulation is carried out without any superimposed noise on amplitude samples. (RIGHT) Application of the signal deconvolution (CEPSTRUM method) based on a reference signal. Time difference between the two delta-like functions 1 and 2 corresponds to the time of flight relative to the path inside the sand layer. The estimated time of flight is 3 ns which corresponds to 0.259 m of sand layer thickness.

domain. In Figure 2 (RIGHT) it is shown the result of applying the deconvolution method to a simulated radar trace and a reference signal for a simple monostatic antenna setup: the time of flight for the pulse propagating into the 0.26 m thick sand layer can be directly evaluated from the time difference of the peaks 1 and 2 of Figure 2 (RIGHT).

Instead, using a bi-static antenna setup — e.g., with a gap between transmitting and receiving antenna of 0.2 m — for a radar scanning in contact with the ground, we obtain two peaks (Figure 3) related to the "mainbang"; these signals are due to the existence of a double path for direct coupling of transmitting and receiving antenna, one path in air and the other into ground (i.e., for the case of Figure 3, sand). The time difference t_D between the two peaks can also be used to evaluate sand propagation velocity.

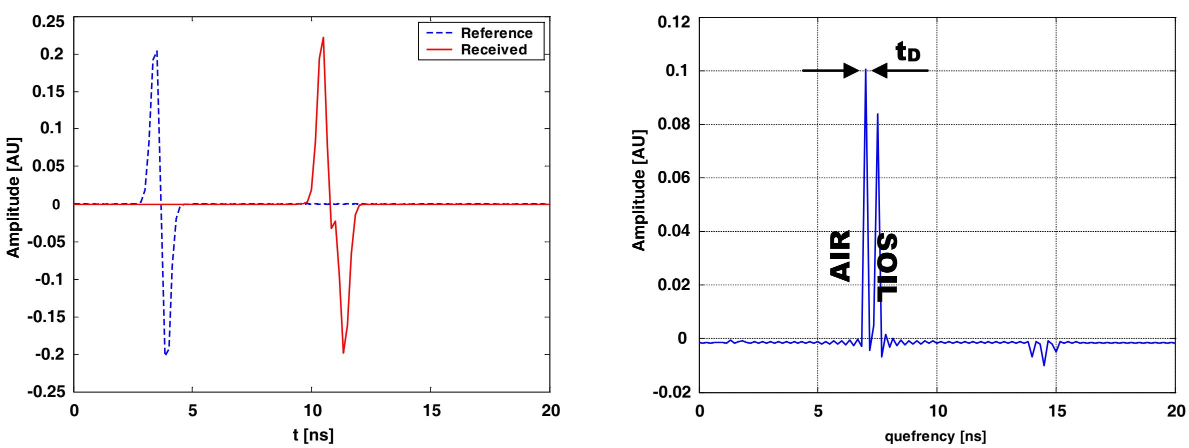


Figure 3: (LEFT) Simulation of the received signal from an infinite layer of sand using an in contact scanning with an UWB radar configured in bistatic mode (0.2 m TX-RX separation). The characteristics of radar and medium are the same as in Figure 2. (RIGHT) The application of the signal deconvolution (CEPSTRUM method) shows two peaks related to the "mainbang"; the first peak corresponds to a wave propagating in the 0.2 m air gap and the second peak corresponds to a wave propagating in the 0.2 m sand gap.

Furthermore, it can be seen in Figure 3 (LEFT) that the received signal is different from the transmitted pulse; in this case, the method was able to separate the *AIR* and *SOIL* signals even for a time delay t_D that is significantly lesser than the pulse duration.

4. AN APPLICATION OF TIME-OF-FLIGHT ESTIMATION TO BURIED OBJECT CHARACTERIZATION

The depth, lateral position and radius of a large buried pipe in a soil with unknown propagation velocity is a challenging problem that can be solved with signal processing methods based on the time-of-flight hyperbolic equation [2–4]:

$$t\hat{o}f_i = tof_i + t_{MB} = \frac{2}{V}\left(\sqrt{(y_i - Y_0)^2 + Z_0^2} - R\right) + t_{MB} \tag{1}$$

where Y_0, Z_0 are the coordinate of the pipe centre, R is the pipe radius ($R > \lambda_{\text{central}} = V/f_{\text{central}}$), V is the medium propagation velocity; tof_i is the time-of-flight measured at the lateral position y_i of a monostatic antenna. According to the analysis of the inversion of the Equation (1) [8], the estimation of the unknown parameters is strongly affected by errors in the *tof* estimation. Moreover, the uncertainty on the estimation of the term t_{MB}, which represents the delay time of the "mainbang" signal, directly reflects on the $\hat{tof}_i$

With the deconvolution method the estimation of the tof_i is straightforward and avoids the problem of estimating t_{MB}, which is rather cumbersome even with instrument calibration procedures; anyway its accuracy is limited by the finite pulse duration.

REFERENCES

1. Dai, R. and C. T. Young, "Transient fields of a horizontal electric dipole on a multilayered dielectric medium," *IEEE Trans. on Ant. And Prop.*, Vol. 45, No. 6, 1023–1031, June 1997.
2. Falorni, P., L. Capineri, L. Masotti, and G. Pinelli, "3-D radar imaging of buried utilities by features estimation of hyperbolic diffraction patterns in radar scans," *Tenth International Conference on Ground Penetrating Radar*, Delft, The Netherlands, 21–24 June, 2004.
3. Grandjean, G., J. C. Gourry, and A. Bitri, "Evaluation of GPR techniques for civil-engineering applications: study on a test site," *Journal of Applied Geophysics*, Vol. 45, 141–156, 2000.
4. Shihab, S., W. Al-Nuaimy, and A. Eriksen. "Radius estimation for subsurface cylindrical objects detected by ground penetrating radar," *Tenth Intern. Conference on Ground Penetrating Radar*, 319–322, Delft, The Netherlands, June 21–24, 2004.
5. Oppenheim, A. V., R. W. Shafer, and T. G. Stockham, "Non linear filtering of multiplied and convolved signals," *Proceedings of the IEEE*, Vol. 56, No. 8, 1264–1291, 1968.
6. Van der Kruk, J., E. C. Slob, and J. T. Fokkema, "Background of ground-penetrating radar measurements," *Geologie en Mijnbouw*, Vol. 77, 177–188, 1999.
7. Fokkema J. T., "Analysis of georadar reflection responses," *2nd Intern. Workshop on Advanced GPR*, 1–4, Delft, The Netherlands, 14–16 May, 2003.
8. Windsor, C., L. Capineri , P. Falorni, S. Matucci, and G. Borgioli, "The estimation of buried pipe diameters using ground penetrating radar," *Insight*, Vol. 47, No. 7, 394–399, July 2005.

Bifurcations of Double Bullet Complexes in Dissipative Systems

N. Akhmediev[1] **and J. M. Soto-Crespo**[2]

[1]Optical Sciences Group, Research School of Physical Sciences and Engineering
The Australian National University, Canberra ACT 0200, Australia
[2]Instituto de Óptica, C.S.I.C., Serrano 121, 28006 Madrid, Spain

Abstract— We show that double bullet complexes in nonlinear dissipative systems can exist in various forms. The most frequent ones have constant energy and rotate in the transverse plane. Others consist of an oscillating structure which rotates while the separation between the two bullets changes periodically, and so does its energy. Several complexes can coexist stably for the same set of values of the parameters. Transitions between them occur in form of bifurcations.

1. INTRODUCTION

The dissipative soliton concept is a fundamental extension of that for solitons in conservative and integrable systems. It includes ideas from three major sources, viz. standard soliton theory developed since the 1960s, ideas from nonlinear dynamics theory, and Prigogine's ideas of systems far from equilibrium [1]. These are basically three sources and three component parts of this novel paradigm. Physically speaking, the major part of standard soliton theory is the notion of the balance between dispersion and nonlinearity that allows stationary localized solutions to exist. For dissipative systems, we have to admit that the major balance is between gain and loss. This is necessary for the solitons to be stationary objects [2]. Nonlinear dynamics inspires us with the idea of soliton bifurcations and the chaotic evolution of solitons. Finally, the theory of systems far from equilibrium tells us that solitons are self-organized formations requiring a continuous supply of matter and energy. As soon as that supply finishes, a dissipative soliton ceases to exist. These are basic cornerstones of the powerful concept of dissipative soliton.

Optical bullets in dissipative systems are a particular example of dissipative soliton. These are (3+1) dimensional spatio-temporal formations that can be generated by wide aperture passively mode-locked lasers. In many respects, their properties are similar to (1+1) dimensional localized structures in dissipative systems. However, additional dimensions can add new properties that are not known for low dimensional systems.

We showed, recently [3], that dissipative optical bullets can be combined into double bullet complexes (DBC). These complexes exist in certain regions of the parameter space that are not the same as for single optical bullets. Double bullet complexes can be considered as new solutions that have unique properties different from those for optical bullets. The separation between the bullets is an additional parameter of the solution that does not exist in the case of an optical bullet. This parameter can be a constant that depends on the parameters of the system. It can also oscillate. This happens as a bifurcation when we change the parameters of the system. Thus, the region of existence of double bullet complexes is subdivided into two parts. The boundary between them separates stationary (but rotating) DBC from oscillating ones. The energy of DBC is a constant in the former case but is an oscillating function in the latter one. In this paper, we give examples of bifurcations from stationary DBCs to oscillating ones.

2. NUMERICAL SIMULATIONS

Our numerical simulations are based on an extended complex cubic-quintic Ginzburg-Landau equation (CCQGLE) model. This model includes cubic and quintic nonlinearities of dispersive and dissipative types, and we have added transverse operators to take into account spatial diffraction. The normalized propagation equation reads:

$$i\psi_z + \frac{D}{2}\psi_{tt} + \frac{1}{2}\psi_{xx} + \frac{1}{2}\psi_{yy} + |\psi|^2\psi + \nu|\psi|^4\psi = i\delta\psi + i\epsilon|\psi|^2\psi + i\beta\psi_{tt} + i\mu|\psi|^4\psi. \tag{1}$$

The optical envelope ψ is a complex function of four real variables $\psi = \psi(t, x, y, z)$, where t is the retarded time in the frame moving with the pulse, z is the propagation distance, and x and y are the two transverse coordinates. Eq. (1) is written in normalized form. The left-hand-side contains the conservative terms, *viz.* $D = +1(-1)$ which is for the anomalous (normal) dispersion propagation regime and ν which is the saturation coefficient of the Kerr nonlinearity. The right-hand-side includes all dissipative terms: δ, ϵ, β and μ are the coefficients for linear loss (if negative), nonlinear gain (if positive), spectral filtering and saturation of the nonlinear gain (if negative), respectively.

This distributed equation could be applied to the modeling of a wide-aperture laser cavity in the short pulse regime of operation. The model includes the effects of two-dimensional transverse diffraction of the beam, longitudinal dispersion of the pulse and its evolution along the cavity. Dissipative terms describe the gain and loss of the pulse in the cavity. Higher-order dissipative terms are responsible for the nonlinear transmission characteristics of the cavity which allow, for example, passive mode-locking. This equation is a natural extension of the one-dimensional complex cubic-quintic Ginzburg-Landau equation (CCQGLE).

We have solved Eq. (1) using a split-step Fourier method. Thus, the second-order derivative terms in x, y and t are solved in Fourier space. Consequently, we apply periodic boundary conditions in x, y and t. All other linear and nonlinear terms in the equation are solved in real space using a fourth-order Runge-Kutta method. Most of the simulations presented in the paper were done using a numerical grid of 256 points in each of the three dimensions x, y and t. We used various values of step sizes along the spatial and temporal domains to check that the results do not depend on the mesh intervals, thus avoiding any numerical artifacts. A typical numerical run presented in this work takes from several hours to several days on a standard modern PC.

In the (1+1)D case, the cubic-quintic CGLE admits various types of soliton solutions [4]. Moreover, several of them can exist for the same set of parameters [5]. Not all of them are necessarily stable. The stability is controlled by the parameters of the equation and by the choice of the soliton branch. In this paper we deal with (3+1)D solitons or optical bullets. We showed in Refs. [6, 7] that, in the case of anomalous dispersion as well as for normal dispersion, Eq. (1) admits 3D dissipative solitons, i.e., optical bullets. We show now that this equation also has stable pulsating solutions in both regimes of dispersion. Pulsating solutions turn out to be double bullet complexes, as we have found. The main task is to find a set of parameters where stationary or pulsating solitons exist. In general, we fix five of the parameters, namely $D, \mu, \nu, \delta, \beta$, and change ϵ when looking for stable localized solutions. The initial conditions must be localized when we are looking for localized structures. Their exact shape is relevant but plays a secondary role if only one type of optical bullet exists for a given set of parameters. The shape becomes highly important when several stable solutions coexist. Once a certain kind of localized solution is found for a given set of equation parameters, it can serve as the initial condition for finding solutions at other nearby values of the parameters. By moving slowly in the parameter space, we are able to determine the regions of soliton existence in a relatively easy way.

The natural control parameter of the solution as it evolves is the total energy Q, given by the three dimensional integral of $|\psi|^2$ over x, y and t:

$$Q(z) = \int_{-\infty}^{\infty} |\psi(x, y, t, z)|^2 \mathrm{d}x\mathrm{d}y\mathrm{d}t. \tag{2}$$

For a dissipative system, the energy is not conserved but evolves in accordance with the so-called balance equation [8]. If the solution stays localized, the energy evolves but remains finite. Furthermore, when a stationary solution is reached, the energy Q converges to a constant value. When the optical field spreads out, the energy tends to infinity. Another possibility is that the solution dissipates, and then the energy goes to zero. However, if the optical bullet is a pulsating one, the energy Q is an oscillating function of z. We observed all these scenarios in numerical simulations. The stationary bullets are radially symmetric objects in the (x, y)-plane.

3. REGIONS OF EXISTENCE OF OPTICAL BULLETS AND DOUBLE BULLET COMPLEXES

Optical bullets do exist in finite regions in the parameter space. In the search for optical bullets, we restricted ourselves to varying only two parameters, (namely ν and ϵ) while fixing the four others [3]. The results of this search are shown in Fig. 1(a). In fact, the red area in Fig. 1(a) a shows a 2-D section of the complete region in the 6-D parameter space where we found stable single optical bullets.

The smaller blue region in Fig. 1(a) is the one which admits stable double bullet complexes. This region resides completely inside the region of existence of stationary single bullets. This is hardly surprising because, to have a stable composite state of two bullets, one would expect that each of them should be stable.

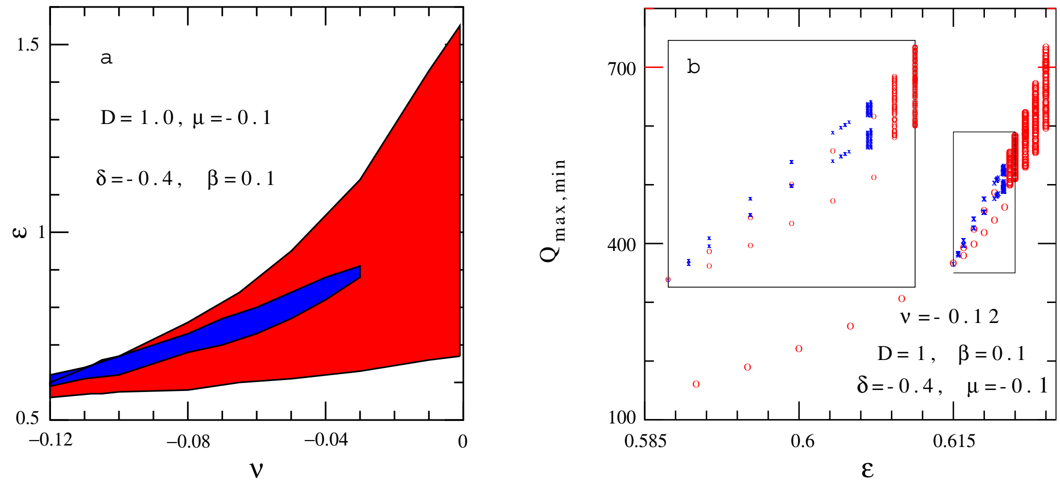


Figure 1: (a) Region of existence of stable stationary optical bullets in the (ϵ, ν)-plane (red area). The blue area shows the region of existence of stable double bullet complexes. (b) Bifurcation diagram for different set of parameters. A second bifurcation results in continuous values for Q. The differently colored points, magnified in the inset, represent different coexisting DBCs.

A consequence of this coexistence is that, at certain conditions, there can be transitions from one type of solution to another. In principle, the two regions may overlap only for the set of the parameters that we have chosen. Changing some of the other 4 parameters of the equation may separate the two regions.

This blue region is not homogeneous. There are several types of DBC in this blue region. Each type of double bullet complex occupies a smaller finite region. The structure of these regions is highly complicated. Transitions from one subregion to another take the form of bifurcations. In Fig. 1(b), we present the maxima and minima of the energy, Q, versus ϵ for the case $\nu = -0.12$ which is the most left point at Fig. 1(a). The interval of ϵ is chosen to cover the whole blue region at $\nu = -0.12$, i.e., for DBC. The energy Q is fixed at each point of ϵ below the value 0.615. This means that double bullet complexes in this region are stationary but rotating objects. Its 2D profile in space stays fixed although its orientation in space changes. However, above the value $\epsilon = 0.615$ the energy Q splits into two branches.

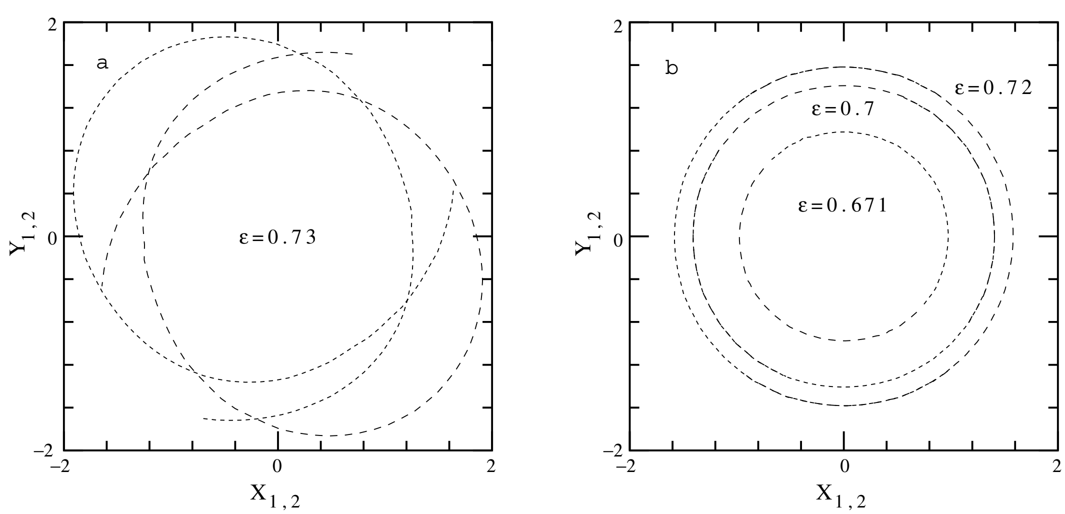


Figure 2: Evolution of the two maxima of the field amplitude profiles. In case (a) the separation between the two maxima oscillates. In case (b) the separation is constant.

These correspond to the lower and upper limits of the oscillations of the energy as the DBC propagates. Oscillations are related to the dynamical pulsations of the double bullet complex. Namely, the separation between the two bullets in the complex changes periodically.

To see this, we show, in Fig. 2(a) the trajectories of the maxima of the optical field. There are two of them on the (x, y) plane, and they occur at $t = 0$. They roughly indicate the position of the two bullets in the soliton complex. The two trajectories are shown by the dotted and dashed lines respectively. The trajectories look more like ellipses than circles, thus clearly indicating the periodic change in the relative separation between the two maxima. They also indicate that in the three-dimensions (defined by the space-time variables x, y, t), the two bullets always appear to be separated. In contrast, the trajectories for the DBC with constant energy are perfect circles indicating that the DBC is rotating but no any oscillations occur (see Fig. 2(b)).

Eq. (1) is translationally invariant in space. Thus, solutions can travel in (x, y) plane with constant velocity. In combination with rotations of DBC, this creates a "boomerang" type of motion. An example of such motion is shown in Fig. 3. The plot 3a shows the trajectories of the two maxima in (x, y) plane while the plot 3b shows oscillations of the energy Q for this motion. These oscillations are relatively small indicating that rotation in the moving frame is almost circular.

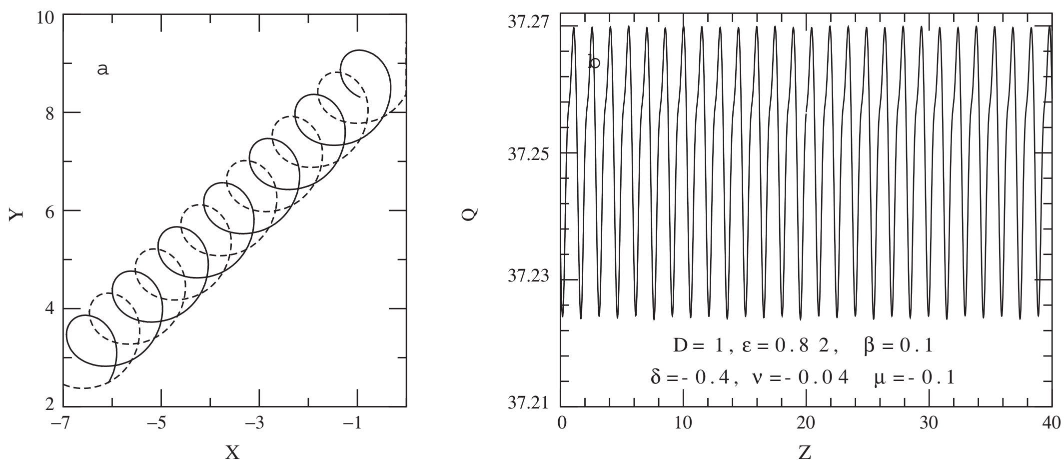


Figure 3: (a) Evolution of the two maxima of the field amplitude profiles for the DBC with rotation and translational motion. (b) Energy Q versus z for this "bumerang" type of motion.

The last value of ϵ in Fig. 1(b) where single periodic pulsations of DBC occur, is around 0.62. The energy Q versus z for these pulsations are shown in Fig. 4(a) (lower curve). The lower and upper values of Q for these oscillations are indicated by the two red horizontal dashed lines. When we further increase the parameter ϵ, simple single frequency oscillations of the DBC are transformed into two-frequency quasi-periodic oscillations. The energy Q versus z for these pulsations are shown in Fig. 4 (upper curve). The upper and lower values of the energy for these oscillations are not constants but periodic functions as well. They are shown by red lines in Fig. 4. This new bifurcation when additional period appears is shown in Fig. 1(b) with a transition from the two branched curve for Q to the bands of Q-values (red vertical stripes).

The two periods of oscillation in the composite solution can be related to a complicated motion inside the complex. One of the motions is the oscillation of the two bullets relative to each other in space, and the other motion involves the pulsations of the shape in the t-domain. In a nonlinear problem, these two motions are inseparable and each period can be found in either of the above-mentioned motions. Both periods are revealed in the upper $Q(z)$ vs z curve of Fig. 4.

More frequencies can appear when we choose other parameters of the system. Fig. 4(b) shows a particular example of such multi-frequency motion. The parameters chosen for this simulation are inside of the blue area in Fig. 1(a). The appearance of each new periodic component in the motion can be attributed to a bifurcation. Thus, the blue area consists of a multiplicity of bifurcations. However, to plot the boundaries for all of them would require an incredibly large amount of numerical work.

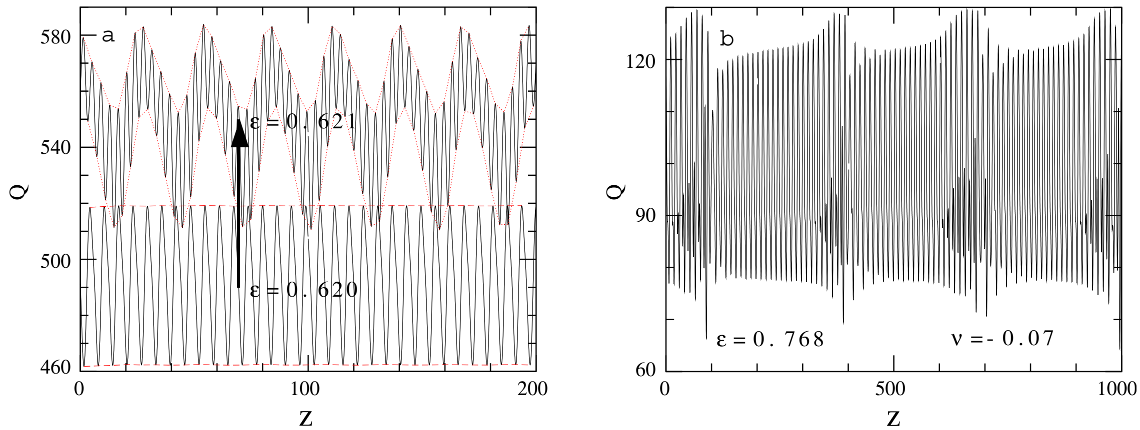


Figure 4: (a) Energy, Q, vs. propagation distance z of a double bullet complex for several values of ϵ. Transition from a single period of oscillations (lower curve) to a double period (upper curve) is clearly seen. The two periods are incommensurate, thus creating a continuous band of Q values in the previous plot. Red dashed lines show the envelopes of the basic frequency of oscillations. (b) An example of multi-frequency oscillations.

ACKNOWLEDGMENT

The work of J. M. S. C. was supported by the M. C. y T. under contracts BFM2003-00427 and FIS2006-03376. N. A. acknowledges support from the Australian Research Council. The authors are grateful to Dr. Ankiewicz for a critical reading of the manuscript.

REFERENCES

1. Nicolis, G. and I. Prigogine, *Self Organization in Nonequilibrium Systems — From Dissipative Structures to Order through Fluctuations*, John Wiley and sons, New York, 1977.
2. *Dissipative solitons*, Editors: N. Akhmediev and A. Ankiewicz, Springer, Heidelberg, 2005.
3. Soto-Crespo, J. M., N. Akhmediev, and P. Grelu, "Optical bullets and double bullet complexes in dissipative systems" *Phys. Rev. E*, in print, 2006.
4. N. Akhmediev, J. M. Soto-Crespo, and G. Town, "Pulsating solitons, chaotic solitons, period doubling, and pulse coexistence in mode-locked lasers: Complex Ginzburg — Landau equation," *Phys. Rev. E*, Vol. 63, 056602, 2001.
5. Soto-Crespo, J. M., N. Akhmediev, and K. Chiang, "Simultaneous existence of a multiplicity of stable and unstable solitons in dissipative systems," *Phys. Lett. A.*, Vol. 291, 115–123, 2001.
6. Grelu, P., J. M. Soto-Crespo, and N. Akhmediev, "Light bullets and dynamic pattern formation in nonlinear dissipative systems," *Optics Express*, Vol. 13, No. 23, 9352–9360, 2005.
7. Soto-Crespo, J. M., P. Grelu, and N. Akhmediev, "Optical bullets and 'rockets' in nonlinear dissipative systems and their transformations and interactions," *Optics Express*, Vol. 14, No. 9, 4013–4025, 2006.
8. Akhmediev, N. and A. Ankiewicz, *Solitons, Nonlinear Pulses and Beams*, Chapman & Hall, London, 1997.

Interaction of Solitary Waves Governed by a Controlled Subcritical Ginzburg-Landau Equation

Y. Kanevsky and A. A. Nepomnyashchy
Technion-Israel Institute of Technology, Israel

Abstract— We consider the influence of a global feedback control which acts on an oscillatory system governed by a subcritical Ginzburg-Landau equation. Exact solutions corresponding to solitary-wave and spatially periodic solutions are obtained. A generalized variational approach is used for studying the indirect interaction between solitary waves caused by the global control. The basic types of dynamics are described.

One of the subjects that started recently attracting growing attention of the researchers working in the area of nonlinear dynamics and pattern formation is an *active feedback control* of pattern forming systems. The aim of feedback control is to achieve the desirable dynamics or a particular pattern. It has been applied to Rayleigh-Bénard convection [1, 2], Marangoni convection [3–5], contact line instability in thin liquid films [6–8], catalytic reactions [9–11], and crystal growth [12]. The effect of feedback control of supercritical oscillatory instabilities was investigated in [13–18]. The possibility of the suppression of a subcritical oscillatory instability by means of feedback control has been demonstrated in [19]. In this paper we present the analytical results of feedback control applied to systems with a subcritical oscillatory instability.

We consider the subcritical complex Ginzburg-Landau equation (CGLE) under a feedback control

$$\partial_t A = A + (1 + ib)A_{xx} - (-1 + ic)|A|^2 A + K(A)A. \tag{1}$$

Feedback control is imposed by adding a term $K(A)A$ to the right-hand side of the Equation (1), with a control functional $K(A)$ of the form:

$$K(A) = -p \max_x |A|. \tag{2}$$

We look for solutions in the form

$$A(x, t) = R(x, t)e^{i\theta(x,t)}, \tag{3}$$

and thus the control functional can be written as $K(A) = -p \max_x R$. Substituting (3) into the Equation (1) and denoting $\max_x R \equiv R_{max}$, we obtain the following system of two real equations

$$\begin{aligned} \partial_t R &= R_{xx} - R\theta_x^2 - b(2\theta_x R_x + R\theta_{xx}) + R^3 + (1 - pR_{max})R, \\ R\partial_t \theta &= b(R_{xx} - R\theta_x^2) + (2\theta_x R_x + R\theta_{xx}) - cR^3. \end{aligned} \tag{4}$$

The system of Equations (4) has the following pulse solutions:

$$R(x, t) = \frac{C}{\cosh \kappa x}, \quad \theta(x, t) = \gamma \ln \cosh \kappa x - \Omega t, \tag{5}$$

where

$$\begin{aligned} C &= \frac{p}{2(1-\alpha)} \left[1 \pm \sqrt{1 - 4(1-\alpha)/p^2}\right], \\ \alpha &= \frac{1}{12(1+b^2)} \left[\sqrt{9(bc-1)^2 + 8(b+c)^2} - 3(bc-1) + 4b(b+c)\right], \\ \gamma &= \frac{1}{b+c} \left[6\alpha(1+b^2) + 3(bc-1) - 2b(b+c)\right], \\ \kappa^2 &= C^2 \frac{1}{3\gamma} \frac{b+c}{1+b^2}, \\ \Omega &= b\kappa^2 - \gamma\kappa^2 + cC^2. \end{aligned} \tag{6}$$

For $\alpha < 1$ the two solutions for C exist for $p > 2\sqrt{1-\alpha}$ and the linear growth rate $\mu = 1 - pR_{\max} < 0$. Therefore, the solutions are stable in this region of parameters, or, in other notation, for $\{-b - 3\sqrt{b^2+1} < c < -b + 3\sqrt{b^2+1}\}$. For $\alpha > 1$ there is only one solution for C and there holds $\mu = 1 - pR_{\max} > 0$. Thus, the solution is unstable there.

There exists also another type of solutions of the system of Equations (4). These solutions are called small-amplitude periodic solutions and they are obtained in the limit $b, c \to \infty$. In the case $p = 0$ (uncontrolled case) and $c/b < 0$, these solutions were derived in [20] in the form:

$$R(x,t) = \sqrt{\frac{2\gamma(t)}{[2-m(t)]\beta}}\, \mathrm{dn}\left(\sqrt{\frac{\gamma(t)}{2-m(t)}}\; x\right), \tag{7}$$

where $\beta = -c/b > 0$, $\gamma(t) > 0$ and $0 \le m(t) \le 1$. The boundary of existence for these solutions is $\beta > 4$ [20]. With the feedback control imposed, we can suppress the limit $\beta > 4$. The obtained solutions exist for all values of $\beta > 0$. One branch of the solutions is stable for all $\beta > 0$, while another branch of the solutions is stable for $\beta > 4$ and has regions of instability for $0 < \beta \le 4$.

The similar periodic solutions are obtained in the limit $b, c \to \infty$ for the case $c/b > 0$:

$$R(x,t) = \sqrt{\frac{-2\gamma(t)m(t)}{[1+m(t)]\beta}}\, \mathrm{sn}\left(\sqrt{\frac{-\gamma(t)}{1+m(t)}} x\right), \tag{8}$$

where $\beta = c/b > 0$, $\gamma(t) < 0$ and $0 \le m(t) \le 1$. The uncontrolled ($p = 0$) solutions are unstable for small values of β. Further, with the increasing of β there appear regions of stability that depend on m. However, the applied feedback control can eliminate these regions of stability even for small values of p.

In order to understand the qualitative behavior of solutions of the CGLE for arbitrary values of b and c, we consider a model of nonlinear evolution using an analytical approach, based on the variational principle, developed for the investigation of the dynamics of stable spatiotemporal solitons (see the review paper [21]). The variational method was extended for treatment of complex dissipative systems by [22] and [23]. We have applied this extended variational method to the Equation (1), with a control functional of the form (2), and used the following ansatz:

$$R_j(x,t) = \frac{C_j(t)}{\cosh[\kappa_j(t)x]}, \quad \theta_j(x,t) = \gamma_j(t)\ln\cosh[\kappa_j(t)x] + \phi_j(t),\ j = 1, 2 \tag{9}$$

for each of the solitary wave. A computer test was provided for the investigation of interaction of two pulse-like solutions according to the obtained variational model

$$\begin{aligned}
\dot{C}_1 &= \frac{1}{9} C_1 \left[9 - 9p\max\{C_1, C_2\} + 8C_1^2 - (\gamma_1^2 + 6b\gamma_1 + 7)\kappa_1^2\right], \\
\dot{\kappa}_1 &= \frac{4}{9}\kappa_1\left[C_1^2 - (3b\gamma_1 + 2 - \gamma_1^2)\kappa_1^2\right], \\
\dot{\gamma}_1 &= \frac{2}{3} C_1^2(c - \gamma_1) + \frac{2}{3}\kappa_1^2(2b + 2b\gamma_1^2 - \gamma_1 - \gamma_1^3), \\
\dot{C}_2 &= \frac{1}{9} C_2 \left[9 - 9p\max\{C_1, C_2\} + 8C_2^2 - (\gamma_2^2 + 6b\gamma_2 + 7)\kappa_2^2\right], \\
\dot{\kappa}_2 &= \frac{4}{9}\kappa_2\left[C_2^2 - (3b\gamma_2 + 2 - \gamma_2^2)\kappa_2^2\right], \\
\dot{\gamma}_2 &= \frac{2}{3} C_2^2(c - \gamma_2) + \frac{2}{3}\kappa_2^2(2b + 2b\gamma_2^2 - \gamma_2 - \gamma_2^3).
\end{aligned} \tag{10}$$

Equations for $\dot{\phi}_j$ do not influence the dynamics and are not written here.

The following possible regimes were obtained:

- existence of a single pulse in the region $\{-b - 3\sqrt{b^2+1} < c < -b + 3\sqrt{b^2+1}\}$, see Figure 1;
- coexistence of two pulses, see Figure 2;
- competition of two pulses, see Figure 3.

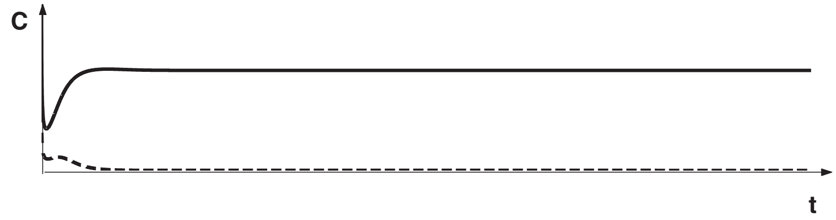


Figure 1: Single pulse-like solution regime; $b = 10$, $c = 3$, $p = 2$.

Figure 2: The coexistence of two pulse-like solutions; $b = -1$, $c = 7$, $p = 1$.

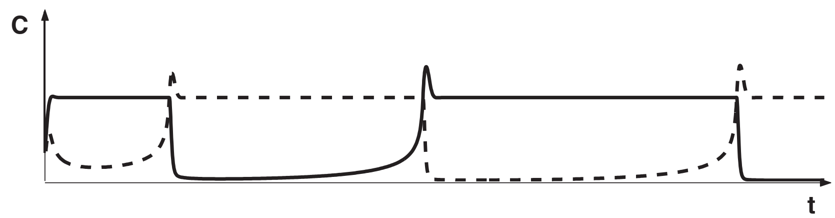


Figure 3: The competition of two pulse-like solutions; $b = 10$, $c = 22$, $p = 2$.

ACKNOWLEDGMENT

The work has been supported by the Israel Science Foundation.

REFERENCES

1. Tang, J. and H. H. Bau, "Stabilization of the no-motion state in Rayleigh-Bernard convection through the use of feedback-control," *Phys. Rev. Lett.*, Vol. 70, No. 12, 1795–1798, 1993.
2. Howle, L. E., "Active control of Rayleigh-Benard convection," *Phys. Fluids*, Vol. 9, No. 7, 1861–1863, 1997.
3. Bau, H. H., "Control of Marangoni-Bernard convection," *Int. J. Heat Mass Transfer*, Vol. 42, No. 7, 1327–1341, 1999.
4. Or, A. C., R. E. Kelly, L. Cortelezzi, and J. L. Speyer, "Control of long-wavelength Marangoni-Benard convection," *J. Fluid Mech.*, Vol. 387, 321–341, 1999.
5. Or, A. C. and R. E. Kelly, "Feedback control of weakly nonlinear Rayleigh-Benard-Marangoni convection," *J. Fluid Mech.*, Vol. 440, 27–47, 2001.
6. Grigoriev, R. O., "Control of evaporatively driven instabilities of thin liquid films," *Phys. Fluids*, Vol. 14, No. 6, 1895–1909, 2002.
7. Grigoriev, R. O., "Contact line instability and pattern selection in thermally driven liquid films," *Phys. Fluids*, Vol. 15, No. 6, 1363–1374, 2003.
8. Garnier, N., R. O. Grigoriev, and M. F. Schatz, "Optical manipulation of microscale fluid flow," *Phys. Rev. Lett.*, Vol. 91, No. 5, 054501, 2003.
9. Bertram, M. and A. S. Mikhailov, "Pattern formation in a surface chemical reaction with global delayed feedback," *Phys. Rev. E*, Vol. 63, No. 6, 066102, 2001.
10. Bertram, M. and A. S. Mikhailov, "Pattern formation on the edge of chaos: Mathematical modeling of CO oxidation on Pt(100) surface under global delayed feedback," *Phys. Rev. E*, Vol. 67, No. 3, 036207, 2003.

11. Beta, C., M. Bertram, A. S. Mikhailov, H. H. Rotermund, and G. Ertl, *Phys. Rev. E*, Vol. 67, No. 4, 046224, 2003.
12. Nepomnyashchy, A. A., A. A. Golovin, V. Gubareva, and V. Panfilov, "Global feedback control of a long-wave morphological instability," *Physica D*, Vol. 199, No. 1–2, 61–81, 2004.
13. Battogtokh, D. and A. Mikhailov, *Physica D*, Vol. 90, No. 1-2, 84–95, 1996.
14. Battogtokh, D., A. Preusser, and A. Mikhailov, *Physica D*, Vol. 106, No. 3–4, 327–362, 1997.
15. Kawamura, Y. and Y. Kuramoto, "Onset of collective oscillation in chemical turbulence under global feedback," *Phys. Rev. E*, Vol. 69, No. 1, 016202, 2004.
16. Echebarria, B. and A. Karma, "Spatio-temporal control of cardiac alternans," *Chaos*, Vol. 12, No. 3, 923–930, 2002.
17. Montgomery, K. A. and M. Silber, "Feedback control of traveling wave solutions of the complex Ginzburg-Landau equation," *Nonlinearity*, Vol. 17, No. 6, 2225–2248, 2004.
18. Kolodner, P. and G. Flätgen, "Spatial-feedback control of dispersive chaos in binary-fluid convection," *Phys. Rev. E*, Vol. 61, No. 3, 2519–2532, 2000.
19. Golovin, A. A. and A. A. Nepomnyashchy, "Feedback control of subcritical oscillatory instabilities," *Phys. Rev. E*, Vol. 73, No. 4, 046212, 2006.
20. Schöpf, W. and L. Kramer, "Small-amplitude periodic and chaotic solutions of the complex Ginzburg-Landau equation for a subcritical bifurcation," *Phys. Rev. Lett.* Vol. 66, No. 18, 2316–2319, 1991.
21. Malomed, B. A., "Variational methods in nonlinear fiber optics and related fields," in *Progress in Optics 43*, Elsevier Science B. V., 69–191, 2002.
22. Chávez Cerda, S., S. B. Cavalcanti, and J. M. Hickmann, "Nonlinear dissipative pulse propagation," *Eur. Phys. J. D*, Vol. 1, No. 3, 313–316, 1998.
23. Skarka, V. and N. B. Aleksić, "Stability criterion for dissipative solitons of the one-, two-, and three-dimensional complex cubic-quintic Ginzburg-Landau equations," *Phys. Rev. Lett.*, Vol. 96, No. 1, 013903, 2006.

Chirped Self-similar Pulse Propagation in Cubic-quintic Media

K. Senthilnathan[1], Qian Li[1], P. K. A. Wai[1], and K. Nakkeeran[2]

[1]Photonics Research Center, Department of Electronic and Information Engineering
The Hong Kong Polytechnic University, Hung Hom, Hong Kong, China
[2]Department of Engineering, Fraser Noble Building, King's College
University of Aberdeen, Aberdeen AB24 3UE, UK

Abstract— We consider nonlinear propagation of optical pulses in a cubic-quintic nonlinear medium wherein the pulse propagation is governed by the generalized nonlinear Schrödinger equation with varying dispersion, nonlinearity and gain/loss. Using the self-similar analysis, we present the generation of chirped bright solitons in the anomalous dispersion regime as well as the normal dispersion regime under the influence of both cubic and quintic nonlinearity. We also find the necessary and sufficient condition for the existence of the resulting self-similar solitary pulses through the physical parameters of the governing systems. Furthermore, our results show that the chirped solitons can be nonlinearly compressed cleanly and efficiently even in the presence of gain/loss.

1. INTRODUCTION

The field of nonlinear optics has blossomed and is undergoing a new revolution in recent years. The nonlinear optical response is now a key element for new emerging technologies. This is particularly true for soliton and other types of nonlinear pulse transmission in optical fibres/nonlinear materials, since this form of light propagation can be used to realize the long-held dream of very high capacity dispersion-free communications. In the recent past, it has been proven beyond doubt that solitons do exist not only in optics but also in many areas of science. Solitons that exist in optics, called "optical solitons", have been drawing a greater attention among the scientific community as they seem to be the right candidates for transferring information across the world through optical fibres [1–3]. More recently, because of the increasing bandwidth demand being put on optical fibre communications (OFC), the emphasis is on producing a source of ultrashort optical pulses that can be used in high-bit rate and long distance optical communication systems. The generation of ultrashort pulses has always been of great scientific and technological interest. Ultrashort pulses in the near infrared spectral region are required for many applications. For instance, in telecommunication applications, pulses with shorter duration are required as the per-channel capacity of communication systems increases beyond 160 Gb/s. At this moment, there are no electronic techniques that can produce 160 Gb/s electrical pulses, so short pulse optical sources are becoming a key technology. Not only do these ultrashort pulses have great potential in OFC, but they also play an indispensable role in ultrafast physical process, infrared time-resolved spectroscopy, sampling systems etc [4]. In general, it is not easy to produce such short pulses even from the best available laser sources. One approach is to utilize optical pulse compression (OPC) techniques to generate the ultrashort optical pulses from broader pulses.

Two types of compressors are commonly used for optical pulse compression. The first type of compression is based on linear effects. After passing through a dispersive delay line, the chirped pulses become narrower as they propagate in grating pairs. The second one is based on nonlinear effect; here, soliton effect compressor and adiabatic compressor are common for achieving pulse compression in the nonlinear regime [4]. Most of the pulse compression techniques rely on chirping. In soliton pulse compression technique, the compressed pulses suffer from significant pedestal generation as the induced chirp is linear only over the central part of the pulse. The pedestal will lead to nonlinear interactions between neighbouring solitons [4]. Adiabatic pulse compression technique has been used to generate a stable train of pedestal free and non-interacting solitons [4]. It has been shown that adiabatic pulse compression can be achieved using a dispersion-decreasing fibre.

As pointed out above, linear chirp is important for pedestal free pulse compression. Thus, recently, much interest has been focused on the generation of linearly chirped pulses. Moores suggested that chirped solitary pulses can be compressed more efficiently if the dispersion decreases approximately exponentially [5]. More recently, the self-similar (SS) analysis has been utilized to generate linearly chirped pulses in optical fibres and fibre amplifiers. Self-similarity is a common phenomenon in nature. Common objects like tree branches, snowflakes, clouds, rivers, or shorelines, appear similar even at a wide range of magnification scales. An object is said to be self-similar if it looks roughly the same on any scale. Thus, self-similarity is defined as the property whereby an object or mathematical function preserves its structure when multiplied by a certain scale factor [6]. Self-similarity is more than a curiosity of nature. Theoretical studies based on the self-similarity analysis of the nonlinear Schrödinger (NLS) equation with constant gain, have revealed that the interplay of normal dispersion, nonlinearity, and gain produces a linearly chirped pulse with a parabolic intensity profile which resists the deleterious effects of optical wave-breaking [7]. Asymptotic SS solitary pulses have been investigated using the NLS-type equation in the presence of gain [8]. Chirped solitary pulse compression has been demonstrated in these optical amplifiers. Very recently, using the self-similar analysis, chirped Bragg solitary pulses have been theoretically generated near the photonic bandgap of a non-uniform fiber Bragg grating, and the possibility of pedestal free Bragg soliton pulse compression is examined [9]. Furthermore, using the variational analysis, we have also theoretically investigated the generation of chirped fundamental optical solitons, which are formed because of the presence of the chirp in contrast to conventional solitons which are chirp free, wherein we have correlated our results with that of the self-similar analysis. For an exponentially decreasing dispersive medium, through the Hirota bilinear method and variational analysis, we have shown that the intensity and the chirp of the chirped soliton increase exponentially while its width decreases exponentially. These properties are consistent with those of the self-similar behavior [10].

Nowadays, the pulse propagation through competing nonlinearities has received much attention since the competition between nonlinearities of different orders, such as cubic and quintic, could cause strong stabilization of the pulse propagation [11]. Competing nonlinearities of different orders may be realized physically if one considers high optical intensities or materials possessing high nonlinear coefficients, for instance, semiconductor doped glasses [11, 12]. One of the fascinating examples of these competing nonlinearties are the stabilization of vortices and vortex tori in cubic-quintic nonlinear media [13–16]. Recently, bright and dark quasi solitons have been studied, for the cubic-quintic nonlinearity, by ansatz method [17]. The main objective of the present paper is to investigate the pulse compression of the chirped soliton pulse, which has been studied with the proper self-similar scaling, under the influence of competing cubic-quintic nonlinearities. The paper is organized as follows. Section 2 deals with the theoretical model where the origin of the next higher order nonlinearity called fifth order (quintic) nonlinearity is addressed. In Section 3, we discuss the generation of chirped pulses for anomalous and normal dispersion regimes and also investigate the corresponding pulse compression studies under the competing nonlinearities. In Section 4, we conclude our investigations.

2. THEORETICAL MODEL

The interest for considering cubic-quintic (CQ) nonlinearity in our model stems from a nonlinear correction to the medium's refractive index in the form $\delta n = n_2 I - n_4 I^2$, where I being the light intensity and the coefficients $n_2, n_4 > 0$ determine the nonlinear response of the media. They are related to third order susceptibility $\chi^{(3)}$ and fifth order susceptibility $\chi^{(5)}$ through $n_2 = 3\chi^{(3)}/8n_0$ and $n_4 = -5\chi^{(5)}/32n_0$, where n_0 being the linear refractive index. Although, formally it may be obtained by an expansion of the saturable nonlinearity $\delta n = n_2 I \left[1 + (n_4/n_2) I\right]^{-1}$, it is restricted under the effect of self-focusing as $d(\delta n)/dI$. However, the CQ model changes the sign of focusing at a critical intensity $I_c = (n_2/2n_4)$. An experimental measurement of the nonlinear dielectric response in the para-toluene salfonate (PTS) optical crystal aptly models the above-mentioned insights [18]. The above mentioned nonlinearity can be obtained by doping a fiber with two appropriate semiconductor materials. One should have positive sign $n_2^{(1)} > 0$ and large saturation intensity and other should have a negative sign $n_2^{(2)} < 0$ with nearly same magnitude but low saturation intensity, i.e., $|n_2^{(1)}| \sim |n_2^{(2)}|$ and $I_{sat}^{(2)} \ll I_{sat}^{(1)}$.

The pulse propagation in the above mentioned competing nonlinearities is governed by the generalized cubic-quintic NLS (CQNLS) equation with distributed linear gain is given by,

$$i\frac{\partial A}{\partial z} - \frac{\beta(z)}{2}\frac{\partial^2 A}{\partial \tau^2} + \gamma(z)\left|A\right|^2 A - \delta(z)\left|A\right|^4 A - i\frac{g(z)}{2}A = 0, \tag{1}$$

where all the physical parameters β, γ, δ and g are functions of the propagation distance z. This equation describes the amplification or attenuation of nonlinear pulses propagating nonlinearly in the single mode optical fiber under the influence of both cubic and quintic nonlinear effects. The sign of the parameter $g(z)$ determines wether it is amplification or attenuation. The function $A(z,t)$ is the slowly varying envelope of the axial electrical field, τ is the retarded time, $\beta(z)$ is the group velocity dispersion and $g(z)$ is the distributed gain function. The cubic and quintic nonlinear parameters are given by $\gamma = 2\pi n_2/(\lambda_0 A_{eff})$ and $\delta = 2\pi n_4/(\lambda_0 A_{eff}^2)$, where λ_0 is the central wavelength.

Now we start to investigate the so called chirped soliton under the influence of cubic-quintic nonlinearity by a scaling analysis known as self-similar analysis. For this purpose, the complex function $A(z,\tau)$ can be written as,

$$A(z,\tau) = Q(z,\tau)\exp\left[i\Phi(z,\tau)\right], \tag{2}$$

where Q and Φ are the amplitude and phase of the envelope function A respectively. In order to study generation of chirped soliton of Eq. (1), the following quadratic phase form is assumed,

$$\Phi(z,\tau) = \alpha_1(z) + \alpha_2(z)(\tau-\tau_c)^2. \tag{3}$$

According to self-similar scaling analysis, the amplitude depends on a scaling variable θ which is a combination of variables $(\tau-\tau_c)$ and some function $\Gamma(z)$ of variable z. Indeed the self-similar solutions possess scaling structure. Hence, we represent the amplitude $Q(z,\tau)$ in the following form,

$$Q(z,\tau) = \frac{1}{\sqrt{\Gamma(z)}}R(\theta)\exp\left(\frac{G(z)}{2}\right). \tag{4}$$

The scaling variable θ and the function $G(z)$ are

$$\theta = \frac{\tau-\tau_c}{\Gamma(z)}, \quad G(z) = \int_0^z g(z')dz', \tag{5}$$

where τ_c is the center of the pulse. Here $\Gamma(z)$ and $F(\theta)$ are some functions which have to be determined. It is also assumed that $\Gamma(0)=1$ without loss of generality. Now substituting Eqs. (2) and (4) into Eq. (1), the quadratic phase coefficient $\alpha_2(z)$ and the function $\Gamma(z)$ are found to be

$$\alpha_2(z) = \frac{\alpha_{20}}{1-\alpha_{20}D(z)}, \quad \Gamma(z) = 1-\alpha_{20}D(z), \tag{6}$$

where $\alpha_{20} = \alpha_2(0) \neq 0$ because the phase should be a quadratic function of variable $(\tau-\tau_c)$ and the cumulative dispersion function $D(z)$ is

$$D(z) = 2\int_0^z \beta(z')dz'. \tag{7}$$

In addition to the above conditions, we also find

$$\frac{d^2R}{d\theta^2} + \frac{2\Gamma^2}{\beta}\frac{d\alpha_1}{dz}R - \frac{2\Gamma\gamma}{\beta}\exp\left[G(z)\right]R^3 + \frac{2\delta}{\beta}\exp\left[2G(z)\right]R^5 = 0. \tag{8}$$

It should be emphasized that the coefficients in Eq. (8) are functions of variable z. However, the function $R(\theta)$ depends only on the scaling variable θ. Therefore, the above equation possesses nontrivial solutions $[R(\theta)\neq 0]$ if and only if the coefficients in Eq. (8) are constants,

$$-\frac{2\Gamma(z)^2}{\beta(z)}\frac{d\alpha_1}{dz} = \lambda_1, \tag{9}$$

$$\frac{\Gamma(z)\gamma(z)}{\beta(z)}\exp\left[G(z)\right] = \lambda_2, \tag{10}$$

$$\frac{\delta(z)}{\beta(z)}\exp\left[2G(z)\right] = \lambda_3, \tag{11}$$

where λ_1, λ_2, and λ_3 are constants. Therefore, the above equations yield

$$\lambda_1 = \frac{-2}{\beta(0)}\left.\frac{d\alpha_1}{dz}\right|_{z=0}, \qquad \lambda_2 = \frac{\gamma(0)}{\beta(0)}, \qquad \lambda_3 = \frac{\delta(0)}{\beta(0)}, \tag{12}$$

because $\Gamma(0) = 1$ and $G(0) = 0$. Thus, in the nontrivial case Eq. (8) can be written as

$$\frac{d^2R}{d\theta^2} - \lambda_1 R + 2\lambda_2 R^3 + 2\lambda_3 R^5 = 0. \tag{13}$$

The solution of Eq. (9) is

$$\alpha_1(z) \quad = \quad \alpha_{10} - \frac{\lambda_1}{2}\int_0^z \frac{\beta(z')dz'}{\left[1-\alpha_{20}D(z')\right]^2}, \tag{14}$$

where α_{10} is an integration constant. Now we proceed to find the distributed gain function using Eqs. (10) and (11)

$$g(z) = \frac{1}{\rho(z)}\frac{d\rho}{dz} - \frac{2\alpha_{20}\beta(z)}{\Gamma(z)}, \tag{15}$$

where we define the function $\rho(z)$ as

$$\rho(z) = \frac{\gamma(z)}{\delta(z)}, \qquad \rho(0) = \frac{\gamma(0)}{\delta(0)}. \tag{16}$$

These are all the necessary and sufficient conditions for the existence of the self-similar solutions of the generalized CQNLS Eq. (1) with distributed coefficients.

3. CHIRPED SELF-SIMILAR BRIGHT SOLITON

In the previous section, we have discussed the phase and the amplitude of self-similar solutions of the generalized NLS equation with distributed coefficients which are given by expressions (15) and (16). On integrating Eq. (13), we investigate the amplitude of the bright solitary wave by applying the following physical condition $\beta(z) < 0$ in Eq. (13) [19, 20]. Here, the integration constant λ_1 is chosen to be $1/\tau_0^2$, where τ_0 is the initial pulse width.

$$A(z,\tau) = \frac{1}{\tau_0\left(1-\alpha_{20}D(z)\right)}\left[\frac{-2\rho_1}{\sqrt{1+\frac{8\rho_1^2}{3\rho_2\tau_0^2(1-\alpha_{20}D(z))^2}}\cosh 2\left(\frac{\tau-\tau_c}{\tau_0(1-\alpha_{20}D(z))}\right)+1}\right]^{1/2}\exp\left(i\Phi\right), \tag{17}$$

where $\rho_1(z) = \beta(z)/\gamma(z)$ and $\rho_2(z) = \beta(z)/\delta(z)$. Equation (17) is the linearly chirped self-similar bright solitary pulse since it propagates in a shape preserving manner in a fiber medium under the influence of cubic-quintic nonlinearity. Here it should be emphasized that the above bright soliton result exactly goes to the well known case of Kerr nonlinearity when the quintic nonlinearity is switched off [8]. The resulting chirped bright solitary pulse is of practical significance as it possesses a strictly linear chirp which eventually leads to efficient pulse compression.

In order to investigate pulse compression of the chirped bright soliton pulse under the competing nonlinearities, we assume that the dispersion and the cubic nonlinearity parameters are distributed according to

$$\beta(z) = \beta_0\exp\left(-\sigma z\right), \qquad \gamma(z) = \gamma_0\exp\left(\eta z\right), \tag{18}$$

where $\beta_0 < 0$, $\gamma_0 > 0$, and $\sigma \neq 0$ ($\sigma > 0$ for dispersion decreasing fibers). At this juncture, we do not know how the next higher order nonlinearity called quintic nonlinearity varies. So, we shall find the condition for the quintic nonlinear parameter in the following using Eqs. (10) and (11)

$$\delta(z) = \frac{\gamma(z)^2\Gamma(z)^2}{\beta(z)}\frac{\lambda_3}{\lambda_2^2}. \tag{19}$$

To discuss the pulse compression, we assume a semiconductor doped (chalcogenide AS_2Se_3) fiber of length $L = 6\,\mathrm{km}$. The effective core area of the fiber is considered as $10\,\mu\mathrm{m}^2$. The cubic and quintic nonlinear coefficients are calculated as $\gamma_0 = 8900\ \mathrm{W}^{-1}\,\mathrm{km}^{-1}$ and $\delta_0 = 32\ \mathrm{W}^{-2}\,\mathrm{km}^{-1}$ respectively. Here, we choose the initial GVD value to be $\beta_0 = -20\,\mathrm{ps}^2\,\mathrm{km}^{-1}$. This initial GVD monotonically decreases to a final value at the end of the dispersion decreasing fiber medium $\beta(z = L)$, which can easily be calculated from Eq. (18). The compression scenario of the bright soliton under the influence of the competing nonlinearities is shown in Fig. 1. Fig. 2 gives the

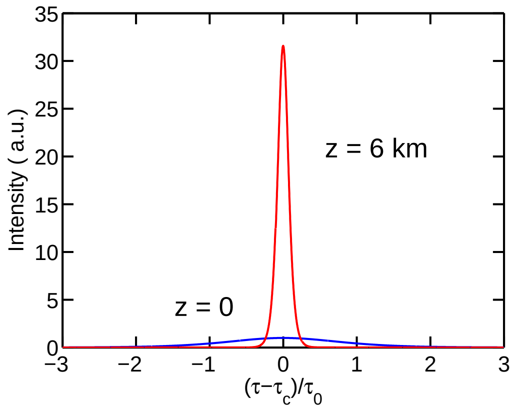


Figure 1: Compression of the chirped bright soliton pulse in time domain for the physical parameters $\tau_0 = 10\,\text{ps}$, $\beta_0 = -20\,\text{ps}^2/\text{km}$, $\sigma = 0.1\,\text{km}^{-1}$, $\eta = 0.1\,\text{km}^{-1}$, $\gamma_0 = 8900\,\text{W}^{-1}\,\text{km}^{-1}$, $\delta_0 = 32\,\text{W}^{-2}\,\text{km}^{-1}$, $\alpha_{20} = -0.005\,\text{THz}^2$ and $z = 6\,\text{km}$.

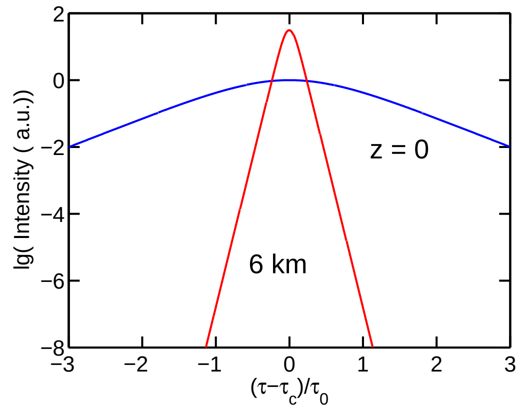


Figure 2: Log-linear plot of the chirped bright soliton pulse compression for the same physical parameters as in Fig. 1.

log-linear plot showing that the compressed bright soliton pulse maintains the shape even after the compression. The compression factor is found to be 10.

From the experimental point of view, it is necessary to know the magnitude of the peak power to excite the chirped bright soliton. Similarly, the soliton period turns out to be another important physical parameter that is involved in the formation of a soliton. Thus, for the chirped bright soliton, we calculate the important and interesting physical parameters such as soliton peak power, energy and pulse width. The energy of the chirped bright soliton is calculated as

$$W = \int_{-\infty}^{\infty} |A|^2 d\tau = -\frac{2\rho_1(z)}{\tau_0 \Gamma \sqrt{1-b^2}} \ln \frac{1+\sqrt{1-b^2}}{b}, \tag{20}$$

where $b = \sqrt{1 + 8\lambda_3/3\tau_0^2\lambda_2^2}$. Similarly, the other physical parameters namely peak intensity and the pulse width in terms of full-width at half-maximum (FWHM) are given by

$$(A^2)_{\max} = -\frac{1}{\tau_0^2\Gamma^2}\frac{2\rho_1(z)}{b+1}, \tag{21}$$

$$\Delta\tau = (\tau_0\Gamma)\ln\left(2 + y + \sqrt{(2+y)^2 - 1}\right), \tag{22}$$

where $y = 1/b$.

From Fig. 3, it is clear that the physical parameters like gain, energy and peak power increase exponentially while the width of the pulse decreases exponentially. This physical process aptly explains the occurrence of nonlinear pulse compression.

While the solitons in Eq. (17) exist in the anomalous dispersion regime, we found that chirped bright soliton like pulses are possible in the normal dispersion regime when the quintic nonlinearity is included. Under this physical condition ($\beta(z) > 0$, $\gamma(z) > 0$ and $\delta(z) > 0$), the bright soliton like pulse in the normal dispersion regime is given by

$$A(z,\tau) = \frac{1}{\tau_0(1-\alpha_{20}D(z))}\left[\frac{2\rho_1}{\sqrt{1+\frac{8\rho_1^2}{3\rho_2\tau_0^2(1-\alpha_{20}D(z))^2}}\cosh 2\left(\frac{\tau-\tau_c}{\tau_0(1-\alpha_{20}D(z))}\right)-1}\right]^{1/2}\exp(i\Phi). \tag{23}$$

The bright soliton like pulse in Eq. (23) is entirely different from the bright soliton like pulse given in Eq. (17) since the latter soliton pulse goes to well known soliton when the quinic nonlinearity coefficient vanishes whereas the former does not. In Eq. (23), if the quintic nonlinearity approaches zero, the energy required for the soliton will approach infinity. This imposes the constraint on the formation of soliton. However, even if a small amount of quintic nonlinearity is present, formation

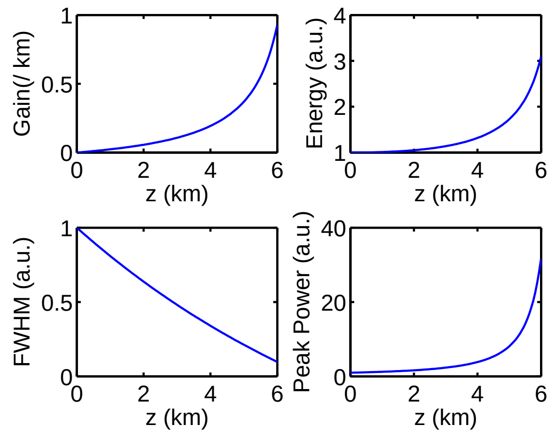


Figure 3: Evolution of the bright soliton pulse: gain (a), width (b), energy (c) and peak power (d) for the same physical parameters.

of this new type of soliton is possible. The Figs. 4 and 5 represent the compression of chirped soliton like pulse. These figures explain that the soliton compression could also be achieved even in the normal dispersion regime only under the influence of competing nonlinearities. Compression factor is also 10 in this case.

We also calculate the energy, peak intensity and pulse width and they are given by

$$\begin{aligned} W &= \int_{-\infty}^{\infty} |A|^2 d\tau = \frac{2\rho_1(z)}{\tau_0 F\sqrt{b^2-1}}\left[\frac{\pi}{2} + \arcsin\left(\frac{1}{b}\right)\right], \\ (A^2)_{\max} &= \frac{1}{\tau_0^2\Gamma^2}\frac{2\rho_1(z)}{b-1}, \\ \Delta\tau &= (\tau_0\Gamma)\ln\left(2 - y + \sqrt{(y-2)^2 - 1}\right), \end{aligned} \tag{24}$$

where $y = 1/b$.

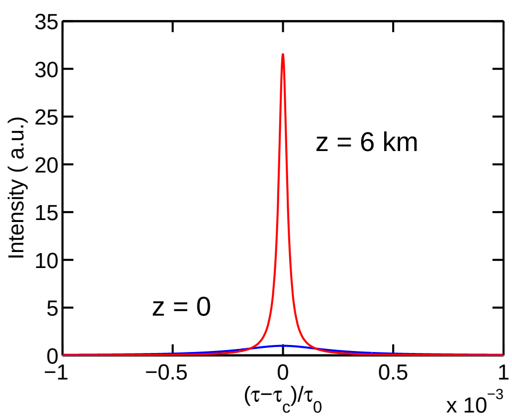


Figure 4: Compression of the chirped bright soliton like pulse in time domain for the physical parameters $\tau_0 = 10\,\text{ps}$, $\beta_0 = 20\,\text{ps}^2/\text{km}$, $\sigma = 0.1\,\text{km}^{-1}$, $\eta = 0.1\,\text{km}^{-1}$, $\gamma_0 = 8900\,\text{W}^{-1}\,\text{km}^{-1}$, $\delta_0 = 32\,\text{W}^{-2}\,\text{km}^{-1}$, $\alpha_{20} = 0.005\,\text{THz}^2$ and $z = 6\,\text{km}$.

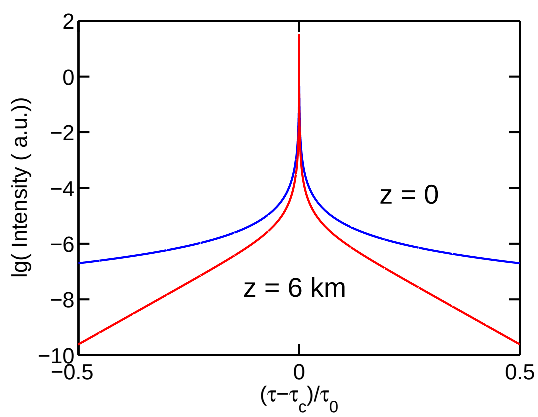


Figure 5: Log-linear plot of chirped bright soliton like pulse compression in the normal dispersion regime for the same physical parameters as in Fig. 4.

In the normal dispersion regime also, the physical parameters like gain, energy and peak power increase exponentially whereas width of the pulse decreases exponentially (Fig. 6). These behaviors are similar to what has been found for the chirped bright soliton pulse.

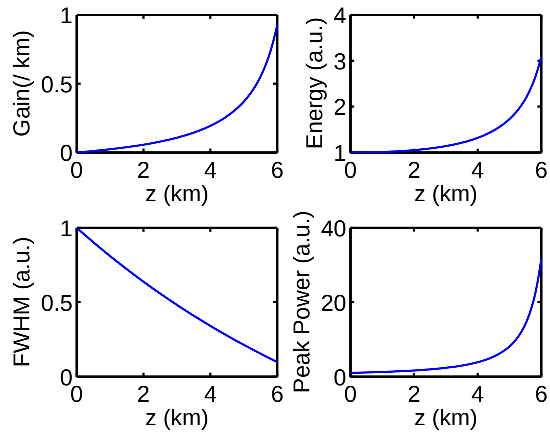


Figure 6: Evolution of the bright soliton like pulse: gain (a), width (b), energy (c) and peak power (d) for the same physical parameters.

4. CONCLUSION

By using self-similar scaling analysis, the generation of chirped bright soliton as well as chirped soliton like pulse has been discussed for anomalous and normal dispersion regimes respectively. It has been found that both the soliton and soliton like pulses differ under the limiting case. In the normal dispersion regime, the system favors soliton like pulse formation only in the presence of quintic nonlinearity. The important and interesting physical parameters like peak power, energy and pulse width of these two chirped soliton and soliton like pulses have also been calculated. From the results, it has been found that the peak power, energy and gain increase exponentially while pulse width decreases exponentially. This physical process has also been explained and demonstrated through the pulse compression of these chirped bright soliton pulses and the compression factor has been found to be 10 for both the cases. Furthermore, the main conclusion of this work is that one could achieve the soliton type pulse compression for both anomalous and normal dispersion regimes under the influence of competing nonlinearities.

ACKNOWLEDGMENT

The support from The Hong Kong Polytechnic University (Project No. G-YW85) is gratefully acknowledged. KN wishes to thank The Nuffield Foundation for their financial support through the Newly Appointed Lecturer Award.

REFERENCES

1. Hasegawa, A. and F. Tappert, "Transmission of stationary nonlinear optical pulses in dispersive dielectric fibers. I. Anomalous dispersion," *Appl. Phys. Lett.*, Vol. 23, No. 3, 142–144, 1973.
2. Agrawal, G. P., *Nonlinear Fiber Optics*, Academic Press, San Diego, 2001.
3. Mollenauer, L. F. and J. P. Gordon, *Solitons in optical fibers: Fundamentals and Applications*, Elsevier Academic Press, San Diego, 2006.
4. Agrawal, G. P., *Applications of Nonlinear Fiber Optics*, Academic Press, San Diego, 2001.
5. Moores, J. D., "Nonlinear compression of chirped solitary waves with and without phase modulation," *Opt. Lett.*, Vol. 21, No. 8, 555–557, 1996.
6. Barenblatt, G. I., *Scaling, Self-Similarity, and Intermediate Asymptotics*, Cambridge University Press, Cambridge, England, 1996.
7. Kruglov, V. I., A. C. Peacock, and J. D. Harvey, "Exact self-similar solutions of the generalized nonlinear Schrödinger equation with distributed coefficients," *Phys. Rev. Lett.*, Vol. 90, No. 11, 113902 (1-4), 2003.
8. Kruglov, V. I., A. C. Peacock, and J. D. Harvey, "Exact solutions of the generalized nonlinear Schrödinger equation with distributed coefficients," *Phys. Rev. E*, Vol. 71, No. 5, 056619(1-11), 2005.
9. Senthilnathan, K., P. K. A. Wai, and K. Nakkeeran, "Pedestal free pulse compression in nonuniform fiber Bragg gratings," *Proceedings of Optical Fiber Communication Conference*, California, USA, JWA19 (1-3), April 2007.

10. Senthilnathan, K., K. Nakkeeran, P. K. A. Wai, and K. W. Chow, "Chirped optical soliton," Communicated to *Europhysics Letters*.
11. Kivshar, Y. S. and G. P. Agrawal, *Optical Solitons: From Fibers to Photonic Crystals*, Academic Press, San Diego, 2003.
12. Akhmediev, N. N. and A. Ankiewicz, *Solitons: Nonlinear Pulses and Beams*, Chapman and Hall, London, 1997.
13. Mihalache, D., D. Mazilu, I. Towers, B. A. Malomed, and F. Lederer, "Stable spatiotemporal spinning solitons in a bimodal cubic-quintic medium," *Phys. Rev. E*, Vol. 67, No. 5, 056608 (1-9), 2003.
14. Towers, I. A., V. Buryak, R. A. Sammut, B. A. Malomed, L. C. Crasovan, and D. Mihalache, "Stability of spinning ring solitons of the cubicquintic nonlinear Schrdinger equation," *Phys. Lett. A*, Vol. 288, No. 5–6, 292–298, 2001.
15. Malomed, B. A., L. C. Crasovan, and D. Mihalache, "Stability of vortex solitons in the cubic-quintic model," *Physica D*, Vol. 161, No. 3–4, 187–201, 2002.
16. Mihalache, D., D. Mazilu, L. C. Crasovan, I. Towers, A. V. Buryak, B. A. Malomed, L. Torner, J. P. Torres, and F. Lederer, "Stable spinning optical solitons in three dimensions," *Phys. Rev. Lett.*, Vol. 88, No. 7, 073902 (1-4), 2002.
17. Hao, R., L. Li, Z. Li, R. Yang, and G. Zhou, "A new way to exact quasi-soliton solutions and soliton interaction for the cubic-quintic nonlinear Schrdinger equation with variable coefficients," *Opt. Comm.*, Vol. 245, No. 1–6, 383–390, 2005.
18. Lawrence, B., W. E. Torruellas, M. Cha, M. L. Sundheimer, G. I. Stegeman, J. Meth, S. Etemad, and G. Baker, "Identification and role of two-photon excited states in a π-conjugated polymer," *Phys. Rev. Lett.*, Vol. 73, No. 4, 597–600, 1994.
19. Maimistov, A. I., B. A. Malomed, and A. Desyatnikov, "A potential of incoherent attraction between multidimensional solitons," *Phys. Lett. A*, Vol. 254, No. 3–4, 179–184, 1999.
20. Porsezian, K., K. Senthilnathan, and S. Devipriya, "Modulational instability in fiber Bragg grating with non-Kerr nonlinearity," *IEEE Qunatum Electron.*, Vol. 41, No. 6, 789–796, 2005.

Simulation of Electromagnetic Pulse Propagation through Dielectric Slabs with Finite Conductivity Using Characteristic-based Method

Mingtsu Ho

Department of Electronic Engineering, WuFeng Institute of Technology
Chia Yi 621, Taiwan, China

Abstract— This paper presents one-dimensional simulation results of electromagnetic pulse propagation through various dielectric slabs. The material used in the numerical model is assumed to be homogeneous and isotropic such that its dielectric constant, relative permeability, and conductivity are not functions of the time and independent of the electric and magnetic field intensities. The computational results are generated by the characteristic-based method and compared in a side-by-side fashion. The effects of medium conductivity on the reflected and transmitted pulses are recorded and shown as time elapses. Also revealed is how the medium conductivity affects the electromagnetic pulses for dielectric slabs that are impedance matched to free space, i. e., $\varepsilon_r = \mu_r$. The frequency-domain results are illustrated that are obtained through the application of FFT to the time-domain data.

1. INTRODUCTION

The method of moment (MoM) and the finite-difference time-domain (FDTD) technique have been the two most popular computational electrodynamics modeling techniques ever since they were proposed in 60 s [1, 2]. The latter can be formulated for either time- or frequency-domain analysis while the former is only in frequency-domain. Numerical simulation results always provide researchers a much more perceptive view of a variety of electromagnetic phenomena and therefore give researchers a better understanding of the problem of interest. This is in contrast to most textbooks where electromagnetic problems are highlighted in frequency-dependent formulations. Furthermore, through proper Fourier transformation one can easily maneuver the computational results, from time-domain transient data to frequency-domain spectrum analysis or *vice versa*.

The characteristic-based method was proposed in mid-90 s and found to produce results in good agreement with data produced by FDTD [3]. Unlike MoM and FDTD, the present numerical method is an implicit numerical procedure, places all field components in grid center, solves Maxwell's equations by balancing all fluxes across cell faces within each computational cell, and yet pays the cost of recasting the governing equations from the Cartesian coordinates into the body-conforming coordinates. It is shown that the characteristic-based method can predict the reflection of electromagnetic fields from moving/vibrating perfect conductor in one dimension [4] and the reflection/transmission of electromagnetic field propagation onto moving dielectric half space [5].

In order to investigate the effects of medium conductivity on the electromagnetic field as it propagates through the medium, one has to include the conductivity in the governing equations, Maxwell's equations. Luebbers et al. simulated the interaction of a Gaussian pulse plane wave with a dielectric slab of finite thickness, constant relative permittivity and conductivity in one-dimension using FDTD [6]. In formulation, FDTD explicitly specifies the medium properties at grid node in accordance with the designation of the electric field and the magnetic field components that are interleaved in time and space. These medium properties are permittivity (ε), permeability (μ), and conductivity (σ). The inclusion of the conductivity term in the governing equation is simply meaning a source term, $-\sigma E$. In the FDTD expression, conductivity is multiplied by the numerical time step (Δt) and then combined with permittivity at each grid point. In the characteristic-based algorithm, it is an extra quantity added to the calculated flux difference for every cell inside the medium.

Since the main objective of this paper is to simulate and demonstrate the computational results of electromagnetic pulse propagation through different conducting dielectric slabs using the characteristic-based method, the comparison between numerical methods will not be covered here. The computational results are illustrated as follows. The effects of medium conductivity on the reflected and transmitted electromagnetic fields are plotted as functions of the time. Especially, for the case when the medium is impedance matched to free space (the reflection coefficient is zero in magnitude), we also demonstrate how conductivity affects the electromagnetic pulse by assuming this impedance matched medium is also conductive. Two groups of plots of the time-domain reflected and transmitted electric fields and their spectra are shown and compared.

2. GOVERNING EQUATIONS

Maxwell's equations are the governing equations of electromagnetic problems. When electromagnetic waves propagate onto a medium with finite conductivity, Maxwell's equations become

$$\frac{\partial \vec{B}}{\partial t} + \nabla \times \vec{E} = 0, \tag{1}$$

$$\frac{\partial \vec{D}}{\partial t} - \nabla \times \vec{H} = -\sigma \vec{E}, \tag{2}$$

$$\nabla \cdot \vec{D} = 0, \tag{3}$$

$$\nabla \cdot \vec{B} = 0, \tag{4}$$

where $\vec{E}$ and $\vec{H}$ are the electric and magnetic field intensities, $\vec{D}$ and $\vec{B}$ are the electric and magnetic flux densities, respectively. The two constitutive relations are $\vec{D} = \varepsilon_o \vec{E}$ and $\vec{B} = \mu_o \vec{H}$ with ε_o and μ_o are the permittivity and permeability of vacuum. Inside medium, these two expressions are given by $\vec{D} = \varepsilon \vec{E} = \varepsilon_r \varepsilon_o \vec{E}$ and $\vec{B} = \mu \vec{H} = \mu_r \mu_o \vec{H}$ with ε_r and μ_r are the dielectric constant and relative permeability of medium. Finally, symbol σ is the medium conductivity. If the medium is conductive, then $\sigma \vec{E}$ represents the current density. Once more, in the numerical model we assume ε, μ, and σ are not functions of the time.

Since the present work focuses on one-dimensional electromagnetic field propagation problem, we can only consider a two-dimensional formulation and make the following assumptions. In the numerical model, the electric field intensity is polarized to z-direction and has a peak value of unity, the incident pulse initially propagates in the positive x-direction and normally illuminates upon a dielectric slab that may be featured with a constant conductivity. As stated earlier, the Maxwell's equations have to be recast from the Cartesian coordinate system (t, x, y) into curvilinear coordinate system (τ, ξ, η) and are written as

$$\frac{\partial Q}{\partial \tau} + \frac{\partial F}{\partial \xi} + \frac{\partial G}{\partial \eta} = -\sigma E_z \tag{5}$$

where $Q = Jq$, $F = J(\xi_x f + \xi_y g)$, $G = J(\eta_x f + \eta_y g)$, $q = [B_x, B_y, D_z]^T$, $f = [0, -E_z, -H_y]^T$, $g = [E_z, 0, H_x]^T$, and J is equal to $\left|\frac{\partial(x, y)}{\partial(\xi, \eta)}\right|$. The numerical formulation is derived by applying the central difference to (5) and are written as

$$\frac{Q^{n+1} - Q^n}{\Delta \tau} + \frac{\delta_i F}{\Delta \xi} + \frac{\delta_j G}{\Delta \eta} = -\sigma E_z \tag{6}$$

where

$$\delta_k(^*) = (^*)_{k+1/2} - (^*)_{k-1/2} \tag{7}$$

is known as the central-difference operator. Symbols Q^{n+1} and Q^n are the variable vectors of two successive time levels, $\Delta\xi$ and $\Delta\eta$ represent the cell size between discrete points in the computational domain, $\Delta\tau$ is the numerical time step, the subscript (k) of the central-difference operator represents the direction in the curvilinear coordinate system and the one-half integer index in (7) indicates that the flux vectors F and G are evaluated at the cell face. The flux vector splitting technique is then applied to (6) and the lower-upper approximate factorization scheme is employed for the solution of the system of linear equations.

The boundary condition treatments in the present simulation simply assure that the tangential components of both electric and magnetic field intensities are continuous across the interface between two media, and that the out-going fields will not be bounced back into the computational domain from the outer computational layer.

3. CASES STUDIED

The problem of interest is defined in Figure 1 where the electric field intensity is shown. As illustrated the incident pulse is Gaussian in form and propagates toward the dielectric slab. The Gaussian pulse has a width of about 0.695 ns, a truncated level of 100 dB, a span of 2 meters, and a highest frequency content of about one GHz. The dielectric slab is 30 cm thick and may be specified by a finite conductivity whose magnitude is normalized by η_0 for the purpose of clear demonstration. For clear demonstration we further assume that the reference time zero is when the peak of the incident pulse is at $x = 1\,\text{m}$ and set a constant time interval $\Delta t = 0.2\,\text{m/C}$ where the symbol C is the light speed and exactly equal to $3 \times 10^8\,\text{m/s}$ in the present simulation. This means the numerical electromagnetic pulse propagates precisely 0.2 meters during the time interval Δt.

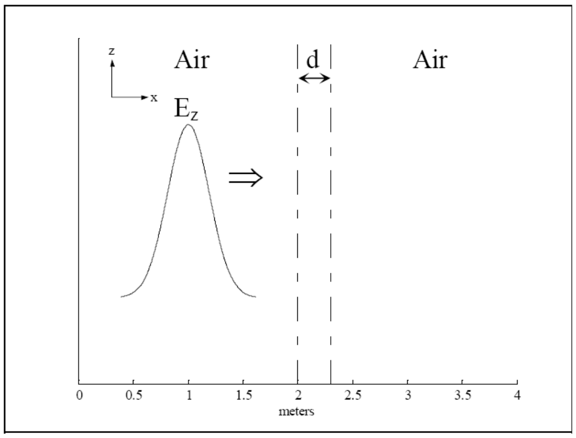


Figure 1: Definition of the problem.

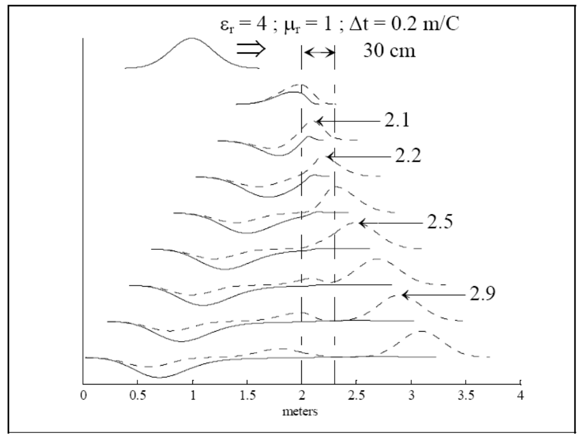


Figure 2: Propagation of electromagnetic pulse through two dielectric slabs. Solid lines: $\sigma = 1$; Dotted lines: $\sigma = 0$.

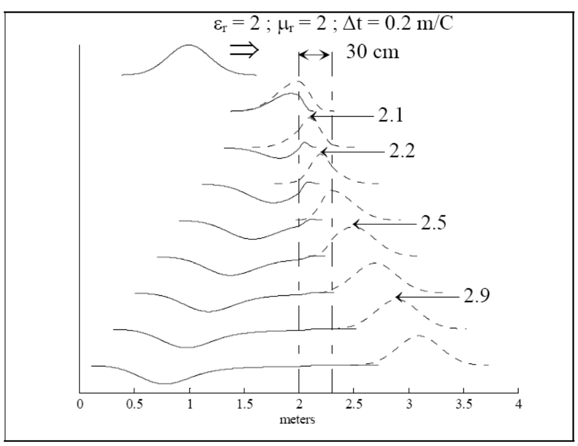


Figure 3: Propagation of electromagnetic pulse through two impedance-matched dielectric slabs. Solid lines: $\sigma = 1$; Dotted lines: $\sigma = 0$.

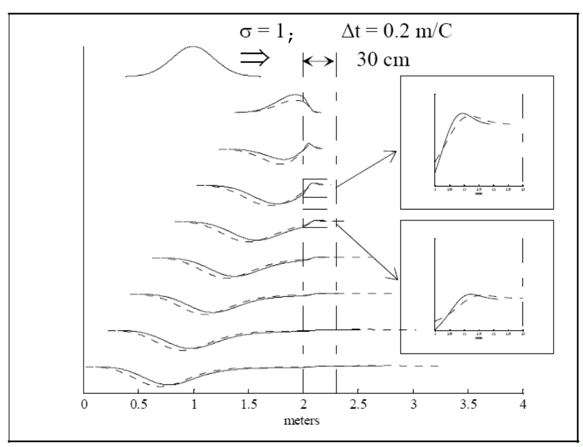


Figure 4: Propagation of electromagnetic pulse through two conducting dielectric slabs. Solid lines: $\varepsilon_r = 2$, $\mu_r = 2$; Dotted lines: $\varepsilon_r = 4$, $\mu_r = 1$.

The medium used in the simulation is assumed to be homogeneous and isotropic. In order to observe the effects of medium conductivity on the electromagnetic fields, we further specify the properties of dielectric slabs as follows. They may be made of either non-magnetic material with $\varepsilon_r = 4$ and $\mu_r = 1$ or material with $\varepsilon_r = 2$ and $\mu_r = 2$ whose impedance is matched to the free space.

4. RESULTS

For the purpose of easy and clear comparison, two sequences of plots of electric fields for electromagnetic pulse propagation through two different dielectric slabs are given in Figure 2. It is shown that the reflected electric field from conducting slab is stronger in magnitude than that from lossless slab, and that the attenuation of the transmitted field as it propagates through the lossy medium. Also noticed are that the reflected electric field bears a dragging tail due to the medium conductivity, and that the electromagnetic pulse is slow down by one half as it propagates through the dielectric slab. This is pointed out by indicating the location of the pulse peak in the plot. In Figure 3, similar setups are used except that the medium is impedance-matched to vacuum. That there exists no reflection of electromagnetic field is evident if the dielectric slab has no conductivity. When the slab becomes conductive, the reflected pulse is observed. For closer examination and comparison on the effects of the medium conductivity on the reflected electric fields from different dielectric slabs, the computational results of two previous conducting slabs were illustrated in Figure 4.

Though having same conductivity and index of refraction ($\sqrt{\varepsilon_r \mu_r} = 2$), based on direct inspections, two facts are observed. The reflected pulse from nonmagnetic slab leads that from magnetic slab and the nonmagnetic slab seems easier to propagate through for electromagnetic pulse than the magnetic slab. Two close-up shots are included in the plot showing details of electric fields inside the slab.

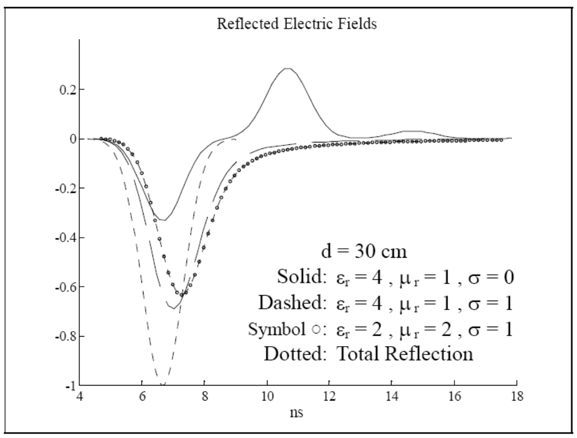


Figure 5: Reflected electric fields from various dielectric slabs.

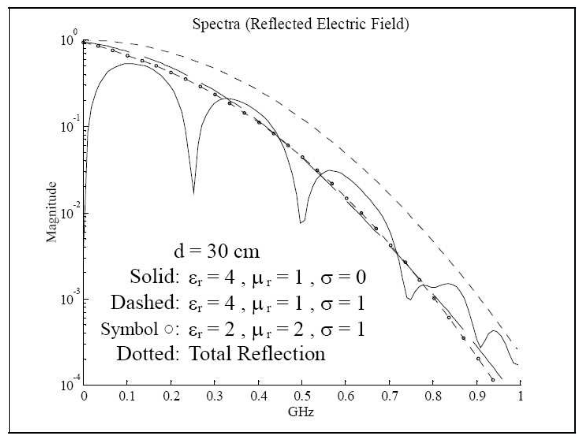


Figure 6: Spectra of reflected electric fields from various dielectric slabs.

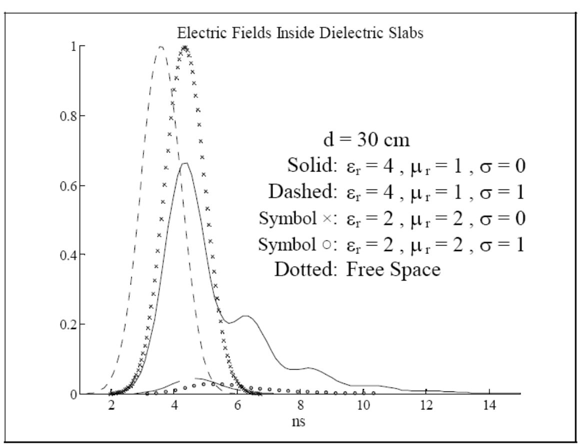


Figure 7: Electric fields inside various dielectric slabs.

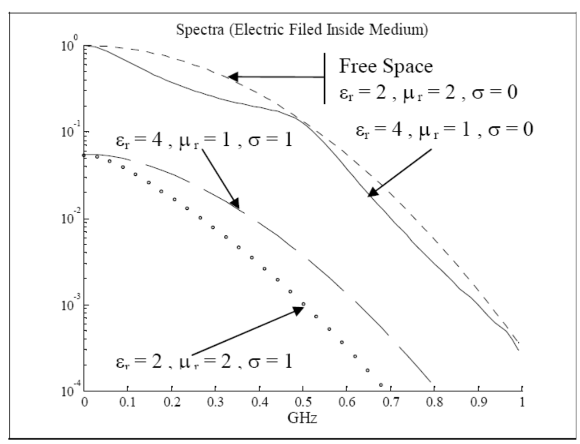


Figure 8: Spectra of electric fields inside various dielectric slabs.

A family of reflected electric fields as functions of the time is plotted in Figure 5 side-by-side for close investigation. Note that the total reflection is included as a reference and that the reflected electric field from slab of $\varepsilon_r = 2$, $\mu_r = 2$, $\sigma = 0$ is not given in the plot for its impedance is matched to the environment. As can be observed, the medium conductivity answers for the attenuations of electromagnetic pulse inside the dielectric slab. This is also shown in Figure 6 that the frequency contents are shrunk and reduced in magnitude for conductive media. Given in Figures 7 and 8 are the electric fields recorded at the middle point of various dielectric slabs and their spectra, respectively. Again, the electric field in free space is given as a reference. In the medium having conductivity, the electromagnetic pulse is attenuated dramatically. It is observed in the spectrum as well. Note that the electric fields both inside the impedance-matched medium and in free space are different merely by a time lag as shown in Figure 7 and have the same spectrum magnitude as indicated in Figure 8.

5. CONCLUSION

In this paper the author showed how the medium conductivity affects the reflected and transmitted electromagnetic pulses as they propagate through various dielectric slabs through the application of the characteristic-based method in one-dimension. The effects are illustrated in both time-domain and frequency-domain. Irrespective of the fact that the comparison of computational results between two numerical methods is not available, it is pointed out that in formulating the numerical procedure the FDTD technique puts extra efforts in teaming up the conductivity with the permittivity term while the present method simply adds an extra term in the governing equations. The computational results give reasonable trend. It is our future goal to extend the existing formulation to more complicated problems.

REFERENCES

1. Yee, K., "Numerical solutions of initial boundary value problems involving Maxwell's equations in isotropic media," *IEEE Transactions on Antennas and Propagation*, Vol. AP-14, 302–307, 1966.
2. Harrington, R. F., *Field Computation by Moment Methods*, MacMillan, New York, 1968.
3. Donohoe, J. P., J. H. Beggs, and M. Ho, "Comparison of finite-difference time-domain results for scattered EM fields: Yee algorithm vs. a characteristic based algorithm," *27th IEEE Southeastern Symposium on System Theory*, March 1995.
4. Ho, M., "Scattering of EM waves by vibrating perfect surfaces simulation using relativistic boundary conditions," *Journal of Electromagnetic Waves and Applications (JEMWA)*, Vol. 20, No. 4, 425–433, 2006.
5. Ho, M., "Propagation of electromagnetic pulse onto a moving lossless dielectric half-space: one-dimensional simulation using characteristic-based method," *Journal of Electromagnetic Waves and Applications (JEMWA)*, Vol. 19, No. 4, 469–478, 2005.
6. Luebbers, R. J., K. S. Kunz, and K. C. Chamberlin, "An interaction demonstration of electromagnetic wave propagation using time-domain finite differences," *IEEE Transactions on Education*, Vol. 33, No. 1, 60–68, 1990.

Broadband PLC Radiation from a Power Line with Sag

Nan Maung
Department of Electrical Engineering & Computer Engineering
New Jersey Institute of Technology, Newark, NJ 07102-1982, USA

Xiao-Bang Xu
Holcombe Department of Electrical & Computer Engineering
Clemson University, Clemson, SC 29634-0915, USA

Abstract— This paper presents a computational model for the analysis of broadband PLC radiation by a practical power line with sag. The numerical results show that the sag may have a significant influence on the magnetic vector potential generated by the power line.

Power line carrier (PLC) is a communication technique that uses existing power lines to carry information, traditionally using frequencies in the KHz range. Lately, there has been an interest in high-speed broadband communication on power lines. The broadband power line carrier communication uses frequencies in the MHz range, which is a challenge, concerning the signal attenuation and its electromagnetic compatibility with other communication systems.

Wave propagation in a *straight* transmission line above ground has been studied in the past. But an overhead power line is actually of *catenary shape* due to its sag. The goal of this paper is to investigate the difference between the electromagnetic (EM) field radiated by a broadband PLC line of catenary shape and that radiated by a straight line.

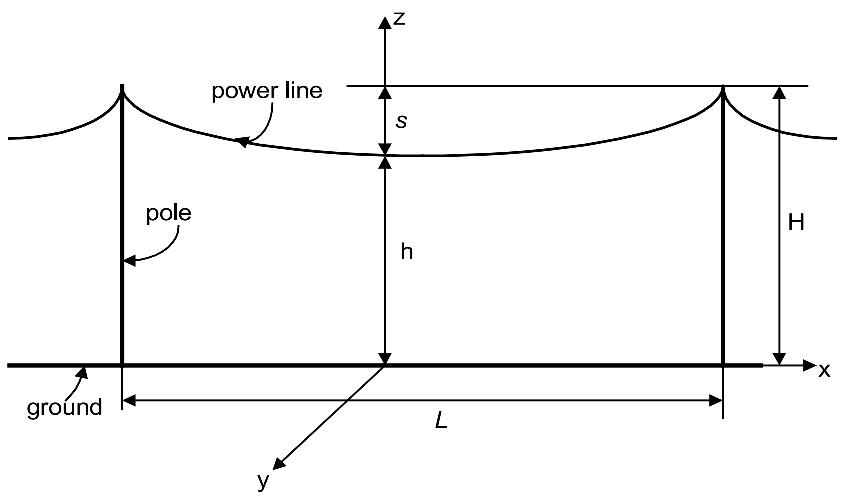


Figure 1: A practical power line with sag.

A practical power line with sag s is illustrated in Fig. 1. As pointed out in [1] and [2], for a practical overhead power line with small relative sag, its catenary shape can be approximated by a parabola,

$$z = h + \frac{4s}{L^2}\left(x - \frac{L}{2}\right)^2, \tag{1}$$

As the first step of studying the electromagnetic radiation by such a line, we partition the line of length L into N short dipoles as depicted in Fig. 2. Each of the dipoles is of horizontal length $\Delta x = L/N$, where N is taken to be sufficiently large so that the length of each dipole is much shorter than the wavelength, and they can be approximated by horizontal dipoles.

Then, we compute the magnetic vector potential generated by each dipole. Finally, the total magnetic potential and subsequently the electromagnetic field radiated by the line can be determined as the sum of the fields produced by each of the short dipoles. In this paper, we focus on

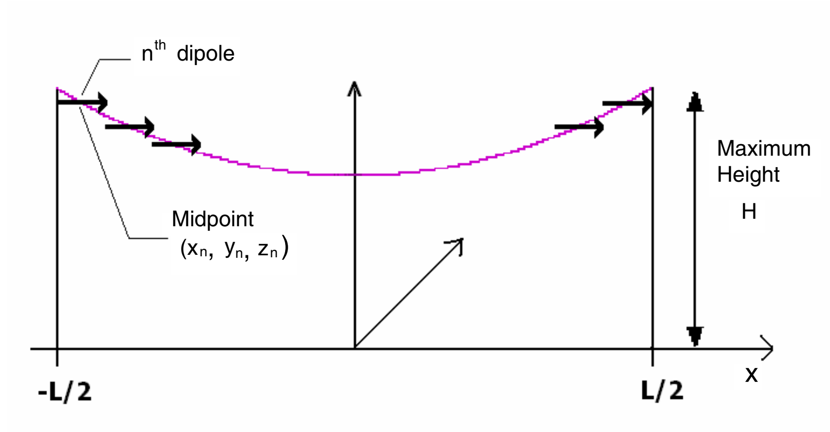


Figure 2: Division of the line into N short dipoles.

the computation of the x-component of the magnetic potential A_{xx}, observed in the vicinity of the line above ground, and A_{xx} is readily found to be

$$A_{xx} = \sum_{n=1}^{N} \int_{\Delta x_n} I G_{xx}(\vec{r}, \vec{r}\,')dx' \tag{2}$$

where I is a known current carried by the line and the Green's function is given by [3–7]

$$G_{xx}(\vec{r}, \vec{r}\,') = \frac{\mu_a}{4\pi} \frac{e^{-jk_a|\vec{r}-\vec{r}\,'|}}{|\vec{r}-\vec{r}\,'|} - j\frac{\mu_a}{4\pi} \int_0^{\infty} R \frac{e^{-j\beta_a(z+z')}}{\beta_a} J_0(\kappa\rho)\kappa d\kappa, \tag{3}$$

where $\mu_a = \mu_0$, $k_a = k_0$ are the permeability and the wave number of air, $J_0(\kappa\rho)$ is the zero-order Bessel function of the first kind, ρ is the distance between the source point and the field point, and

$$R \equiv \frac{\dfrac{\beta_a}{\mu_a} - \dfrac{\beta_b}{\mu_b}}{\dfrac{\beta_a}{\mu_a} + \dfrac{\beta_b}{\mu_b}}, \tag{4}$$

in which $\beta_{\substack{a\\b}}$ is defined by $\beta_{\substack{a\\b}}^2 = k_{\substack{a\\b}}^2 - \kappa^2$, where k_b is the wave number of the earth. The Sommerfeld-type integral [8] appearing in Equation (3) is evaluated using a Gaussian-quadrature numerical integration scheme.

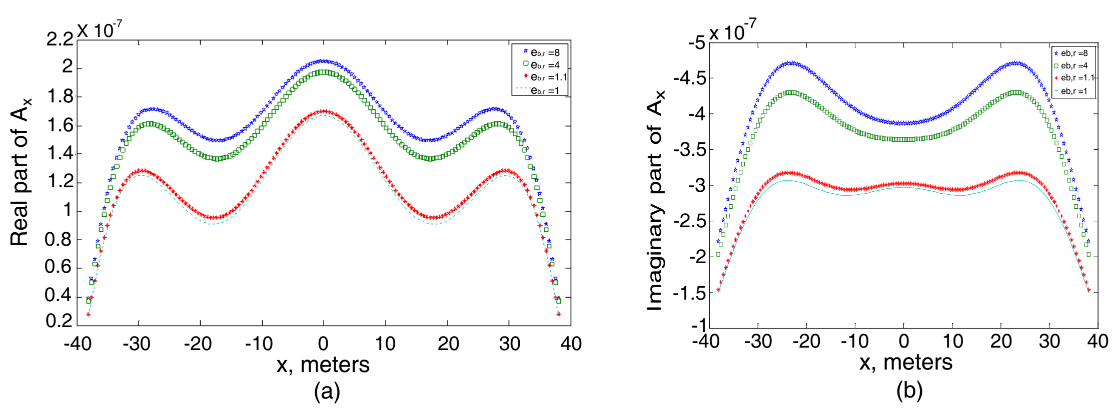
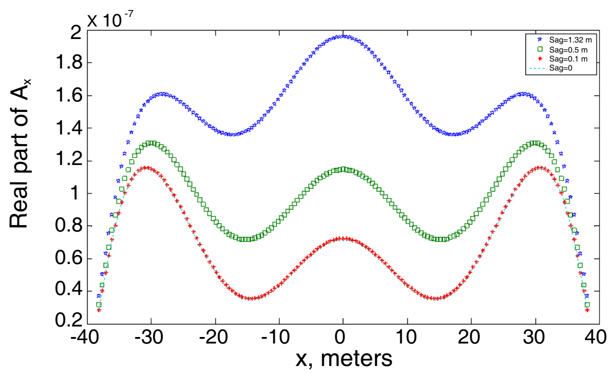


Figure 3: Magnetic vector potential computed for various region b relative permittivities.

Using Equations (2)–(4) together with (1), the magnetic potential A_{xx}, in the vicinity of a broadband PLC line is computed for a practical power line of length $L = 76.2\,\text{m}$ and sag $s = 1.32\,\text{m}$.

The frequency is taken to be $f = 10\,\mathrm{MHz}$. At this frequency, the line is partitioned into 151 dipoles and the dipole length is indeed much smaller than the wavelength. The potential is observed at a point 3-meters right below the lowest point of the line with sag. In Figs. 3(a) and 3(b) are shown the real and imaginary part of A_{xx}, computed for various relative permittivity of the lower half space (region b) ε_{br}. One observes that as ε_{br} gradually decreases to one, the potential reduces to that of free space, as expected.

Figures. 4(a) and 4(b) depict data of the potential corresponding to different line sags. One notices a significant difference between the potential generated by a practical power line with sag $s = 1.32\,\mathrm{m}$ and that produced by a straight line. As the sag decreases to zero, the potential gradually becomes that due to a straight line.



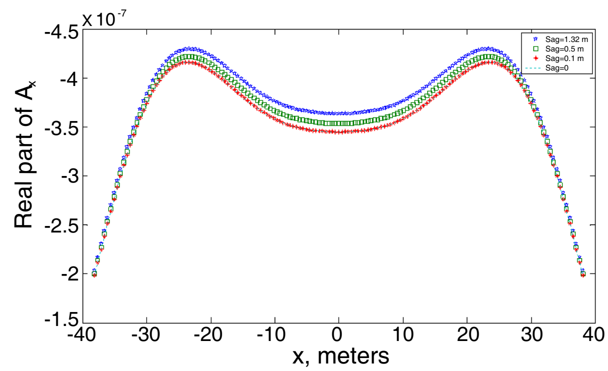


Figure 4: Magnetic vector potential computed for different line sags.

REFERENCES

1. Mamishev, A. V., R. D. Nevels, and B. D. Russel, "Effects of conductor sag on spatial distribution of power line magnetic field," *IEEE Transactions on Power Delivery*, Vol. 11, No. 3, 1571–1576, July 1996.
2. Xu, X.-B., X. Cheng, and K. Craven, "An analysis on magnetically induced subsequent fault in utility line topologies," *Electric Power Systems Research*, Vol. 63/3, 161–168, October 2002.
3. Tai, C.-T., *Dyadic Functions in Electromagnetic Theory*, Second Edition, IEEE Press, New York, 1993.
4. Michalski, K. A. and D. Zheng, "Electromagnetic scattering and radiation by surfaces of arbitrary shape in layered media, part I: theory," *IEEE Transactions on Antennas and Propagation*, Vol. 38, No. 3, 335–344, March 1990.
5. Michalski, K. A. and D. Zheng, "Electromagnetic scattering and radiation by surfaces of arbitrary shape in layered media, part II: implementation and results for contiguous half-spaces," *IEEE Transactions on Antennas and Propagation*, Vol. 38, No. 3, 345–352, March 1990.
6. Michalski, K. A. and J. R. Mosig, "Multilayered media Green's functions in integral equation formulations," *IEEE Transactions on Antennas and Propagation*, Vol. 45, No. 3, 508–519, March 1997.
7. Kaires, R. G., "Analysis of narrow conducting strips on a grounded dielectric slab by means of the modified diakoptic theory," Ph.D. Dissertation, Clemson University, 1989.
8. Sommerfeld, A., *Partial Differential Equations*, Academic, New York, 1949.

A Novel Broadband Quasi-fractal Binary Tree Dipole

Xing Chen
College of Electronics and Information Engineering, Sichuan University, China

Abstract— With the development of fractal theory, fractal geometries are widely used in the antenna design. In this paper, a novel dipole antenna based on the quasi-fractal binary tree is proposed. To achieve broadband characteristics, the proposed antenna is optimized in an automated design, making use of the Genetic Algorithm (GA) in conjunction with NEC (Numerical Electromagnetic Codes) and cluster parallel computation. A design result is presented. A prototype of the designed antenna has been fabricated and tested. Results of the measurement verify the designed antenna is able to provide a broad impedance bandwidth.

1. INTRODUCTION

Fractals are geometrical shapes, which are self-similar, repeating themselves at different scales. With the development of the fractal theory, the nature of fractal geometries has caught the attention of antenna designers. The utilization of fractal geometries in antenna design has led to the evolution of a new class of antennas, called fractal shaped antennas [1].

As one class of fractal geometries, the fractal tree has already been exploited in antenna designs to produce multi-band characteristics or to achieve miniaturization [2–4]. The fractal tree includes several families such as the binary, three dimensional, ternary, and etc. As illustrated in Fig. 1, the structure of a canonical binary fractal tree can be defined by the following parameters: a length of the trunk LT; Branch angle 2θ or branch half angle θ; Scale ratio S that is the length ratio between a child branch and its parent branch as well as between a first level branch and the trunk; The number of iteration N.

In this paper, a dipole antenna based on a quasi-fractal binary tree is presented. Different from a canonical binary fractal tree, the length of the trunk of a quasi-fractal binary tree is independent from the length of branches, i.e., the scale ratio only denotes the length ratio between a child branch and its parent branch. The configuration of a quasi-fractal binary tree dipole is shown in the Fig. 2. Besides parameters LT, θ, S, and N, we need one more parameter, i.e., the length of the first level branch LB, to descript the structure of the proposed antenna.

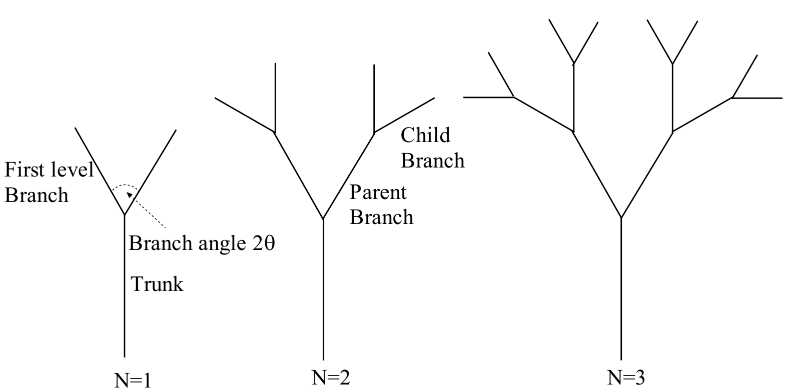


Figure 1: The first 3 iterations of a binary fractal tree.

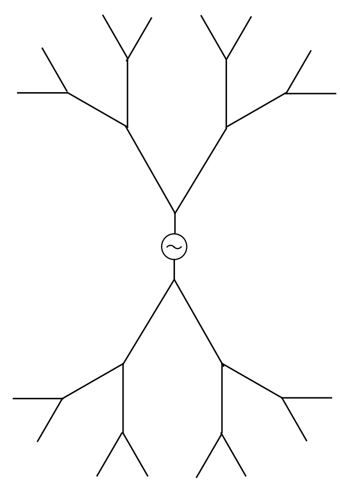


Figure 2: Configuration of a 3rd iterated quasi-fractal binary tree dipole.

2. AUTOMATED ANTENNA DESIGN USING GA AND NEC

To achieve broadband characteristics, the proposed antenna will be automated designed making use of parallel Genetic Algorithm (GA) in conjunction with NEC (Numerical Electromagnetic Codes) on a cluster system.

The GA [5] is a non-linear, robust stochastic optimization algorithm based on the Darwinian theory of decent with modification by natural selection. It has found great utility in electromagnetic optimization tasks including the design of various antennas [6, 7].

NEC is a Method-of-Moments (MoM) simulator for wire antennas, which was developed at Lawrence Livermore National Laboratory in the early 1980's. As a well-known antenna simulator, it is widely used in antenna simulation and design [6–9]. In this research, the second version of NEC (NEC2) is used for computing the input impedance of the proposed antenna.

An automated antenna design based on GA usually invokes hundreds or even thousands numerical simulations, hence is computationally intensive. Since the GA exhibits an intrinsic parallelism and allows a very straightforward implementation on parallel computers, we implement the GA-based antenna design into parallel computation to make the computation more effective. The computation is parallelized in a master-slave model and is carried out in a Beowulf cluster system, which is composed of 16 AMD 1700+ processors interconnected by a fast 100 Mb/s Ethernet and uses the message passing interface (MPI) library. One processor, named the master processor, controls the antenna design procedure. While the other processors, called slave processors, carry out the numerical simulations using NEC2.

3. AUTOMATED DESIGN OF THE QUASI-FRACTAL BINARY TREE DIPOLE

In this paper, the quasi-fractal binary antenna is assumed to be made of 1 mm-diameter copper wires, operate at a center frequency of 2450 MHz, and be fed by a 50 Ω coaxial cable through a $1 : K$ balun (contraction for "balanced to Unbalanced") [10], which is employed for eliminating the unbalanced currents and for impedance transformation.

The design goal of the proposed antenna is to achieve a broad impedance bandwidth at the center frequency. Therefore, the fitness function is defined as

$$F = C^{*}BW, \tag{1}$$

where C is a weight factor and set to be 0.2, BW is the impedance bandwidth in GHz.

In the GA optimization procedure, total 5 parameters, i.e., LT, BT, θ, S, N and K, will be optimized. LT and BT are confined within the range of 5 mm to 100 mm, θ is restricted to be from 5° to 60°, S is between 0 and 1, and both N and K are chosen from integers of 1 to 6.

A set of GA-based automated design processes are executed. In each of them, a binary GA carries out 200 generations with the number of individuals in a population $N_{ind} = 200$, the probability of crossover $P_c = 0.5$, and the probability of mutation $P_m = 0.2$.

4. RESULTS AND DISCUSSION

The result of the automated design is: $LT = 8.0$ mm, $BT = 41.7$ mm, $\theta = 8.5°$, $S = 0.7$, $N = 5$, and $K = 4$. Simulated by NEC, the input impedance and the corresponding VSWR (Voltage Standing Wave Ratio) of the designed quasi-fractal binary tree dipole are illustrated in the Fig. 3. One can observe that the designed antenna has a wide 2 : 1 VSWR impedance band from 1150 MHz to 3700 MHz.

To validate the result of the automated antenna design, a prototype of the designed antenna has been fabricated and tested. As shown in Fig. 4, the prototype antenna is on an organic plastic

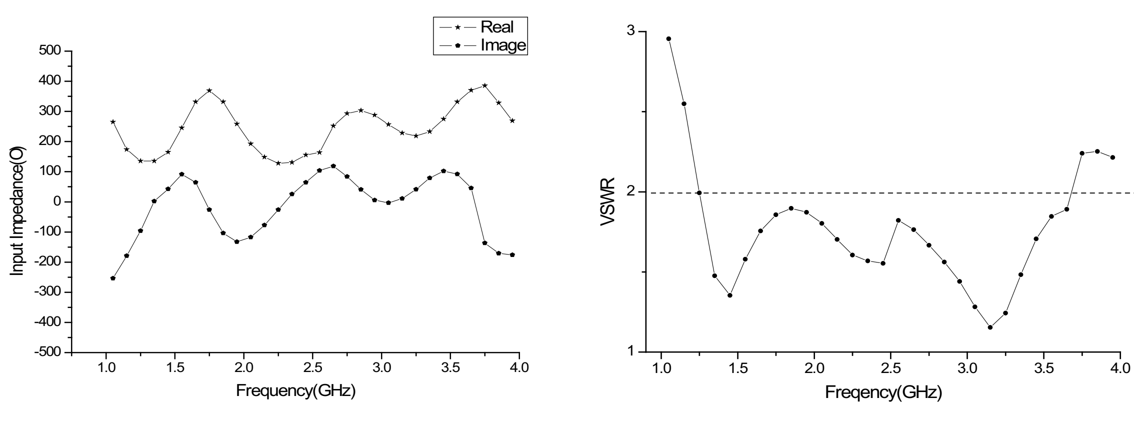


Figure 3: NEC computed input impedance and VSWR of the design quasi-fractal binary tree dipole.

board, and fed by a $50\,\Omega$ coaxial cable through a $1:4$ balun. The return loss of the prototype antenna was measured using a Hewlett-Packard 8510 Network Analyzer. The Fig. 5 compares the computed and measured VSWR of the designed antenna. The measurement verified the designed quasi-fractal binary tree dipole possesses an encouraging capability of broadband.

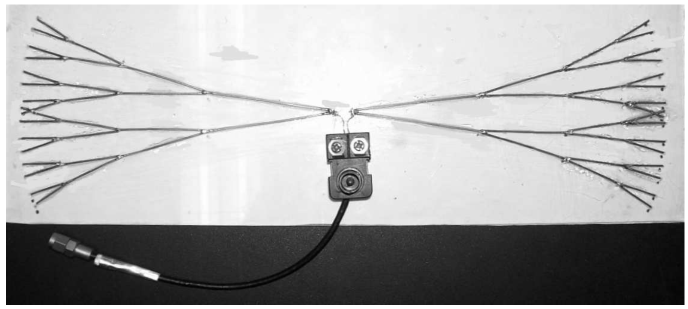


Figure 4: The photo of the fabricated quasi-fractal binary tree dipole.

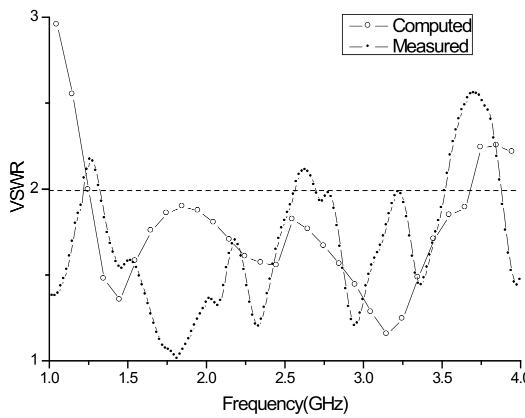


Figure 5: Comparison of the computed and measured VSWR for the quasi-fractal binary tree dipole.

5. CONCLUSION

This paper proposes a dipole antenna based on the quasi-fractal binary tree. The GA in conjunction with the NEC and cluster parallel computation is employed for automated design of the proposed antenna to achieve a wide impedance bandwidth. The designed antenna was fabricated and tested. Results of the measurement show the proposed quasi-fractal binary tree dipole possesses an encouraging capability of broadband.

ACKNOWLEDGMENT

The work was supported by a grant from the National Natural Science Foundation of China (No. 10476013).

REFERENCES

1. Gianvittorio, J. P. and Y. Rahmat-Samii, "Fractal antennas: a novel antenna miniaturization technique, and applications," *IEEE Antenna's and Propagation Magazine*, Vol. 44, No. 1, 20–36, 2002.
2. Petko, J. S. and D. H. Werner, "Miniature reconfigurable three-dimensional fractal tree antennas," *IEEE Trans. on Antennas Propagat.*, Vol. 52, No. 8, 19451956, 2004.
3. Puente, C., J. Claret, and etc., "Multiband properties of a fractal tree antenna generated by electrochemical decomposition," *Electronics Letters*, Vol. 32, No. 25, 2298–2299, 1996.
4. De Werner, H., A. R. Bretones, and B. R. Long, "Radiation characteristics of thin-wire ternary fractal trees," *Electronics Letters*, Vol. 35, No. 8, 609–610, 1996.
5. Goldberg, D. E., *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison-Wesley, New York, 1989.

6. Linden, D. S., "Automated design and optimization of wire antennas using genetic algorithms," Ph.D. Thesis, MIT, Sep. 1997.
7. Jones, E. A. and W. T. Joines, "Design of Yagi-Uda antennas using genetic algorithms," *IEEE Trans. on Antennas Propagat.*, Vol. 45, No. 9, 1386–1391, 1997.
8. Richie, J. E. and H. R. III. Gangl, "EFIE-MFIE hybrid simulation using NEC: VSWR for the WISP experiment," *IEEE Trans. on Electromag. Compat.*, Vol. 37, No. 2, 293–296, 1995.
9. Peng, J., C. A. Balanis, and G. C. Barber, "NEC and ESP codes: guidelines, limitations, and EMC applications," *IEEE Trans. on Electromag. Compa.*, Vol. 35, No. 2, 125–133, 1993.
10. Stutzman, W. L. and G. A. Thiele, *Antenna Theory and Design*, John Wiley and Sons, Inc., 1981.

Hybrid Nyström-PO Method for 3D EM Scattering and Its Application

Shuguang Liu[1,2] **and Xiaojuan Zhang**[2]

[1]Graduate University of Chinese Academy of Science, China
[2]Key Lab of High Power Microwave and EM Radiation, IECAS, China

Abstract— An Nyström-PO Method is introduced to compute the scattering properties of 3D objects with conductor surfaces. With the current-based hybrization, Nyström method and PO method are combined to deal with electrical large objects of arbitrary shape. Examples have shown this method can effectively deal with objects with various surfaces and features.

1. INTRODUCTION

The most recently developed method for electromagnetic scattering computation is the high order Nyström method. It is a point based method utilizing high order surface describing, high order basis functions and high order numerical quadrature rules. High order Nyström method is superior over traditional MoM using RWG basis function [1]. Nyström method is easy to implement, more accurate due to the high order feature and offers the user error control mechanisms [2]. With the recently invented high order interpolatory vector basis [3], the local correction that is needed by the high order Nyström method can be omitted, which makes it more simple.

Nyström method solves the integral equations directly and employs subdomain basis functions which make its use limited by the electrical size of the target. With the advances of modern computers technology, the field that MoM and Nyström method can be employed has expanded a lot, but there are still many interesting objects that can not employ these mature methods to calculate due to their electrical size. High frequency asymptotic techniques such as PO, PTD etc. can be used to calculate the scattering properties. These asymptotic methods are only suit for objects with smooth surfaces and simple edges. For objects that have small feature on surfaces, these methods usually can not get accurate results that reflect these small but important scattering features.

The hybrid Nyström-PO method introduced in this paper is intent to fill this gap. Electrical large objects with fine features on surfaces are common in practical research and no current available methods can handle it easily. The hybrid Nyström-PO method provides an easy way to solve such problems.

2. METHODS

An arbitrarily shaped 3D surface is divided into Nyström region and PO region and further into small patches. The currents are represented by interpolatory vector basis functions over each patch. This is possible because both Nyström and PO methods are current based methods [4]. Define the unitary vector basis over each patch as $(\hat{u}, \hat{v})$. Patch currents can be expressed as

$$\vec{J}_p = \sum_{i=1}^{I_p} \frac{1}{\mathrm{J_{jcb}}} L_{(i,p)}(u,v) \left[\hat{u} \left(g_{22}\hat{u} - g_{12}\hat{v} \right)_{(i,p)} + \hat{v} \left(g_{11}\hat{v} - g_{21}\hat{u} \right)_{(i,p)} \right] \cdot \vec{J}_{(i,p)} \tag{1}$$

The unknown interpolatory current $\vec{J}_{(i,p)}$ can be obtained by PO methods in PO region and by Nyström method in Nyström region.

The PO currents are easy to find out:

$$\vec{J}_{po} = \begin{cases} 2\hat{n} \times \vec{H}_i & \text{illuminated region} \\ 0 & \text{shadowed region} \end{cases} \tag{2}$$

The currents in Nyström region needs to solve an integral equation:

$$\left(\vec{L}_J^E + \vec{L}_J^E K \vec{L}_J^H \right) \left\{ \vec{J}_{\mathrm{N}} \right\}_{\tan} = -\vec{E}_{i,\tan} - \vec{L}_J^E K \vec{H}_i \tag{3}$$

This equation is more complicated than the EFIE used by Nyström method because it has reflected the coupling between Nyström region and PO region. The operator K is the Equation (2) and the operators $\vec{L}_J^E$ and $\vec{L}_J^H$ are:

$$\vec{L}_J^E(\vec{J}) = -\frac{j\omega\mu}{4\pi}\iint_S \left[1+\frac{1}{k^2}\nabla\nabla\right]\vec{J}(\vec{r})\, g\left(\vec{r},\vec{r}'\right) ds' \tag{4}$$

$$\vec{L}_J^H(\vec{J}) = \frac{1}{4\pi}\iint_S \vec{J}(\vec{r})\times\nabla g\left(\vec{r},\vec{r}'\right) ds' \tag{5}$$

$$g\left(\vec{r},\vec{r}'\right) = \frac{1}{|\vec{r}-\vec{r}'|}e^{-jk|\vec{r}-\vec{r}'|} \tag{6}$$

Choose the testing function:

$$\vec{t}(\vec{r}) = \vec{\alpha}\delta\left(\vec{r}-\vec{r}'\right) \tag{7}$$

where $\vec{\alpha}$ is a vector tangential to the surface and in practice it is chosen as $\hat{u}$ and $\hat{v}$.

When an appropriate quadrature rule is applied over each patch to (3), a linear equation system is formed by Nyström discretization scheme.

$$\sum_{i=1}^{P}\sum_{j=1}^{I_P}\begin{bmatrix}A_{11} & A_{12}\\ A_{21} & A_{22}\end{bmatrix}\begin{bmatrix}J_{ij,u}\\ J_{ij,v}\end{bmatrix} = \begin{bmatrix}V_u\\ V_v\end{bmatrix} \tag{8}$$

where

$$A_{\alpha\beta} = \vec{t}_\alpha\cdot\left(\vec{L}_J^E+\vec{L}_J^E K\vec{H}_i\right)\cdot\left[\frac{L_{(j,i)}(u,v)}{\vartheta}\vec{t}_\beta\right] \tag{9}$$

$$V_\alpha = \hat{t}_\alpha\cdot\left(\vec{E}_{i,\tan}+\vec{L}_J^E K\vec{H}_i\right) \tag{10}$$

Solve these equations and obtain currents in Nyström region. With the current in PO region from (2), the RCS can be computed as:

$$\vec{E}^s = -j\omega\mu\left\{\frac{e^{-jkr}}{4\pi r}\int_S \vec{J}\left(\vec{r}'\right)e^{jk\vec{r}'\cdot\vec{r}}ds'\right\}_{\tan} \tag{11}$$

$$\sigma\left(\theta,\phi,\theta^i,\phi^i\right) = \lim_{r\to\infty}4\pi r^2\frac{\left|\vec{E}^s(\theta,\phi)\right|^2}{\left|\vec{E}^i\left(\theta^i,\phi^i\right)\right|^2} \tag{12}$$

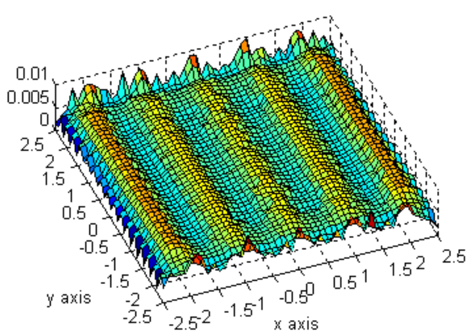


Figure 1: Current distribution of the x component obtained by Nyström method.

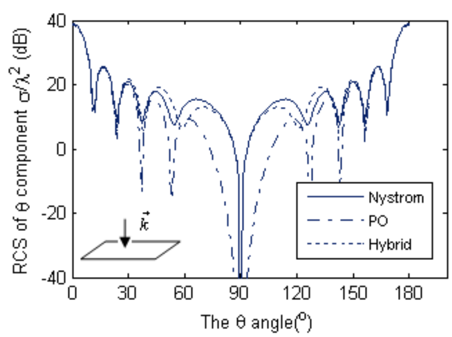


Figure 2: RCS of a $5\lambda\times 5\lambda$ conductor plate, obtained by Nyström method, PO method, and the Nyström-PO hybrid method.

3. APPLICATIONS

To validate the Nyström-PO hybrid method, a simple case is chosen to compute the RCS of flatten perfect conductor surface. The perfect conducting plate has a $5\lambda\times 5\lambda$, size and meshed into 560 patches. The current distribution is shown in Fig. 1 and the RCS are shown in Fig. 2.

The figure shows that the current distribution is reasonable. The computed RCS by Nyström method, PO method and Nyström-PO hybrid method are exactly the same when the observation directions are near the normal direction of the surface. When the direction of observation is near tangential to the surface, the difference between the three methods are increased. But the hybrid method is far more close to Nyström method which is served as reference. The case shows the Nyström-PO hybrid method can give accurate result for the EM scattering problems.

4. CONCLUSIONS

In this paper, a Nyström-PO hybrid method is introduced to compute the scattering properties of conductor surfaces. This method can be applied to both 2D and 3D surfaces of arbitrary shape. Examples have shown this method can give desired accuracy and fast computation. Further analysis will be presented later.

ACKNOWLEDGMENT

Supported by the National Natural Science Foundation of China (No. 60371003).

REFERENCES

1. Rao, S. M., D. R. Wilton, and A. W. Glisson, "Electromagnetic scattering by surfaces of arbitrary shape," *IEEE Trans. Ant. Prop.*, Vol. Ap-30, 409–416, May 1982.
2. Gedney, S., "Application of high-order Nyström scheme to the integral equation solution of electromagnetic interaction problem," *Antennas and Propagation Society, 2000 IEEE International Sym.*, Washington DC, Vol. 1, 356–359, 2000.
3. Graglia, R. D., D. R. Wilton, and A. F. Peterson, "Higher order interpolatory vector bases for computational electromagnetics," *IEEE Trans. Antennas Propagat.*, Vol. 45, 329–342, Mar. 1997.
4. Jakobus, U. and F. J. C. Meyer, "A hybrid physical optics/method of moments numerical technique: Theory, investigation and application," in *Proc. IEEE AFRICON'96*, Stellenbosch, South Africa, Vol. 1, 282–287, 24–27 Sep. 1996.

A Novel Mesh-free Method for Electromagnetic Scattering from a Wire Structure

Min Zhang[1], **Lingxia Li**[1], **Ping Zhou**[2], **and Xiangyang Zhang**[2]

[1]School of Science, Xidian University, Box 270, Xi'an 710071, China
[2]National Electromagnetic Scattering Laboratory, Beijing 100854, China

Abstract— A novel solution scheme of electromagnetic scattering from thin-wire by using method of mesh-free is presented. Based on the mixed potential integral equation (MPIE), the currents of the wire is constructed in terms of moving least squares (MLS) approach, the control function is built according the weighted residual method (WRM). Finally, the axial equivalent current of the wire is obtained and its radar cross section (RCS) is calculated. The effects of influence radius for the accuracy are emphatically investigated.

Mesh-free methods are now proven as robust numerical methods for many field analyses of EM problems. Such methods avoid the onerous mesh generation and adaptive updating, thereby resulting in continuous differentiable approximations that are smooth function and require no post-processing, and therefore, they are fit for many EM problems with large structure and moving conductors. However, by now we discover most of the researchers are working in the static electromagnetic field problems, and few is working on the electromagnetic scattering problem. In this paper, we present a novel solution of electromagnetic scattering from thin-wire excited by an incident plane wave by using mesh-free method. Using MLS with WRM, we calculate the axial equivalent current and RCS of the wire structure. The influence radius of the control equation for the accuracy are emphatically investigated.

The moving least-squares interpolation $u^h(x)$ of the function $u(x)$ is defined in the domain Ω by

$$u^h = \sum_{j=1}^{m} p_j(x)a_j(x) \equiv P^T(x)a(x) \tag{1}$$

where $p_1(x) = 1$ and $p_j(x)$ are monomials in the space co-ordinates $X^T = [x,\, y]$ so that the basis is complete.

A linear basis in a two-dimensional domain is chosen by

$$P^T = (1,\, x,\, y), \quad m = 3, \tag{2}$$

or by a quadratic basis function

$$P^T = (1,\, x,\, y,\, x^2,\, xy,\, y^2), \quad m = 6, \tag{3}$$

with a cubic basis function

$$P^T = (1,\, x,\, x^2,\, xy,\, y^2,\, x^3,\; x^2y,\; xy^2,\, y^3), \quad m = 10 \tag{4}$$

here, m is the numbers of basis functions.

Thus, we could present the current J as

$$J = \sum_{I=1}^{n} w\left(x - x_i\right)\left[P^T(x)a(x) - u_i\right]^2 \tag{5}$$

where, n is the number of points in the influence domain for which the weight function $w\left(x - x_i\right) \neq 0$, and u_i is the node value of u at $x = x_i$.

The stationary of J in (5) with respect to $a(x)$ leads to the following linear relation between $a(x)$ and u_i:

$$A(x)a(x) = B(x)u, \tag{6}$$

where $A(x)$ and $B(x)$ are matrices defined by

$$A(x) = \sum_{I=1}^{n} w(x - x_i)p(x_i)p^T(x_i) = p^T w(x)p \tag{7}$$

$$B(x) = \sum_{I=1}^{n} w(x - x_i)p(x_i) = p^T w(x), \tag{8}$$

Correspondingly, $u^T = [u_1, u_2, \ldots, u_n]$.
Hence, we have

$$u^h(x) = \sum_{i=1}^{n} N_i(x)u_i = N(x)u, \tag{9}$$

where the shape function $N(x)$ is defined by

$$N(x) = p^T(x)A^{-1}(x)B(x), \tag{10}$$

The scattered field $\boldsymbol{E}^s$ is

$$\boldsymbol{E}^s = -j\omega\bar{\boldsymbol{A}} - \nabla\Phi, \tag{11}$$

$n \times \left(E^i + E^s\right) = 0$ here, $A = jk\eta \int J \frac{e^{-jkR}}{R} dx'$ and $\Phi = \frac{\eta}{jk} \int \sigma \frac{e^{-jkR}}{R} dx'$, according the boundary condition on the surface of thin-wire, the MPIE is built as

$$\boldsymbol{n} \times \left(\boldsymbol{E}^i + \boldsymbol{E}^s\right) = 0, \tag{12}$$

Discrete equations using Galerkin method, use the test function as N_m in Ω zone, $-N_m$ on its boundary Γ, thus the weak form functional corresponding to the above boundary value problem is:

$$(u_i) = \delta\Pi \int_{\Omega} N_m \left(\boldsymbol{E}^s + j\omega\bar{\boldsymbol{A}} + \nabla\phi\right) dx + \int_{\Gamma} -N_m \left(\boldsymbol{E}^s + \boldsymbol{E}^i\right) dx = 0 \tag{13}$$

In order to solve for the induced axial current on this body, we use mesh-free method solution with the expansion function, which can be written in form as

$$J = \sum_{n=1}^{N} I_n N_n \tag{14}$$

where, I_n are the unknown coefficient, N_n are the shape function got by moving least squares approximation.

By substituting Eq. (14) into Eq. (13), we have

$$ZI = V \tag{15}$$

Eq. (15) can now be solved to obtain the currents induced along the axis of the thin-wire due to the incident field, and then we can developing other electromagnetic scattering properties.

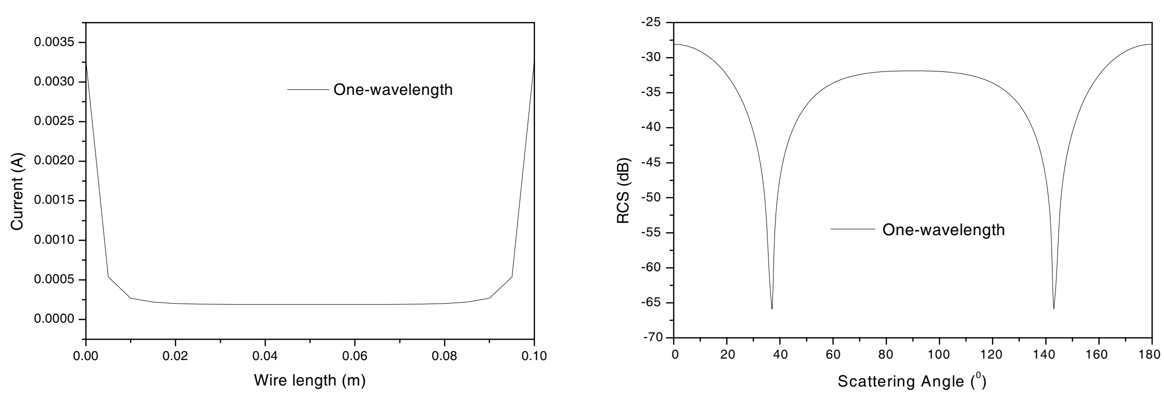


Figure 1: The induced current and RCS of a 1λ wire structure.

We present the thin-wire as example to test the validity of the mesh-free method for electromagnetic scattering. A one-wavelength (1λ) and a ten-wavelength (10λ) straight wire illuminated by an

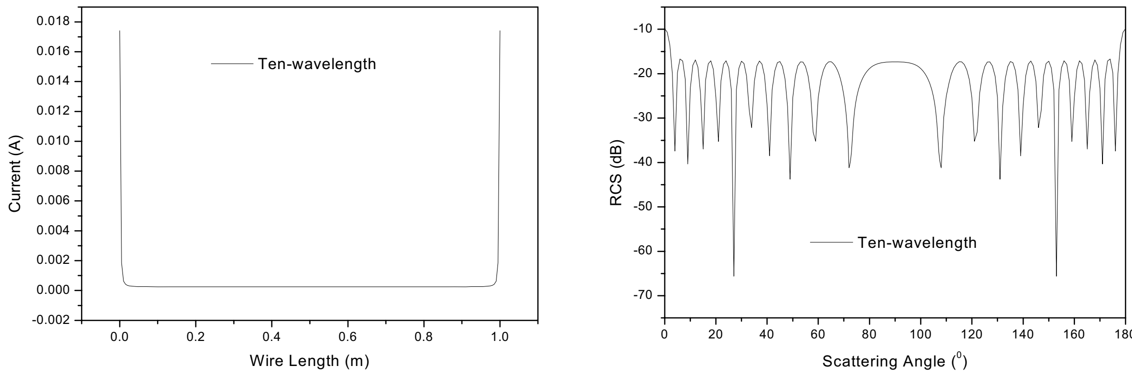


Figure 2: The induced current and RCS of a 10λ wire structure.

axially polarized are considered respectively, the frequency of normally incident wave is 3 GHz. By the mesh-free method, the wire is divided into 20 symmetrical cell per wavelength, and influence radius is 3.4. Fig. 1 shows the induced current and back RCS in TM case of a 1λ wire structure. Correspondingly Fig. 2 indicates the current and RCS of a 10λ wire structure.

It is discovered that the influence radius acts major point on the accuracy of modeling. A proper selection of influence radius of the test function and trial function is very important in mesh-free method. In present method, the basis function and the trial function in the element-free Galerkin method come from the same function space, so the influence radius of the weighting function decided to influence radius of the test function and the trial function. Fig. 3 presents the choice different influence radius, such as 3.4, 3.5 and 3.6, respectively, making for the different results of induced current and RCS. As we known, by now most of the influence radius is the experience value.

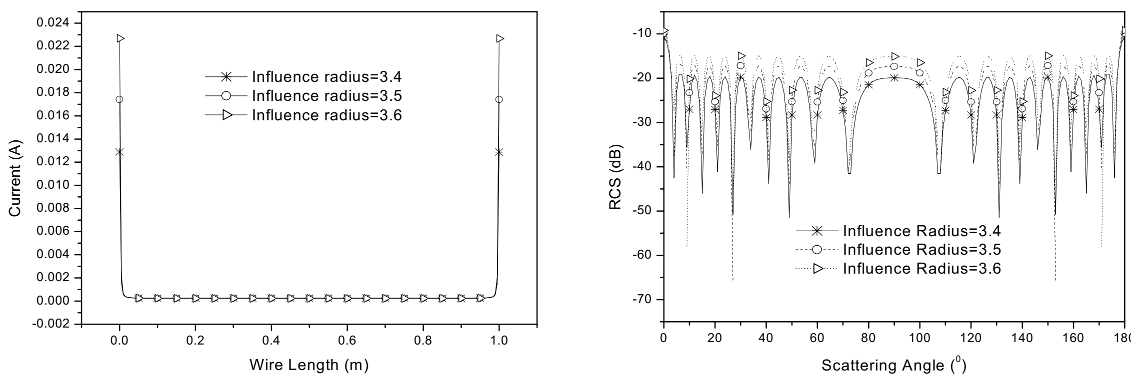


Figure 3: The infection for the induced current and RCS by different influence radius.

In conclusion, the Garlerkin-based mesh-free method was applied for electromagnetic scattering from a wire structure. The solution method presented in this work is efficient and will be used in scattering problem of large structure in our future work.

REFERENCES

1. Cingoski, V., N. Miyamoto, and H. Yamashita, "Element-free Galerkin for electromagnetics field computations," *IEEE Trans. on Mag.*, Vol. 34, No. 5, 236–3239, 1998.
2. Ho, S. L., J. Yang, M. Machado, and H. C. Wong, "Application of a meshless method in electromagnetics," *IEEE Trans. on Mag.*, Vol. 37, No. 5, 3198–3202, 2001.
3. H6rault, C. and Y. Mar6chal, "Boundary and interface condition in meshless methods," *IEEE Trans. on Mag.*, Vol. 35, No. 3, 1450–1453, 1999.
4. Zhang, Y., K. R. Shao, D. X. Xie, and J. D. Lavers, "Meshless method based on orthogonal basis for computational electromagnetics," *IEEE Trans. on Mag.*, Vol. 41, No. 5, 1432–1435, 2005.

Evaluating Surface Impedance Models for Terahertz Frequencies at Room Temperature

S. Lucyszyn

Imperial College London, UK

Abstract— Commercial electromagnetic modeling software employs overly-simplified models for the terahertz simulation of metal structures. For the first time, this paper gives a unique review of various modeling strategies (classical, semiclassical and quantum mechanical based) for normal metals and discusses their limitations with frequency at room temperature.

1. INTRODUCTION

High frequency CAD software packages employ overly-simplified models for the electromagnetic simulation of metal structures; using either classical skin-effect or classical relaxation-effect models. At room temperatures, these models are accurate beyond the upper edges of the microwave and sub-millimeter-wave parts of the frequency spectrum, respectively. However, semiclassical models are needed to extend modeling well into the terahertz region or at significantly lower temperatures. Here, issues relating to the specular or diffuse nature of electron reflections at the air-metal interface become apparent at very low temperatures.

Within the near-infrared, visible and ultra-violet parts of the frequency spectrum, Commercial CAD software packages again may employ overly-simplified empirically-fitted relaxation-effect models, which only work over relatively narrow spectral bandwidths. However, to be accurate, an analytical model must be adopted that employs a quantum mechanical treatment, as this takes into account both energy dispersion and electron wavefunctions.

The author has investigated modeling strategies for normal metals. In one study, experimental measurements that suggested the possibility of anomalous room-temperature conduction losses were examined between DC and 12.5 THz [1]. It was found that the classical relaxation-effect model was still valid up to these frequencies. In another study, an elaborate semiclassical model to describe anomalous excess conduction losses at room temperature was found to be completely erroneous [2]. In order to create accurate analytical models, it is important to develop semiclassical modeling strategies [3] or develop quantum mechanical treatments. To this end, and for the first time, this paper will review various approaches to the modeling of normal metals at room temperature. More importantly, their limitations will be discussed in detail. It will be shown that a number of well-know approaches have severe limitations to general applications.

2. CLASSICAL TREATMENT

Drude's model of intraband transitions describes an ideal system of free electrons having a spherical Fermi surface. The classical relaxation-effect model takes into account electron-phonon collisions, represented by the following expression for surface impedance, Z_{SR}, in terms of Drude's model for intrinsic bulk conductivity, σ_R:

$$Z_{SR} = \sqrt{\frac{j\omega\mu_o\mu_r}{\sigma_R + j\omega\varepsilon_o}} \quad \text{where} \quad \sigma_R = \frac{\sigma_o}{(1+j\omega\tau)} \tag{1}$$

where, angular frequency, $\omega = 2\pi f$; and f = frequency of the driving electric field; μ_o = permeability of free space; μ_r = relative permeability; ε_o = permittivity of free space; σ_o = intrinsic bulk conductivity at DC; and τ = phenomenological scattering relaxation time for the free electrons (i.e., mean time between collisions).

Note that the complex operator $+j$ is used throughout and that this replaces $-i$ used in some cited references. Failure to adopt a consistent notation can result in errors. It is found that at room temperature and at sufficiently low frequencies, (1) reduces to the classical skin-effect model:

$$Z_{SR} \approx \sqrt{\frac{j\omega\mu_o\mu_r}{\sigma_0}} = R_o(1+j) \quad \text{when} \quad \omega\tau << 1 \tag{2}$$

where, R_o = classical skin-effect surface resistance and the displacement current term can been ignored.

A rigorous investigation into the robustness of the classical relaxation-effect model, when compared to measured data, found that it works well from DC to the lower edge of the mid-infrared frequency range, for normal metals at room temperature [1]. This finding helped to dispel a long-standing myth that an anomalous intrinsic conduction loss in normal metals exists at room temperature.

There are several reasons why the classical relaxation-effect model is not sufficient for a qantitative account of experimental observation, even when interband transitions can be safely ignored or calculated separately:

1) Conduction bands with normal metals can be significantly different from the ideal spherical energy band (that assumes an effective electron mass, m).
2) Scattering relaxation time is assumed to be independent of energy $E(k)$, where k = wave vector of the electron, even though it depends on both energy and position. For this reason, τ is considered to be a semi-empirical parameter.
3) A local-response regime is assumed.

For the above reasons, to achieve even better modeling accuracy, it is necessary to move towards the more complicated semiclassical treatment.

3. SEMICLASSICAL TREATMENT

It is well known that, at sufficiently high frequencies, the conductivity of normal metals exhibits both temporal (i.e., frequency) and spatial (e.g., one-dimensional) dispersion. Harrison introduced a frequency- and wave number-dependent dielectric function for a semiclassical free-electron gas [4]. This model describes the screened Coulomb potential effect that a spatial charge-density fluctuation has on a free electron as it travels through a periodic lattice of fixed positive ions. It can be helpful to think of induced conduction current within a normal metal as flowing in lamina-type sheets: almost parallel to the surface (x-y-plane) of the conductor and having an amplitude that decays exponentially from its surface into the bulk material. The electric and magnetic fields and conduction current distributions inside the metal have time t and spatial (positive z-direction) variations of the form:

$$e^{(j\omega t-\gamma z)} \tag{3}$$

where, propagation constant, $\gamma = \alpha + j\beta$; α = attenuation constant; β = phase constant; and $\gamma = jq$, where modified wave number of the driving electric field, $q = q' - jq''$.

Wang assumed that the periodic nature of the conduction current density gives rise to a screening potential effect on free electrons as they travel in a direction perpendicular to the surface of the air-metal boundary [5]. With spatial charge-density fluctuations being attributed to lamina-type sheets of conduction currents, as an analogy to a periodic lattice of fixed positive ions, Wang adopted Harrison's screening potential theory. To this end, Harrison's semiclassical expression for intrinsic bulk conductivity, σ_H, was used to calculate Wang's surface impedance, Z_{SW} [2]:

$$Z_{SW}=\sqrt{\frac{j\omega\mu_o\mu_r}{\sigma_H(s,\omega)}} \text{ where } \sigma_H(s,\omega)=\sigma_R\left\{\frac{-3j\omega\tau}{s^2}\cdot\frac{\left[2s-\ln\left(\frac{1+s}{1-s}\right)\right]}{2sj\omega\tau+\left[2s-\ln\left(\frac{1+s}{1-s}\right)\right]}\right\} \text{ and } s=\frac{\pm jql_m}{(1+j\omega\tau)} \tag{4}$$

where, l_m = mean distance traveled by the electron between collisions (i. e., mean-free path length).

In deriving (4), it was assumed that electrons at all angles, with respect to the metal surface in the half-space, will contribute to the conductivity within this semiclassical free-electron gas analysis [5]. Moreover, the modified wave number considered here is at least two orders of magnitude lower than the Fermi wave number and, therefore, a quantum mechanical analysis is not necessary. It was poorly assumed that with such long wavelengths of potential there is no difference between conductivities for transverse and longitudinal fields in cubic materials [5]. Wang then applied his spatial dispersion theory of excess conduction loss to room temperature measurements of copper, from DC up to 6.7 THz, using the measured data at 35 GHz and 70 GHz reported by Tischer [1, 2].

It is evident that the expression for intrinsic bulk conductivity, quoted by both Harrison and Wang, is derived for longitudinal wave propagation [2]. As a result, his model has no meaning for surface impedance and excess conduction loss calculations (as they are based on transverse wave propagation for normal incidence). Details of a rigorous de-construction of Wang's model have previously been published [2]. From this detailed investigation, it was clearly shown that not only are there serious discrepancies within the general methodology, but even the data on which it is based has unacceptable errors.

An accurate semiclassical treatment was first reported by Reuter and Sondheimer, back in 1948, for transverse wave propagation [6]. Their approach gives an exact solution for all frequencies and temperatures. Moreover, their general methodology can be for any value of specular reflection coefficient, p; from diffuse (completely random) scattering, $p = 0$, to specular (mirror-type) surface reflections, $p = 1$. With the latter, surface impedance Zs_{RS} can be calculated from the following:

$$\begin{aligned} Zs_{RS} &= \frac{j\omega\mu_o\mu_r}{\gamma(o)} \quad \text{where} \quad \gamma(o) = \left[-l_m \frac{f(0)}{f'(0)}\bigg|_{p=1} \right]^{-1} \\ \text{and} \quad \frac{f(0)}{f'(0)}\bigg|_{p=1} &= \frac{2}{j\pi(1+j\omega\tau)} \int_0^{\frac{j\infty}{(1+j\omega\tau)}} \frac{1}{s^2 + \frac{\omega^2\mu_o\mu_r\varepsilon_o l_m^2}{(1+j\omega\tau)^2} - \xi K(s)} \cdot ds \\ \text{and} \quad \xi &= \frac{j\varsigma}{(1+j\omega\tau)^3}; \quad \varsigma = \frac{3}{4}\omega\mu_o\mu_r\sigma_o l_m^2; \quad K(s) = \frac{1}{s^3}\left[2s - (1-s^2)\ln\left(\frac{1+s}{1-s}\right) \right] \end{aligned} \tag{5}$$

This modeling approach assumes the following:

1) Skin depth is much less than linear dimensions of metal, thus regarding it as planar and infinite in extent.
2) Normal incidence of propagation, thus simplifying to a one-dimensional problem.
3) Conduction electrons are quasi-free, having a kinetic energy $E(k) = \hbar|k|^2/2m$ with a parabolic band approximation, where $\hbar$ = modified Planck's constant.
4) Collision mechanism is always described in terms of l_m.
5) A fraction p of electrons arriving at the surface is scattered specularly, while the rest are scattered diffusely.
6) This one-band free-electron model does not apply to multi-valent metals, in which the electrons occupy more than one energy band (e.g., aluminium or tin).
7) $K(s)$ is derived for transverse conductivity, and holds for the whole (q, ω) plane, for a spherical energy band.
8) While the semiclassical treatment takes into account partially-filled conduction band energy dispersion, $E(k)$, it ignores electron wavefunctions.

It is interesting to note that when $|s| \ll 1$, e.g., with the long-wavelength limit of $q \to 0$, then $K(s) = 4/3$ in (5) and the associated intrinsic bulk conductivity is equal to Drude's model for intrinsic bulk conductivity, σ_R, given in (1).

4. QUANTUM MECHANICAL TREATMENT

Zhang and Pan use a quantum mechanical approach to derive, in detail [7], the following expression for the dielectric function of the free-electron gas [7, 8]:

$$\varepsilon_r(q,\omega) = 1 + \frac{2e^2}{q^2\varepsilon_o}\frac{1}{V}\sum_k \frac{f(k) - f(k+q)}{E(k+q) - E(k) - \hbar\omega + j\zeta} \tag{6}$$

where, e = electron charge; V = volume; $f(k)$ = Fermi-Dirac distribution function. This is a standard derivation and the result is known as the Lindhard dielectric function; also know as the self-consistent-field (SCF) or random phase approximation (RPA) dielectric function [9]. By replacing $f(k+q)$ with $f(-k)$ and taking the long wavelength limit, Zhang and Pan highlight the following [7, 8]:

$$\frac{1}{V}\sum_k f(k) = \frac{n}{2}; \quad \sigma_o = \frac{ne^2\tau}{m}; \quad v_f = \frac{\hbar k_f}{m}; \quad \zeta = \frac{\hbar}{2\tau} \tag{7}$$

where n = electron density, v_f = velocity between collisions of the free electron, having kinetic energy at the Fermi level and k_f = Fermi wave number; as a means to somehow derive an expression for their intrinsic bulk conductivity, in expressions for dielectric function and surface impedance [7, 8]:

$$\varepsilon_{rZP}(q,\omega) = 1 + \frac{\sigma_{ZP}(q,\omega)}{j\omega\varepsilon_o} \quad \text{where} \quad \sigma_{ZP}(q,\omega) = \frac{\sigma_o}{1 + j\omega\tau\left[1 - \frac{3}{5}\left(\frac{ql_m}{\omega\tau}\right)^2\right]}$$
$$\therefore Zs_{ZP} = \sqrt{\frac{j\omega\mu_o\mu_r}{\sigma_{ZP}(q,\omega) + j\omega\varepsilon_o}} \tag{8}$$

Unfortunately, the original expression for the dielectric function given in (6) is only valid for longitudinal wave propagation [9]. As a result, the Zhang and Pan model has no meaning for their surface impedance, Zs_{ZP}, and excess conduction loss calculations.

It is evident from the publications of Zhang and Pan that they have made the same mistake as Wang, by trying to fit the wrong type of theoretical model to measured data. It has been previously demonstrated that the relaxation-effect model is adequate for characterising the intrinsic frequency dispersive nature of normal metals at room temperature, even into the terahertz frequency range. Indeed, the results from a quantum mechanical model should actually converge onto those from the relaxation-effect model at these frequencies, at room temperature. Since the room temperature results, from the model of Zhang and Pan, deviate from the classical relaxation-effect model at relatively low frequencies then this alone points to a fundamental error.

5. PROOF OF CONTRADITIONS WITH MODELS (4) AND (8)

Apart from the lack of any physical insight to justify the need for new models and the obvious misuse of longitudinal wave propagation terms for calculating surface impedance, the work published by Wang and also by Zhang and Pan also share a couple of fundamental contradictions.

5.1. First Contradiction

Intrinsic bulk conductivity exhibits spatial dispersion in the non-local response regime and, therefore, has a wave number dependency in reciprocal q-space. For one-dimensional propagation along the positive z-axis, the general expression for conduction current density, Jc, in terms of intrinsic bulk conductivity and electric field, E, is given by the non-local constitutive equation in real space [10]:

$$Jc_y(z)\Big|_{p=1} = \int_{-\infty}^{+\infty} \sigma(z - z')E_y(z').dz' \tag{9}$$

Now, at the surface of the metal, conduction current density vector, $Jc_y(0)$, and surface current density vector, Js, are related by the propagation constant at the surface, $\gamma(0)$, as follows:

$$Jc_y(0) = \gamma(0)Js \quad \text{where} \quad Js = n \times H_x(0) \tag{10}$$

where $n =$ unit vector pointing out normal to the surface of the metal and $H_x(0) =$ magnetic field vector at the surface of the metal. Equation (9) can be expressed in terms of a simple convolution (denoted by the symbol *) of the intrinsic bulk conductivity and electric field as follows:

$$\therefore Jc_y(z)\Big|_{p=1} \equiv \{\sigma(z)^* E_y(z)\} \tag{11}$$

But, surface impedance for normal incidence is related to the electric and magnetic fields as follows using Ohms law [9, 10]:

$$Zs = \frac{E_y(0)}{H_x(0)} \tag{12}$$

It can be easily seen that determining surface impedance is not so straight forward in the non-local response regime, since:

$$\left\{\sigma(z)^* E_y(z)\right\}\big|_{z=0} \equiv \gamma(0)H_x(0) \tag{13}$$

Now, taking the Fourier transform of both sides of (11) gives the non-local constitutive equation in reciprocal q-space [10]:

$$Jc(q)\Big|_{p=1} = \sigma(q)E(q) \tag{14}$$

Surface impedance, by its very nature, must be expressed within real space and so (14) is of no direct use. In the local response regime, however, the generalised conductivity behaves like a δ-function, i.e., $\sigma(z-z') \to \sigma\delta(z-z')$.

$$\therefore \int_{-\infty}^{+\infty} \delta(z-z')E_y(z').dz' = E_y(z) \quad \text{and} \quad Jc_y(z)\Big|_{p=1} \to \sigma E_y(z) \tag{15}$$

Therefore, in the local response regime, which represents the long-wavelength limit $q \to 0$ in reciprocal q-space, and ignoring displacement current, the simple expression for surface impedance can be easily determined using (10), (12) and (15):

$$Zs = \frac{\gamma(0)}{\sigma} = \sqrt{\frac{j\omega\mu_o\mu_r}{\sigma}} \quad \text{where} \quad \sigma \neq f(q) \tag{16}$$

Evidently, (16) could only have been derived here if the intrinsic bulk conductivity is in the local response regime. Since Wang's semiclassical model and the work by Zhang and Pan both adopt (16) with intrinsic bulk conductivity having spatial dispersion, i.e., in a non-local response regime, this is direct proof of a fundamental contradiction.

5.2. Second Contradiction

Angular frequency ω and modified wave number q are assumed to be real variables by both Wang [5] and also by Zhang [11], but this must be a complex variable when modeling normal metals at room temperatures and at frequencies below the plasma frequency, as previously stated by Lucyszyn [2]. In general, either the angular frequency ω or the modified wave number q must be a complex term, in order to account for an exponentially decaying wave as it propagates into the metal. The dispersion relations for electromagnetic waves in an isotropic homogeneous medium has either q being complex and $q = f(\omega)$, where ω is real, or ω being complex and $\omega = f(q)$, where q is real [9]. Only if there is no dissipation of electromagnetic energy can both q and ω be real [9]. Therefore, since both are treated as real variables, this contradicts the principle of conservation of energy for wave propagation within a normal metal; once again challenging the validity of (4) and (8).

6. CONCLUSIONS

It would be inappropriate to use either Wang's semiclassical model or the work by Zhang and Pan to support the view that anomalous intrinsic frequency dispersion exists within normal metals at room temperature. When the findings from this investigation are combined with those recently published on the lack of experimental evidence, any myth associated with anomalous behavior at room temperature can be finally dispelled; the relaxation-effect model is sufficiently accurate to describe the natural behavior of normal metals at room temperature. It has already been shown that measured data up to 12.5 THz actually fits the classical relaxation-effect model quite well [1]. This is good news for those working between circa 30 GHz and 12 THz. For example, within future measurement systems, the mathematically simple Drude model can be used to characterize the surface impedance for calibration standards (with accurate values for only σ_o and τ being needed by the metrologist). This approach would lead to much more accurately calibrated sub-mm-wave measurement systems. For even greater accuracy, the complicated semiclassical approach based on the work of Reuter and Sondheimer are recommended, especially when modeling is to be undertaken for applications well below room temperature.

REFERENCES

1. Lucyszyn, S., "Investigation of anomalous room temperature conduction losses in normal metals at terahertz frequencies," *IEE Proc. — Microwaves, Antennas and Propagation*, Vol. 151, No. 4, 321–329, Aug. 2004.
2. Lucyszyn, S., "Investigation of Wang's model for room temperature conduction losses in normal metals at terahertz frequencies," *IEEE Transactions on Microwave Theory Tech.*, Vol. 53, No. 4, 1398–1403, Apr. 2005.
3. Lucyszyn, S., "Accurate CAD modelling of metal conduction losses at terahertz frequencies," *11th IEEE International Symposium on Electron Devices for Microwave and Optoelectronic Applications (EDMO2003)*, 180–185, Orlando, USA, Nov. 2003.
4. Harrison, W. A., *Solid State Theory*, 280–290, McGraw-Hill, 1970.

5. Wang, Y.-C., "The screening potential theory of excess conduction loss at millimeter and submillimeter wavelengths," *IEEE Trans. on Microwave Theory Tech.*, Vol. MTT-26, No. 11, 858–861, Nov. 1978.
6. Reuter, G. E. H. and E. H. Sondheimer, "The theory of the anomalous skin effect in metals," *Proc. Royal Society*, Vol. A195, 336–364, London, 1948.
7. Zhang, X., W. Pan, and S. Lucyszyn, "A new theoretical modeling of surface resistance in normal metals at terahertz frequencies," *International Journal of Infrared and Millimeter Waves*, Vol. 25, No. 11, 1611–1620, Springer Science+Business Media Inc., Nov. 2004.
8. Pan, W. and X. Zhang, "Study of excess conduction loss in normal metals at/below submillimeter wavelengths," *International Journal of Infrared and Millimeter Waves*, Vol. 27, No. 3, 455–464, Springer Science+Business Media Inc., Mar. 2006.
9. Fuchs, R. and P. Halevi, "Basic concepts and formalism of spatial dispersion," *Spatial Dispersion in Solids and Plasmas*, Halevi, P. (Editor), Chapter 1, Published by North-Holland, 1992.
10. Grosso, G. and G. P. Parravicini, *Solid State Physics*, Chapter 11, 389–414, Published by Academic Press, 2000.
11. Zhang, X., *Private Communication*, Sep. 2004.

Reduction of Monostatic RCS by Switchable FSS Elements

M. Khosravi and M. S. Abrishamian

K. N. Toosi University of Technology, Tehran, Iran

Abstract— A special kind of active FSS element incorporating switched PIN diodes is simulated. There are two different geometries; each one is related to a different on-off state of diode switches. Both geometries, square loop and four-legged loaded element, are excited by a uniform plane wave which its angle of incidence is adjustable. The transmission characteristic is identified easily without complex resonances obscuring the results.

1. INTRODUCTION

Frequency selective surfaces (FSSs) are periodic structures that can provide frequency filtering to incoming electromagnetic waves [1]. Frequency-selective surfaces have found many applications in antenna systems. These include multibanding reflector antennas and providing frequency windows in radomes. Many element geometries have been studied ranging from dipoles and Jerusalem crosses to square and circular loops [2].

Moreover, the capabilities of the FSS have been extended by the addition of active devices embedded in the unit cell of the periodic structures. Such structures are also called active grid arrays [1]. An active frequency selective surface is capable of changing its frequency transmission response from a resonating or reflecting structure, to one which is virtually transparent except for a small insertion loss [3]. This change in response, which is electronically controlled [4], could find applications in areas including RCS reduction, reconfigurable reflector systems and active waveguide horns [5].

In this paper, we present a kind of active FSS elements incorporating switched PIN diodes. The fundamental design principle was to create FSS elements whose geometry changed significantly on switching the diodes on and off; hence there would be a significant difference in the frequency response. As shown in Fig. 1, the L-sections are connected using PIN diodes in two different cases: First, the geometry is a square loop; and second, the state in which the geometry becomes a four-legged loaded element. Square loop geometry is commonly employed as FSS element. It has a simple frequency response consisting of a transmission band followed by a reflection band at which the elements resonate. This allows the transmission characteristics to easily be identified without complex resonances obscuring the results.

Ansoft HFSS (High Frequency Structure Simulator) is an interactive software package for calculating the electromagnetic behavior of a structure. HFSS uses Finite Element Method (FEM) to compute solutions. It allows for the computation of the basic electromagnetic field quantities, S-parameters, and resonance of a structure. In this paper, HFSS was used to model various FSSs.

2. ACTIVE FSS WITH SQUARE LOOP/FOUR-LEGGED LOADED ELEMENTS

In the model proposed here, PIN diodes placed in the structure could be used to switch the basic geometry of the elements and fundamentally change the frequency characteristics. The diodes are connected into the arms so that their axes are parallel to the incident electric field vector. Thin diode bias lines are introduced which are perpendicular to the electric field. Inserting the diodes as mentioned has the property of reducing the overall equivalent capacitance of the structure as discussed in [3]. When the diodes are forward biased they form a low resistance path and the sides are connected together and with the reverse biased case they act essentially as an open circuit between the L-sections. So by biasing various proper groups of diodes, the surface looks like an infinite array of elements which has its own resonance frequency.

Square loops have a simple frequency response consisting of a transmission band followed by a reflection band at which the elements resonate. This allows the transmission characteristics to easily be identified without complex resonances obscuring the results. Another possible geometry is four-legged loaded loop element which its development is quite instructive. Such as other loaded elements, one of the most important features of this element is the ability to control bandwidth by modification of the load impedance.

3. HFSS MODELING OF FSS ELEMENTS AND RESULTS

Because of periodicity inherent in the geometry, only the fields in the "unit cell" are unique and need to be determined the edges of the unit cell. In Ansoft Designer, an infinite array represents a part of model geometry (unit cell) that repeats itself infinitely. When exciting the array with a plane wave, the structure will act as a frequency-selective surface (FSS), or a filter receiving the excitation.Geometries such as square loops are commonly employed as FSS elements. Fig. 2 shows the basic geometry and dimensions of the square loop element. The dimensions of the surface were approximately $d = 7.0\,\text{mm}$, $w = 1\,\text{mm}$ and $g = 3.2\,\text{mm}$ and the elements were printed on the RT Duroid substrate 0.25 thick with $\varepsilon_r = 2.33$. It was excited by a uniform plane wave for two different angles of incidence: first normal incidence ($\theta = 0°$) and then oblique one ($\theta = 45°$). The simulation had a frequency range from 8 to 18 Giga-Hertz. Fig. 3 and Fig. 4 demonstrate the reflectivity and transmission response for the active square loop array, respectively. Responses are in agree with results in [3]. For the four-legged loaded element everything is the same and just frequency range has changed (from 0 to 10 GHz). The frequency responses of this active FSS array are shown in Figs. 5 and 6. The four legged loop FSS modeled is one whose frequency response is published in Munk [2].

4. FIGURES

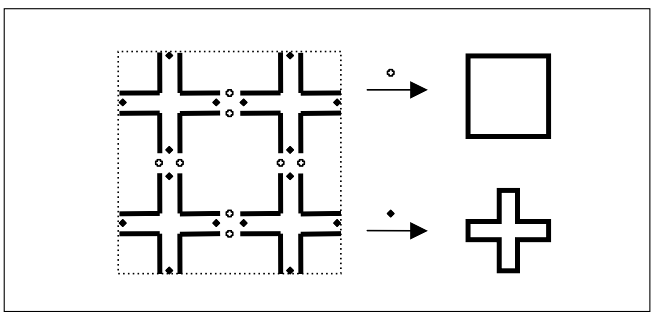


Figure 1: Active FSS square loop/ four-legged loaded element.

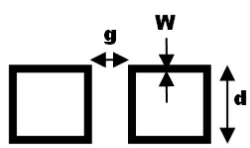


Figure 2: Active square loop geometry (depicting dimension parameters).

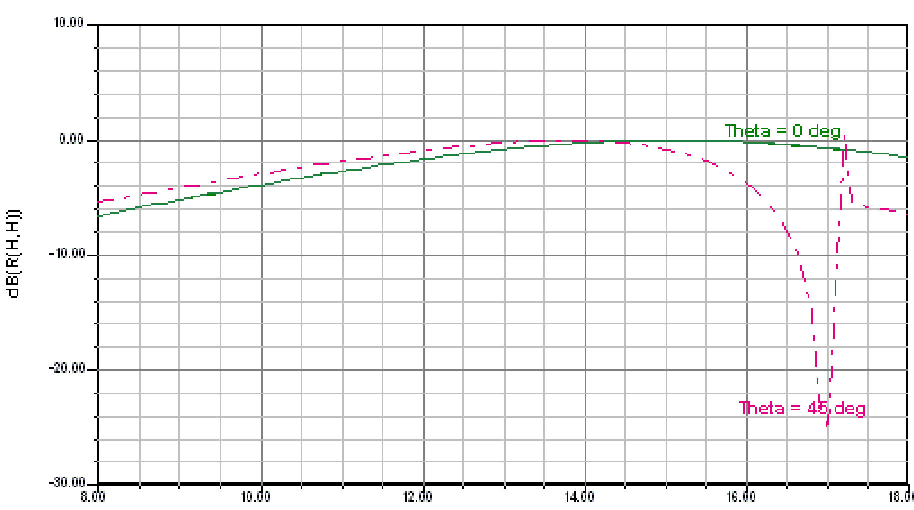


Figure 3: Reflectivity response of square loop element FSS.

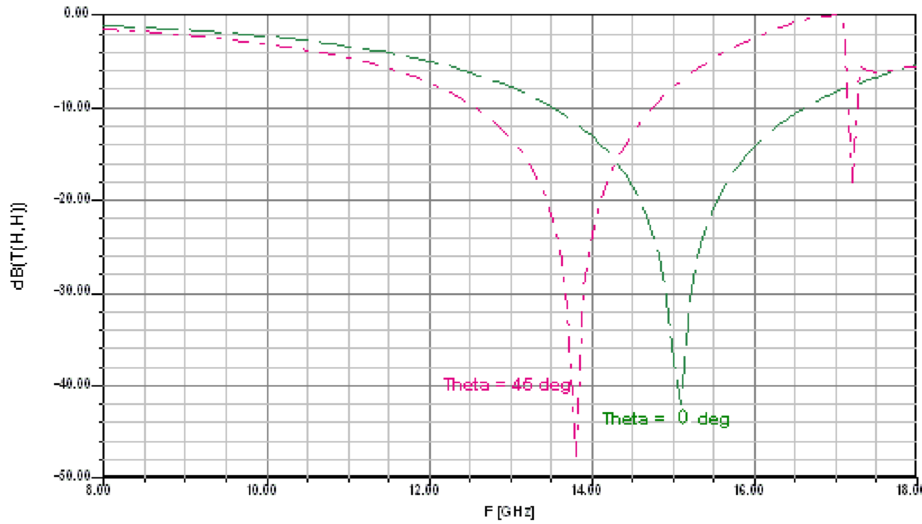


Figure 4: Transmission response of square loop element FSS.

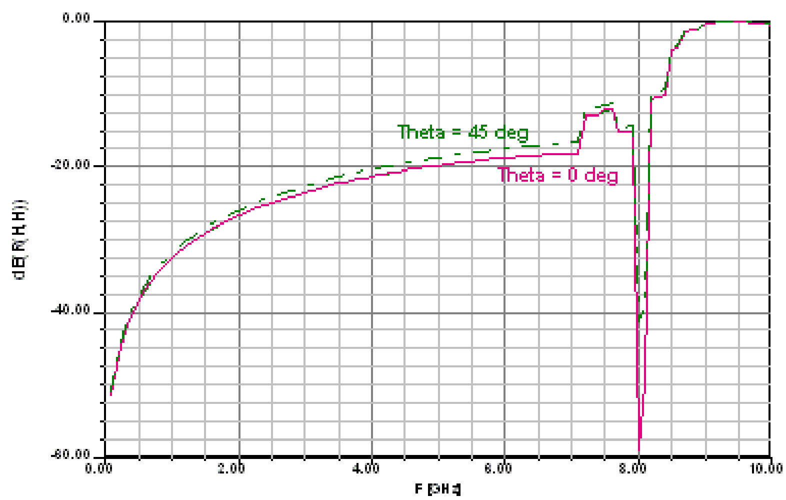


Figure 5: Reflectivity response of four-legged loaded element FSS.

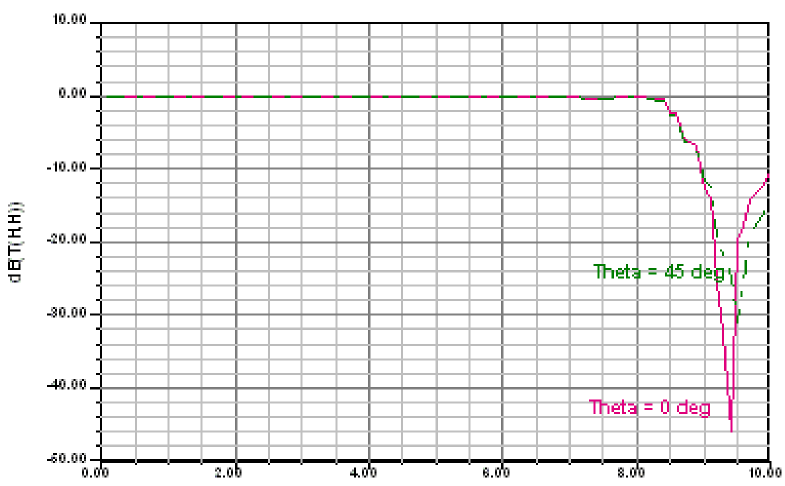


Figure 6: Transmission response of four-legged loaded element FSS.

5. CONCLUSIONS

The operation of active FSS whose reflectivity can be tuned and switched electronically as a function of time have been demonstrated. It has been shown that active devices placed in FSS elements fundamentally change the geometry and frequency response of the arrays. It has also been shown that HFSS can be used to model infinite frequency selective surfaces. As final results show, square loop element is convenient for higher frequencies and four-legged loaded for lower ones. Therefore, desired frequency response is achievable by switching proper diodes.

ACKNOWLEDGMENT

Authors appreciate Iranian Telecommunication Research Center (ITRC) which provided funding for this work.

REFERENCES

1. Wu, T. K., *Frequency Selective Surfaces and Grid Arrays*, Wiley-Interscience, New York, 1995.
2. Munk, B. A., *Frequency Selective Surfaces*, Wiley-Interscience, New York, 2000.
3. Chang, T. K., R. J. Langley, and E. A. Parker, "Active frequency selective surfaces," *IEE Proc.-Microw. Antennas Propag.*, Vol. 143, No. 1, 62–66, February 1996.
4. Alpaslan, A. and P. Edenhofer, "Electronically tunable frequency selective surfaces," *Proceedings of International Symp. On Electromagnetic Theory*, Vol. 1, 130–132, Thessaloniki, Greece, May 1998.
5. Mias, C., "Demonstration of wideband tuning of frequency-selective surfaces in a waveguide setup," *Microwave and Optical Technology Lett.*, Vol. 44, No. 5, 412–416, March 2005.

Shimming Permanent Magnet of MRI Scanner

Liming Hong and Donglin Zu

IHIP, Physical School, Peking University
Beijing 100871, China

Abstract— In this paper an approach of ferromagnetic shimming for permanent MRI magnet is presented. It is designed to reduce unhomogeneity of magnetostatic field of permanent magnet to meet the stringent requirement for MRI applications. An optimal configuration of ferromagnetic pieces is generated through calculation according to the initial field map and the demanded final homogeneity specifications. This approach uses a minimization technique that makes the sum of squared magnetic moment minimum to restrict the amount of the ferromagnetic material used and the maximal thickness of shim pieces stacked at each position on the shimming boards. Simulation results verify that the method is effective and efficient.

As well known there are two kinds of MRI [1] magnets: closed cylindrical superconductive(SC) magnet and open biplanar-pole permanent magnet such as C-shaped one. The open space of the C-magnet helps the patient overcome any feelings of claustrophobia that may be experienced in a closed designed magnet. Both superconductive and permanent magnet must be shimmed to reduce the unhomogeneity of the magnetic field in the working magnetic field volume to within a predetermined specification, i. e., within a few parts per million for use in medical diagnosis. However, to the permanent magnet due to approximation in the design as well as magnetizing and fabricating tolerances, the final homogeneity of main magnetic field is often far away from the acceptable level. Therefore, shim technology is of major importance in the design and manufacture of permanent MRI magnet. Many papers [2–4] and patents [5, 6] published focus on passive steel shimming of the SC magnet. As for active coil shimming of the SC magnet, likewise there are many papers [7, 8] published. Recently target field method [9–11] has been used for active shimming of SC magnet.

Conventional electromagnetic shimming follows the approach based on representing the field as a spherical harmonics series:

$$B_z(r,\,\theta,\,\phi)=\sum_{n=0}^{\infty}\sum_{m=0}^{n}r^nP_n^m(\cos\theta)\left(A_n^m\cos m\phi+B_n^m\sin m\phi\right)$$

Amplitudes of harmonic components $(A_n^m,\,B_n^m)$ are calculated from the magnetic field measured in or around imaging region. To the MRI permanent magnet [12, 13] including spaced-apart first and second pole faces, a common active current shimming technique is the use of biplanar correction coils [14, 15]. These biplanar coils consist of windings placed on parallel planes, and the magnetic field of interest is created in the space between them. These coils are designed to produce corresponding spherical harmonics [7] and have the merit of not interfering with each other. Generally speaking, active current shimming is used for corrections after a patient's access. Recent progress is that Forbes et al. [16, 17] apply target-field method to design the biplanar shim coils. An alternative is the passive shimming technique [18, 19] using passive ferromagnetic shim configuration to produce different spherical harmonics, the complexity of which increases with harmonic order. In addition, the magnetic moment of a passive iron piece is uni-orientation, its magnitude depends on the strength of the local magnetic field unless saturated, and the passive iron or steel shims have magnetic coupling between them and interfere with each other. Despite the passive ferromagnetic shim approach gives rise to technological difficulties, especially compensating high ordered harmonics, passive shimming is still the preferred method due to its advantages that no power is required and passive ferromagnetic shims are less expensive than active current shimming coils.

Three-dimensional and two dimensional finite element analysis (FEA) techniques have also been used for preliminary mechanical shimming [20]. However, the application of FEA has encountered its bottle-neck because its precision is inadequate to meet the stringent technical demand in MRI. Though mechanical shimming has the merit that requires no extra power and implementation cost is low, it has generally been empirical.

The aim of this paper is to extend Dorri and Vermilyea's method [3, 5] from cylindrical SC magnet to biplanar-pole permanent one and substituting passive ferromagnetic shims with active ferromagnetic shims. However, the active ferromagnetic shimming method presented here is a sort of linear mapping that directly links magnetic field to the sources without corresponding to spherical harmonics. The reason why active ferromagnetic material not iron is used in this kind of shimming is because active ferromagnetic shim have invariable magnetic dipole moment.

1. METHOD

Typical permanent magnet with a pair of opposing parallel shimming boards attached to each pole face. Holes containable of shims are evenly positioned on the shimming boards as shown in Fig. 1(a) and Fig. 1(c). These active ferromagnetic shims with magnetization are placed in holes on the shimming boards and the magnetic field of interest is created in the space between them. This construction provides the ease of fabrication and assembly of shims with required thickness. The shimming process begins by measuring the magnetic flux density at all the sampling points over the imaging region, and then the distribution of ΔB, which is the deviation of the flux density from its expected value, obtained by measurement can be achieved. The object of shimming is to find a reasonable placement of active shims on the shimming boards so that ΔB is eliminated, thereby to better reduce three-dimensional magnetic field unhomogeneity.

In this method the thickness and orientation of the shim pieces on the shimming boards is the design variable to be determined. The magnetic moment of the shim pieces at each position may be positive or negative. Therefore a good algorithm and an optimization computer code are needed.

2. THEORY

According to Maxwell electromagnetic theory, the magnetic induction field generated by a point magnetic dipole is

$$\vec{B} = \frac{\mu_0}{4\pi}\frac{3\left(\vec{\mu}\cdot\vec{r}\right)\vec{r} - \vec{\mu}r^2}{r^5} \tag{1}$$

here $\vec{\mu}$ the magnetic moment of a magnetic dipole and μ_0 the vacuum magnetic permeability. In the case of active ferromagnetic shimming, the ferromagnetic shims are circular or square. The shims piled at each position on the shimming boards can be deemed as magnetic dipoles orientating to $\pm z$-axis while the shimming boards lie at $z' = \pm z_0$ respectively. In practice the shim pieces should be inserted into the predetermined holes of shim boards. We only need to concern the B_z-component within the central sphere region generated by these dipoles because $\vec{B}$ field is determined to be homogeneous if one of its three components is homogeneous. Suppose $\mathbf{r}_i$ represents the coordinate of the ith measured field point; $\vec{r}'_j$ of the jth source point. The magnetic induction field at $\vec{r}_i$ generated by the jth shim dipole can be written as

$$B_z^j\left(\vec{r}_i\right) = \frac{\mu_0\mu_j}{4\pi}\frac{2\left(Z_i - Z'_0\right)^2 - \left(X_i - X'_j\right)^2 - \left(Y_i - Y'_j\right)^2}{\left|\vec{r}_i - \vec{r}'_j\right|^5} \tag{2}$$

We let $\vec{M}$ express the m dimensional serial of magnetic dipoles, $\vec{B}$ the n dimensional serial of compensation induction field. The linear relationship is expressed as

$$\vec{B} = \vec{\vec{A}}\vec{M} \tag{3}$$

Here $\vec{\vec{A}}$ is the $n \times m$ dimensional transformation matrix. This problem can be solved directly by inversing the matrix $\vec{\vec{A}}$ if n equals to m. However, $\vec{M}$ calculated in this way is distinctly definite and may not be controlled. In order to overcome this difficulty, we developed a method that minimizes the sum of squared magnetic moment. The dimension m should be greater than n in order to give space for confinement.

The confinement can be written in matrix form

$$\text{MINMIZE } F = \vec{M}^{\mathrm{T}}\vec{M} \tag{4}$$

Using Lagrange's method of undetermined multipliers, target function is constructed as

$$G = \vec{M}^{\mathrm{T}}\vec{M} - 2\vec{\lambda}^{\mathrm{T}}\left(\vec{\vec{A}}\vec{M} - \vec{B}\right) \tag{5}$$

where $\vec{\lambda}$ representing Lagrange's multiplier vector. Applying $\dfrac{\partial G}{\partial M_i} = 0$ one can get

$$\vec{M} = \vec{\vec{A}}^{\mathrm{T}} \vec{\lambda} \tag{6}$$

Inserting vector $\vec{M}$ into original field equation, one can derive

$$\vec{B} = \vec{\vec{A}} \vec{M} = \vec{\vec{A}} \vec{\vec{A}}^{\mathrm{T}} \vec{\lambda} \tag{7}$$

Here $\vec{\vec{A}}\vec{\vec{A}}^{\mathrm{T}}$ is an $n \times n$ dimensional squared matrix. Assuming the existence of its inverse matrix, vector $\vec{\lambda}$ can be determined as

$$\vec{\lambda} = \left(\vec{\vec{A}} \vec{\vec{A}}^{\mathrm{T}} \right)^{-1} \vec{B} \tag{8}$$

The optimized magnetic dipole moment vector is calculated by the equation

$$\vec{M} = \vec{\vec{A}}^{\mathrm{T}} \vec{\lambda} = \vec{\vec{A}}^{\mathrm{T}} \left(\vec{\vec{A}} \vec{\vec{A}}^{\mathrm{T}} \right)^{-1} \vec{B} = \vec{\vec{A}}_v \vec{B} \tag{9}$$

After that the optimal thickness of the jth dipole consists of shim pieces can be determined if the remanence of the ferromagnetic material is known:

$$t_j = \frac{\mu_0}{B_r S} M_j \tag{10}$$

here B_r represents the remanence and S the sectional area of the shim pieces.

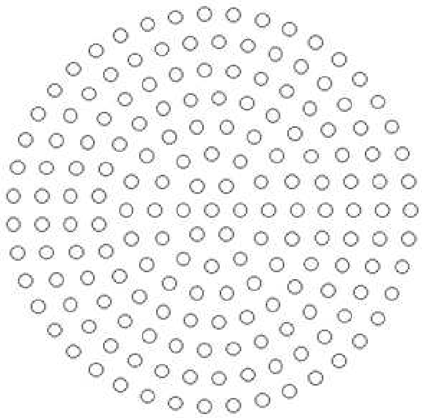
Table 1: Parameters of the Helmholtz system

radius of main coils (m)	0.5
distance between coils (m)	0.5
superconducting current supposed (A)	10000
central magnetic field (Gs)	179.8353
radius of shimming boards (m)	0.3
distance between up & low boards (m)	0.5
radius of shim pieces (cm)	1
radius of target imaging region (m)	0.10
configuration of sampling	$3 \times 3 \times 3$
lattice constant of sampling points (m)	0.08
number of sampling points n	27
unhomogeneity before shimming (ppm)	1763.498*
unhomogeneity after shimming (ppm)	117.469*
total number of shim positions m (case I)	346
maximal thickness of shim pieces (mm)	5.1
total number of shim positions m (case II)	558
maximal thickness of shim pieces (mm)	3.6

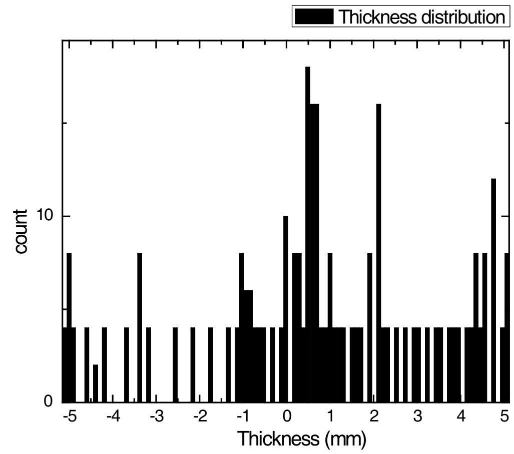
*observed on z-axis

3. RESULTS

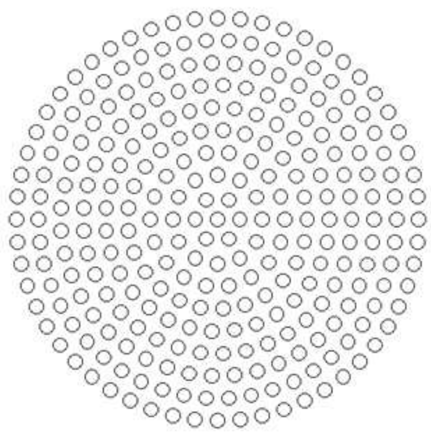
A numerical inspection is presented to test the approach. Consider the quasi-homogeneous magnetic field generated by Helmholtz coil pair instead of the permanent magnet for the sake that its magnetic field can be analytically determined. A series of field points are chosen and the numerical values of magnetic field are used as input to the shimming method. Then the homogeneity of compensated field is investigated to demonstrate the effectiveness of the method. Table 1 shows geometric parameters, associate parameters of investigated Helmholtz coil system, and corresponding results. Measured field points are chosen as equidistant cubic lattices. For comparison of the dependency relationship of the maximal shim thickness to the number of shimming positions, we investigate two cases. Fig. 1(a) shows the configuration of shim pieces on upper shimming board and that of lower



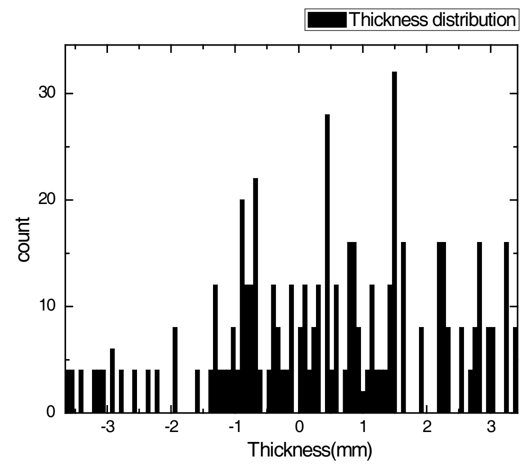
(a)



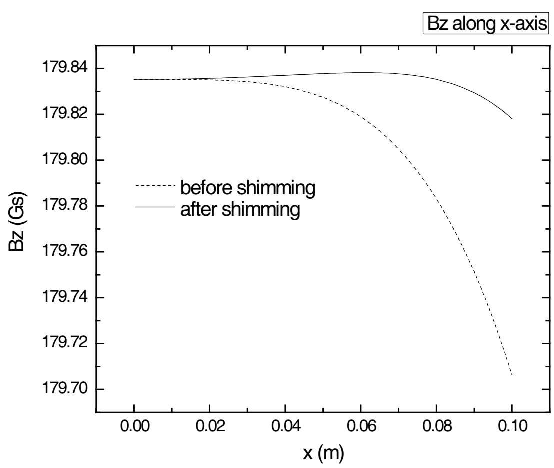
(b)



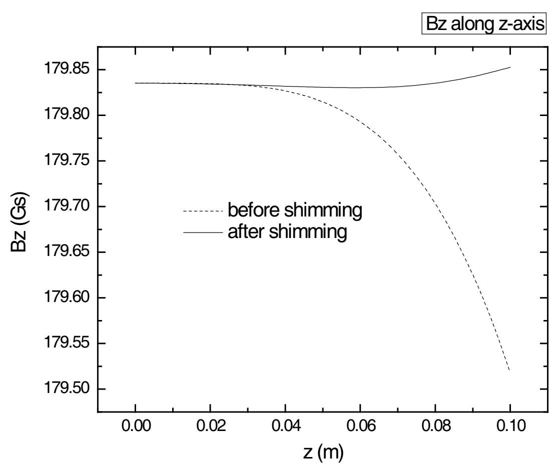
(c)



(d)



(e)



(f)

Figure 1: Improve the field homogeneity from 1763 ppm to 117 for the larder region; from 235 ppm to 15 ppm for the smaller region. (a) The configuration of the shim pieces positioned on each shimming plate. Notice the total number is 346, (b) The statistical thickness distribution of magnetic dipoles positioned at the two shimming plates. Notice the maximal thickness is 5.1 mm, (c) The configuration of the shim pieces positioned on each shimming plate. Notice the total number is 558, (d) The statistical thickness distribution of magnetic dipoles positioned at the two shimming plates. Notice the maximal thickness is 3.6 mm, (e) Field distribution along x-axis before (dashed) and after (solid) shimming, (f) Field distribution along z-axis before (dashed) and after (solid) shimming.

board is identical. Notice that in case I, there are totally 346 positions on the two shimming boards. Among them only 4 positions have the maximal thickness of 5.1 mm. Fig. 1(b) shows the statistics of thickness distribution for those determined dipoles that are used in the shimming and a negative thickness means it is set up inversely. In case II, there are totally 558 positions on the two shimming boards (see Fig. 1(c)) among which only 4 positions have the maximal thickness of 3.6 mm as is shown in Fig. 1(d). In both cases no apex exists at any positions on the two shimming boards, and the field homogeneity is improved over one order of magnitude from 1763 ppm to 117ppm as is indicated in the last two rows in Table 1. Figs. 1(e) and 1(f) show field distributions before and after shimming for the larger region (20 cm diameter sphere region). Corresponding to the smaller region (13.6 diameter sphere region) where the homogeneity is improved from 235 ppm to 15 ppm.

For the same coil system, same target homogeneity requirement, the maximal shim- thickness will decrease while the number of shim-positions increases (see Table 1 and Fig. 1). Because the gap space between opposing parallel pole faces is expensive, increasing the number of shim-positions on the shimming boards resulting in decreasing the thickness of the shimming boards is of practical significance.

4. CONCLUSION

A complete analytical designing methodology for active ferromagnetic shimming of permanent MRI magnet has been formulated and presented. This approach is adaptive to not only the initial rough shimming procedure but the final refined one too as long as sufficient locations are available in both shimming boards. The simulation results prove that this technique is feasible without a spatial harmonics consideration. Moreover the method is proved to be effective and efficient. It will significantly curtail shimming phase of MRI apparatus. The method presented is currently tryout in MRI magnet manufactory.

ACKNOWLEDGMENT

The authors are willing to thank Prof. Pingchou Xia for many helpful discussions.

REFERENCES

1. Zu, D. L., *Magnetic Resonance Imaging (in Chinese)*, Higher Education Press, Beijing, 2004.
2. Hoult, D. I. and D. Lee, "Shimming a superconducting nuclear magnetic resonance imaging magnet with steel," *Ref. Sci. Instrum.*, Vol. 56, No. 1, 131–135, 1985.
3. Dorri, B. and M. E. Vermilyea, "Passive shimming of MR magnets: algorithm, hardware, and results," *IEEE Trans. Appl. Superconductivity*, Vol. 3, No. 1, 254–257, 1993.
4. Belov, A., V. Bushuev, M. Emelianov, et al., "Passive shimming of the superconducting magnet for MRI," *IEEE Trans. Applied Superconductivity*, Vol. 5, No. 2, 679–681, 1995.
5. Vermilyea, M. E., "Method of passively shimming MR magnets," US Patent 4771224, 1988.
6. Dorri, B., "Method for passively shimming a magnet," US Patent 5677854, 1997.
7. Hoult, D. I. and R. Deslauriers, "Accurate shim-coil design and magnet-field profiling by a power-minimization-matrix method," *J. Magn. Reson. Ser. A*, Vol. 108, 9–20, 1994.
8. Forbes, L. K. and S. Crozier, "Asymmetric zonal shim coils for magnetic resonance applications," *Med. Phys.*, Vol. 28, 1644–51, 2001.
9. Forbes, L. K. and S. Crozier, "A novel target-field method for finite-length magnetic resonance shim coils: Part 2. Tesseral shims," *J. Phys. D, Appl. Phys.*, Vol. 35, 839–849, 2002.
10. Forbes, L. K. and S. Crozier, "A novel target-field method for magnetic resonance shim coils: Part 3. Shielded zonal and tesseral coils," *J. Phys. D, Appl. Phys.*, Vol. 36, 68–80, 2003.
11. Brideson, M. A., L. K. Forbes, and S. Crozier, "Determining complicated winding patterns for shim coils using streamfunctions and the targetfield method," *Concepts Magn. Reson.*, Vol. 14, 9–18, 2002.
12. Abele, M., *Structures of Permanent Magnets*, John Wiley and Sons, New York, 1993.
13. Xia, P., *Structures of Permanent Magnets (in Chinese)*, BGD Press, Beijing, 2000.
14. Anderson, W. A., "Electrical current shims for correcting magnetic field," *Rev. Sci. Instru.*, Vol. 32, No. 3, 241–250, 1961.
15. Coley, M. J. E., "Field homogenizing coils for nuclear spin resonance instrumentation," *Rev. Sci. Instru.*, Vol. 29, No. 4, 313–315, 1958.

16. Forbes, L. K., M. A. Brideson, and S. Crozier, "A target-field method to design circular biplanar coils for asymmetric shim and gradient fields," *IEEE Trans. Magn.*, Vol. 41, No. 6, 2134–44, 2005.
17. Forbes, L. K. and S. Crozier, "Novel target-field method for designing shielded biplanar shim and gradient coils," *IEEE Trans. Magn.*, Vol. 40, No. 4, 1929–1938, 2004.
18. Battocletti, J. H., H. A. Kamal, T. J. Myers, and T. A. Knox, "Systematic passive shimming of a permanent magnet for P-31 NMR spectroscopy of bone mineral," *IEEE Trans. Magn.*, Vol. 29, No. 3, 2139–2151, 1993.
19. Zhang, Y. L., D. X. Xie, and P. C. Xia, "Passive shimming and shape optimization of an unconventional permanent magnet for MRI," *Proc. Sixth ICEMS*, Vol. 2, 899–902, Beijing, China, Nov. 9–11, 2003.
20. Battocletti, J. H. and T. A. Knox, "A permanent magnet for wholebody NMR imaging," *IEEE Trans. Magn.*, Vol. 21, No. 5, 1874–1876, 1985.

Research on Target-field Method for Designing Gradient Coil in Permanent-magnet MRI System

Wentao Liu and Donglin Zu
Institute of Heavy Ion Physics, and Beijing Key Laboratory of Medical Physics and Engineering
Physics School, Peking University, Beijing 100871, China

Abstract— Based on the target-field method, an ameliorated approach is presented in this paper for designing gradient coils that can be used in biplanar pole system of permanent-magnet MRI. To restrict the current on the coil plane within a finite radius, the current density is pre-expanded with a suitable series of orthogonal basis, setting the target field points, solved out the current density by inverse matrix calculation, and then the winding patterns can be gained from dispersing the continuous distribution of current density by using the stream-function method. Applying the method, the computer emulation design of some transverse gradient coils and self-shield coils result in the optimized parameters, and the effect of the approach has been illustrated.

1. INTRODUCTION

In 1986, to design gradient coils, Turner devised the target-field method [1], which has become the mainstream method of MRI gradient coil design nowadays. The target-field method is applicable in designing not only the cylindrical gradient coils of superconducting (SC) MRI [2–5] system, but also the planar gradient coils of permanent-magnet MRI [6–9] system. It is clear that the design of gradient coils is a difficult task, since the mathematical problem is often extremely ill-posed. In fact, for coils of finite size, it is the case that there is no unique winding patter that will give the desired target field, and many different choices are possible. To restrict the distribution of current on the coil plane within a finite area, the current density is pre-expanded with a suitable series of orthogonal basis, then setting the target field points, solved out the current density by inverse matrix calculation. As for what orthogonal basis to be chosen, that depends on whether it is convenient or not. Trigonometric functions are the simplest ones. Liu and Truwit [8] use Cartesian coordinates to restrict the currents within a square area. Morrone [9] uses polar coordinates to restrict the currents within a roundness area. However, the expression he referred for calculating magnetic induction B_{z} is not correct. In this paper, adopting Morrone's trigonometric functions in the polar coordinates as basis functions, we educe the correct expression of B_{z} straightforward from BiotSavart Law. Then setting the target field values of B_{z} in the imaging region as the process, solving out the current density by inverse matrix calculation, finally, dispersing the continuous distribution of current density by the stream-function method [10], and the winding pattern comes out. Using this approach, we can design the primary gradient coils as well as the shield coils.

2. THEORY

Give a certain restriction to the distribution of current density based on the permissibility of factual magnet. For instance, to restrict the current between radius ρ_0 and ρ_m, the form of current density can be expressed as:

$$\begin{cases} J_\rho = \sum\limits_{\mathrm{q}=1}^{\mathrm{Q}} U_{\mathrm{q}} \frac{1}{\rho} \sin[qc(\rho-\rho_0)] \sin\varphi \\ J_\varphi = \sum\limits_{\mathrm{q}=1}^{\mathrm{Q}} U_{\mathrm{q}} qc \cos[qc(\rho-\rho_0)] \cos\varphi \end{cases} \tag{1}$$

In the expression, $c = \frac{\pi}{\rho_m - \rho_0}$. Obviously, when $\rho = \rho_0$ and $\rho = \rho_m$, then $J_\rho = 0$, it assures that the current cannot flow out of the region $\rho_0 < \rho < \rho_m$. Setting current density of expression (1) as known condition, we can get the expression of the distribution of space magnetic induction B, and only its components along the axis z are interested in the design research. It is not difficult to write out the analytic expression of B_{z}. Using target-field method, setting target-field point in the imaging region, solved the coefficient U_{q} in formula (1), and the current density can be obtained.

The relationship between values of magnetic induction of the target-field point and U_{q} is:

$$B_{\mathrm{z}} = \sum_{\mathrm{q}=1}^{\mathrm{Q}} U_{\mathrm{q}} D_{\mathrm{q}} \tag{2}$$

In the expression, D_{q} is a double surface integral of space location. For the certain design demands, setting the corresponding space location and target value of target-field points, we can get a matrix equation:

$$\begin{bmatrix} B_1 \\ B_2 \\ B_3 \\ \vdots \\ B_{\mathrm{Q}} \end{bmatrix} = \begin{bmatrix} D_{11} & D_{12} & D_{13} & \dots & D_{1\mathrm{Q}} \\ D_{21} & D_{22} & D_{23} & \dots & D_{2\mathrm{Q}} \\ D_{31} & D_{32} & D_{33} & \dots & D_{3\mathrm{Q}} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ D_{\mathrm{Q}1} & D_{\mathrm{Q}2} & D_{\mathrm{Q}3} & \dots & D_{\mathrm{QQ}} \end{bmatrix} \begin{bmatrix} U_1 \\ U_2 \\ U_3 \\ \vdots \\ U_{\mathrm{Q}} \end{bmatrix} \tag{3}$$

Solve the equation for U_{q}, put it back to expression (1), and get J_ρ, J_φ. Then get the map of conductor wires of the gradient coils from dispersing the current density by using the stream-function method. To deal with the eddy current problem, we can add shield coils at the outside of gradient coils, to counteract the magnetic field out of coil space, reducing the flux into the iron pole of the magnet, and making the eddy current as small as possible. The process of design and calculation for the shield coils are similar to the gradient coils.

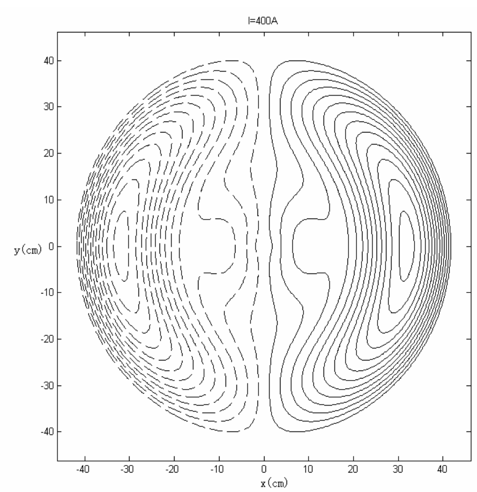


Figure 1: The pattern of conducting wire of the gradient coil computed using Equations (1)–(3).The maximal radius of the coil is 0.42 m, the center aperture 0.01 m, distance between the gradient coil and the center plane is 0.23 m. The imaging region as the linear magnetic field is $-0.2\,\mathrm{m} < y$, $x < 0.2\,\mathrm{m}$, $-0.15\,\mathrm{m} < z < 0.15\,\mathrm{m}$. In this region the gradient strength $G_x = 25\,\mathrm{mT/m}$, current 400 A not account for any mirror current.

3. RESULTS

Based on the equations listed above, we implement some computer simulations by using MATLAB. Two results of computer simulations are given here, respectively illustrating the designs of unshielded transverse gradient coils and shielded transverse gradient coils. The map of conductor wires of unshielded transverse gradient coil is shown in Fig. 1; the dashed line indicates the reverse current. The magnetic field caused by the coil is linear in the imaging region: $-20\,\mathrm{cm} < x < 20\,\mathrm{cm}$, increasing along with the axis x, as shown in Fig. 2.

The self-shield gradient coil consists of a primary coil and a shield coil. The parameter of the primary coil is the same as the gradient coil mentioned above. The distance between the shield coil and the center plane ($z = 0$) is 0.29 m. The shielded region is the space: $|z| \geq 29.5\,\mathrm{cm}$. Figs. 3(a) and (b) respectively show the wire maps of the primary coil and the shield coil. Fig. 4 shows the magnetic field in the imaging region and outside space generated by the self-shield gradient coil. When the current is 500 A, in the imaging region the gradient field is $G_x = 25\,\mathrm{mT/m}$. The slant dashed is ideal gradient; the upright dashed indicate the boundary of imaging region, and the 'horizontal' dashed show that the magnetic field outside the shield coil near to zero. It illustrates the self-shield coil system's effect of shielding.

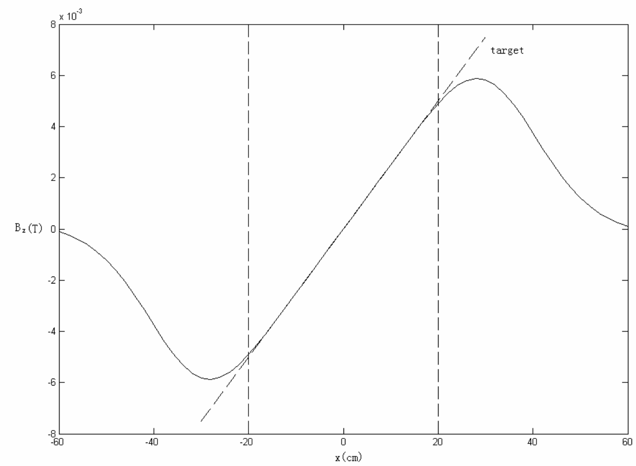
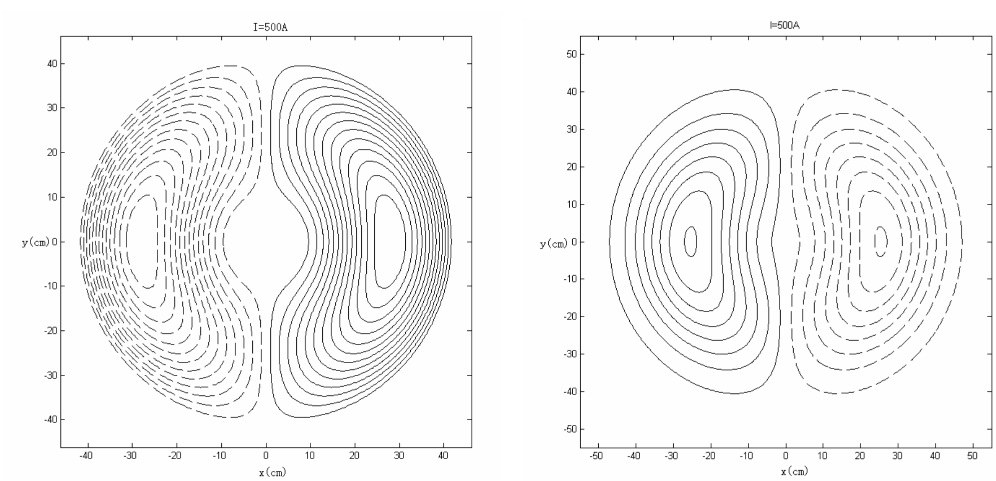


Figure 2: The magnetic field generated by gradient coil shown in Fig. 1. At plane $z = 0$, in the area $-20\,\text{cm} < x < 20\,\text{cm}$, The field is linearly vary with coordinate x. The slant dashed is ideal gradient field. The two upright dasheds indicate the boundary of target imaging region.



(a) (b)

Figure 3: Eddy current self-shield gradient coil system. (a) primary coil; (b) shield coil.

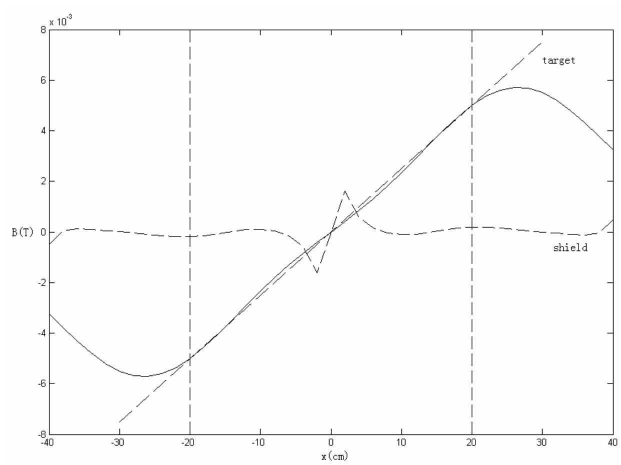


Figure 4: The gradient field produced by self-shield gradient coil shown in Fig. 3. The slant dashed line indicates the ideal gradient field along the x-axis, the upright dashed lines indicate the boundary of imaging region. The 'horizontal' dashed line shows the magnetic field d at $z = \pm 29.5\,\text{cm}$.

4. CONCLUSION

In original target-field method the mathematical problem is often extremely ill-posed. Whether the gradient coil or the self-shielded gradient coil designed here, the current density satisfies the condition of producing the desired linear field B_z over the target region $-20\,\text{cm} < x,\ y < +20\,\text{cm}$, $-15\,\text{cm} < z < +15\,\text{cm}$ and could be implemented into a practical winding pattern for a transversal gradient coil. It confirms that this kind of new target-field method is able to solve for a coil of genuinely finite size. The shape of the plane coil can be conveniently controlled according to the design demands. While the eddy current problem can be solved to some extent by designing self-shielded gradient coil. It is evident that this approach is robust, convenient and successful. In a word, this paper provides a suit of simple design approach and program for gradient coils of permanent-magnet MRI system. In addition, the method is also applicable in the design of the zonal gradient coils and the shimming coils.

REFERENCES

1. Turner, R., "A target field approach to optimal coil design," *J. Phys. D: Appl. Phys.*, Vol. 19, 147–151, 1986.
2. Turner, R., "Minimum inductance coils," *J. Phys. E: Sci. Instrum.*, Vol. 21, 948–952, 1988.
3. Turner, R., "Gradient coil design: a review of methods," *Magnetic Resonance Imaging*, Vol. 11, 903–920, 1993.
4. Blaine, A., et al., "Constrained length minimum inductance gradient coil design," *Magnetic Resonance in Medicine*, Vol. 39, 270, 1998.
5. Tang, X., D. Zu, and S. Bao, "A new design method for asymmetrical head gradient coils used for superconducting MRI scanner," *Progress in Natural Science*, Vol. 14, No. 9, 753–757, 2004.
6. Forbes, L. K., M. A. Brideson, and S. Crozier, "A target-field method to design circular biplanar coils for asymmetric shim and gradient fields," *IEEE Trans. Magn.*, Vol. 41, No. 6, 2134–44, 2005.
7. Forbes, L. K. and C. Stuart, "Novel target-field method for designing shielded biplanar shim and gradient coils," *IEEE Trans. Magn.*, Vol. 40, No. 4, 1929–1938, 2004.
8. Liu, H. and C. L. Truwit, "True energy-minimal and finite-size biplanar gradient coil design for MRI," *IEEE Trans. Med. Imag.*, Vol. 17, No. 5, 826–830, 1998.
9. Morrone, T., "Optimized gradient coils and shim coils for magnetic resonance scanning systems," United States Patent 5,760,582, Jul. 23, 1992.
10. Tomasi, D., "Stream function optimization for gradient coil design," *Magn. Reson. Med.*, Vol. 45, 505, 2001.

A New Target Field Method for Optimizing Longitudinal Gradient Coils' Property

Feng Qi[1], **Xin Tang**[2], **Zhe Jin**[1], **Le Wang**[2], **Donglin Zu**[2], **and Weimin Wang**[1]

[1]Key Laboratory for Quantum Information and Measurements, Ministry of Education
School of Electronics Engineering & Computer Science, Peking University
Beijing 100871, China
[2]Institute of Heavy Ion Physics, and Beijing Key Laboratory of Medical Physics and Engineering
Physical School, Peking University
Beijing 100871, China

Abstract— A new approach has been presented in the paper to optimize the longitudinal gradient coil performance. First, traditional spherical harmonic target field method is deduced, the relation between the magnetic field and the current distribution is described by a matrix equation. Then, simulated annealing method is introduced to the optimization procedure, and those high order coefficients which are used to vanish become the variables designed for optimization. Finally, stream function method is used to transform the current density into discrete gradient coils. Comparison between traditional method and the optimized method shows that the inhomogeneity in the region of interest can be reduced from 13.59% to 4.9%, the coil efficiency is increased from 11.56 mT/m/A to 18.07 mT/m/A, and the minimum distance of the discrete current wire is raised from 0.95 mm to 2.02 mm.

In MRI (Magnetic Resonance Imaging) system, the gradient field are used to encode the NMR returned signals from patients. It is very important to design a high-quality gradient coil. Nowadays, there are mainly two main methods have been used in the design of gradient coils: TF (target field) method and SA (simulated annealing) method. TF method is a widely used method for designing gradient coils, which was first put forward by Turner [1, 2] in 1986. In this method, the desired magnetic fields are first specified in the DSV region, and then the current density is evaluated by inverse Fourier transformation. Finally, stream functions [3] are used to change the current density into discrete wires. Terry uses spherical harmonics to represent the target field [4] and provides us a very convenient way to design biplanar gradient and shim coils. However, there are some deficiencies in traditional TF method. One particular drawback is that the current distribution is continuous current density, which can only be approximated by a discrete distribution, and the discrete process may bring some errors to the ideal field. In addition, longitudinal gradient coils designed by traditional TF method tend to concentrate to the periphery of the pole-piece, which increases the inductance and makes gradient coils extremely hard to built. Finally, it is usually a difficult job for TF method to optimize all the coil performance simultaneously. Another way for gradient coil design is stochastic optimization. One of the best tools is SA (Simulated Annealing) method which was first published by Metropolis in 1953 and further developed by Kirkpatrick and co-workers. In 1993, Crozier applied the algorithm to the design of gradient coils [5, 6]. The advantage of this method is that we can make an overall control of coil performance by defining a suitable energy function E. However, an obvious shortage is that it needs a large amount of time to make the optimization procedure to converge to the optimal solution. Here we follow Morreone's target field method for the design of transverse gradient coils, a new approach have been introduced to optimize the current geometry with simulated annealing method. A high quality gradient coil has been built using the new optimization method.

1. METHODS

1.1. Target Field Method

The divergence of magnetic induction $\mathbf{B}(\mathbf{r})$ vanish everywhere, $\nabla \cdot \mathbf{B} = 0$. $\mathbf{B}$ can be written by taking the curl of the vector potential $\mathbf{A}(\mathbf{r})$ in the form (1)–(2). $\mathbf{A}(\mathbf{r})$ is defined to describe the magnetic field at the observation point $\mathbf{r}$, which is generated by the current density $\mathbf{J}(\mathbf{r}')$. The vacuum permeability $\boldsymbol{\mu}_0 = 4\pi \cdot 10^{-7}\,\mathrm{N/A^2}$.

$$\mathbf{A}(\mathbf{r}) = \frac{\mu_0}{4\pi}\int \frac{\mathbf{j}(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}d\tau' \tag{1}$$

$$\mathbf{B}(\mathbf{r}) = \nabla\times\mathbf{A}(\mathbf{r}) = \frac{\mu_0}{4\pi}\int \nabla\frac{1}{|\mathbf{r}-\mathbf{r}'|}\times\mathbf{j}(\mathbf{r}')\,d\tau' \tag{2}$$

In biplanar gradient coil design, current distribution is constrained to the pole pieces, and steady state magnetic phenomena are characterized by no change in the net charge density anywhere. Then we can easily get $\nabla\cdot\mathbf{J(r')} = 0$ and $\mathbf{J_z}(\mathbf{r}') = 0$. So current density can be expressed in terms of the curl of the stream function $\mathbf{S}$ (4–6), and stream function $\mathbf{S}$ is defined by the form of series expansion (3).

$$\mathbf{S} = S\mathbf{e_z} = -\sum_{q=1}^{Q} U_q \sin\left[qc\left(\rho'-\rho_{\min}\right)\right]\mathbf{e_z} \tag{3}$$

$$J'_\rho = \frac{1}{\rho'}\frac{\partial S}{\partial\varphi'} = 0 \tag{4}$$

$$J'_\varphi = -\frac{\partial S}{\partial\rho'} = \sum_{q=1}^{Q} U_q qc\cos\left[qc\left(\rho'-\rho_{\min}\right)\right] \tag{5}$$

$$c = 2\pi/\left(\rho_{\max}-\rho_{\min}\right) \tag{6}$$

where $\rho_{\max}$ and $\rho_{\min}$ are the maximum and minimum winding radius, and variable Q indicates that there are totally Q terms of coefficient U_q. Then we will return to (2) and substitute $|\mathbf{r}-\mathbf{r}'|^{-1}$ by the spherical harmonic expansion.

$$\begin{aligned}\nabla\frac{1}{|\mathbf{r}-\mathbf{r}'|} &= \sum_{n=0}^{\infty}\sum_{m=-n}^{n}\frac{4\pi}{2n+1}\nabla\frac{r^n}{r'^{n+1}}Y_{nm}(\theta,\varphi)\,Y^*_{nm}\left(\theta',\varphi'\right)\\ &= -\mathrm{Re}\sum_{n=0}^{\infty}\sum_{m=0}^{n}\frac{2(n-m)!}{\delta_m(n+m)!}r^n\mathrm{P}_\mathrm{n}^\mathrm{m}(\cos\theta)\,\mathrm{e}^{im\varphi}\nabla'\left(\frac{\mathrm{e}^{-im\varphi'}}{r'^{(n+1)}}\mathrm{P}_\mathrm{n}^\mathrm{m}\left(\cos\theta'\right)\right)\end{aligned} \tag{7}$$

Since we put the original point O in the center of the main magnet, as is shown in Figure 1, the distance of observation point $\mathbf{r}$ is smaller than the distance of source point $\mathbf{r}'$. And let $x = \cos\theta$, we can rewrite Legendre function $P_\mathrm{n}^\mathrm{m}(x)$ by means of Rodrigues's formula. Thus, equation (7) is obtained. Longitudinal gradient coil are required to provide a field $\boldsymbol{B}$ such that the z component B_z is in direct proportion to z. And since longitudinal gradient coils are axial symmetric, we can easily find that coefficient $m = 0$. For terms $m \neq 0$, expression $r^n P_\mathrm{n}^\mathrm{m}(x)$ should be canceled in the process of magnetic field integration.

After expand the gradient operator of (7) in cylindrical coordinates, z component of the the magnetic induction intensity $\boldsymbol{B}$ is given by the equation (8)

$$\begin{aligned}B_z^a(r,\theta,\varphi) &= \frac{\mu_0}{2}\sum_{n=0}^{\infty}r^nP_n(\cos\theta)\int\Bigg\{\sum_{q=1}^{Q}U_q\rho' qc\cos\left[qc\left(\rho'-\rho_{\min}\right)\right]\\ &\quad\times\left[\frac{-(n+1)}{r'^{(n+2)}}P_n\left(\cos\theta'\right)\sin\theta' + \frac{1}{r'^{(n+2)}}\frac{\mathrm{d}P_n\left(\cos\theta'\right)}{\mathrm{d}\theta'}\cos\theta'\right]\Bigg\}\mathrm{d}\rho'\end{aligned} \tag{8}$$

For biplanar gradient coils, there are two same current sheets fixed at a distance of $\pm a$ from the origin O. Because of the inherent symmetry associated with the gradient fields, it is convenient to consider symmetric current distribution $\mathbf{J}_{z=a} = -\mathbf{J}_{z=-a}$. Therefore, we can write z-component of the magnetic field as

$$\begin{aligned}B_z(r,\theta,\varphi) &= B_z^{a+}+B_z^{a-} = \mu_0\sum_{n=1}^{\infty}r^{2n-1}P_{2n-1}(\cos\theta)\int\Bigg\{\sum_{q=1}^{Q}U_q\rho' qc\cos\left[qc\left(\rho'-\rho_{\min}\right)\right]\\ &\quad\times\left[\frac{-2n}{r'^{(2n+1)}}P_{2n-1}\left(\cos\theta'\right)\sin\theta' + \frac{1}{r'^{(2n+1)}}\frac{\mathrm{d}P_{2n-1}\left(\cos\theta'\right)}{\mathrm{d}\theta'}\cos\theta'\right]\Bigg\}\mathrm{d}\rho'\end{aligned} \tag{9}$$

According to coefficient n and m, we can rewrite equation (9) into the the form of matrix equation (10). C_n is the weighting factor of term $r^{2n-1}P_{2n-1}^{m}(x)$. Coefficient A_{qn} can be calculated from equation (9).

$$\begin{bmatrix} A_{11} & \cdots & A_{1Q} \\ \vdots & \ddots & \vdots \\ A_{N1} & \cdots & A_{NQ} \end{bmatrix} \begin{bmatrix} U_1 \\ \vdots \\ U_Q \end{bmatrix} = \begin{bmatrix} C_1 \\ \vdots \\ C_N \end{bmatrix} \tag{10}$$

So, if coefficient C_1, C_2, ..., C_N is given, coefficient U_1, U_2, ..., U_Q is obtained by equation (10). Then, according to stream function theory, we can transform current density distribution into discrete wires. And the current intensity carried by each wire is written as

$$I_0 = (S_{\max} - S_{\min})/N_0 \tag{11}$$

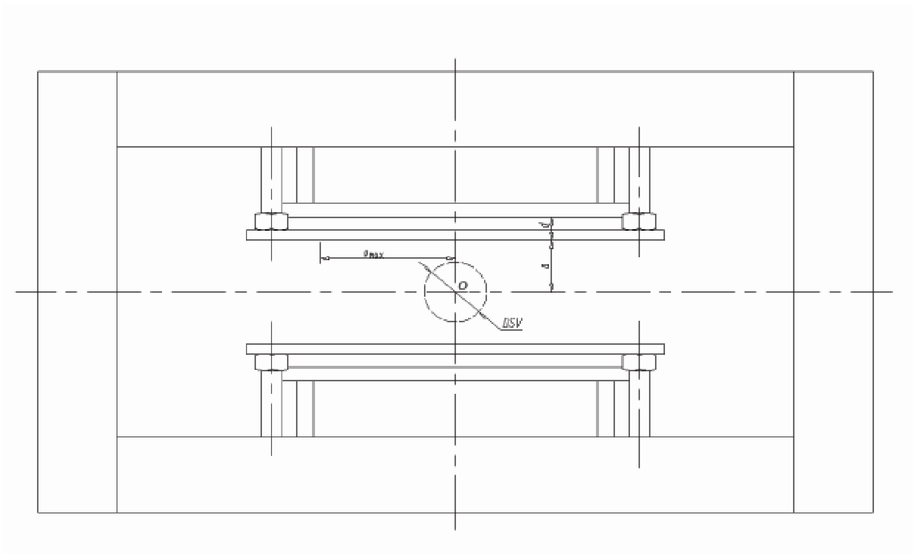


Figure 1: Main magnet structure of the micro magnetic resonance imaging system.

1.2. Simulated Annealing Method

SA method is a global optimization technique, the basis of which is rooted from Monte Carlo iteration. The SA program begins at a high temperature, and the energy function of the system is optimized step by step through the process of annealing. In order to circumvent being trapped in local energy minima, we use Metropolis accept rule which allows the system to make positive energy move with a probability linked to Boltzman statistics. It is very convenient for SA to make an overall control of the gradient-coil properties by defining energy function E,

$$E = k_1 \frac{I_0}{MD} + k_2^{(per-5\%)} \tag{12}$$

$$per = \frac{2\,(grad_{\max} - grad_{\min})}{(grad_{\max} + grad_{\min})} \tag{13}$$

where MD is minimum distance of the discrete gradient coil. per is an index of inhomogeneity. $grad_{\max}$ and $grad_{\min}$ are the maximum and minimum gradient values in DSV (Diameter of spherical volume) region. k_1 and k_2 are weighting factors for efficiency and uniformity.

1.3. Optimization

We found a way to optimize the property of the gradient coil. In traditional target field method, we only consider the term $n = 1$ to design longitudinal gradient coils, and the other coefficients C_2, ..., C_N are used to design high order shim coils. However, take into consideration of the error brought by current density discretization, Coefficients C_2, ..., C_N can be used to reduce the error and improve the coil efficiency. Simulated annealing method has been used to optimize these coefficients, and the energy function can be written as equation (12).

2. RESULTS AND DISCUSSION

Gradient coils of the Micro-MRI (Magnetic Resonance Imaging) system are fixed on two pole pieces of the permanent magnet, as shown in Figure 1. $a = 0.0236$ m, the longitudinal distance between the current plane and the origin O; $d = 0.025$ m, the distance between the current plane and the pole plate; $\rho_{\max} = 0.065$, maximum radius of the gradient coils; $\rho_{\min} = 0$ m, minimum radius of the gradient coils; $N_0 = 12$, coil turns; In SA method, we use *geometric* rule as the annealing schedule rule, $T(i) = T_0 \times 0.95^i$, where $T_0 = 15.3$ is the initial optimization temperature; $M = 50$, maximum sample steps at each temperature; $S = 100$ annealing steps, so the total number of iterations is 5000; Since the existence of image interference, its 4rd-order influence must be taken into account in the model. Other parameters are listed: $k_1 = 10^{-4}$ and $k_2 = 20$, weighting factors for efficiency and homogeneity respectively; $DSV = 0.03$ m, diameter of spherical volume.

Table 1: Coil performance designed by the traditional TF method and the optimized TF method.

C_1	C_2	C_3	C_4	per	Effi(mT/m/A)	MD(mm)
0.06	0	0	0	13.95%	11.56	0.95
0.06	37.50	-6.64×10^4	-2.25×10^6	4.91%	18.07	2.02

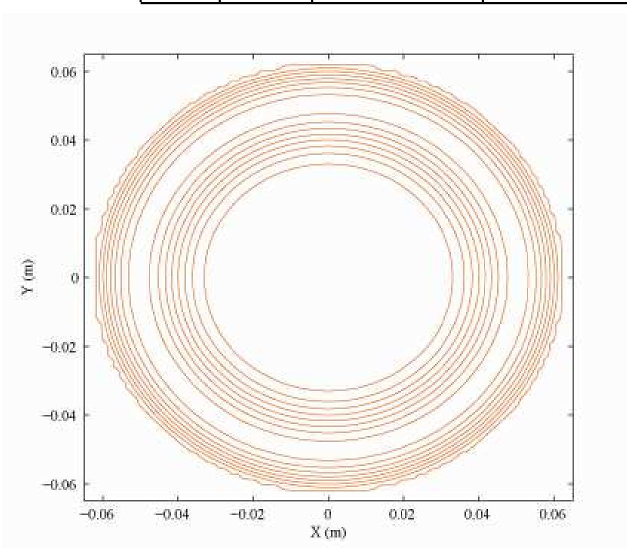


Figure 2: z-gradient coil designed by traditional target field method.

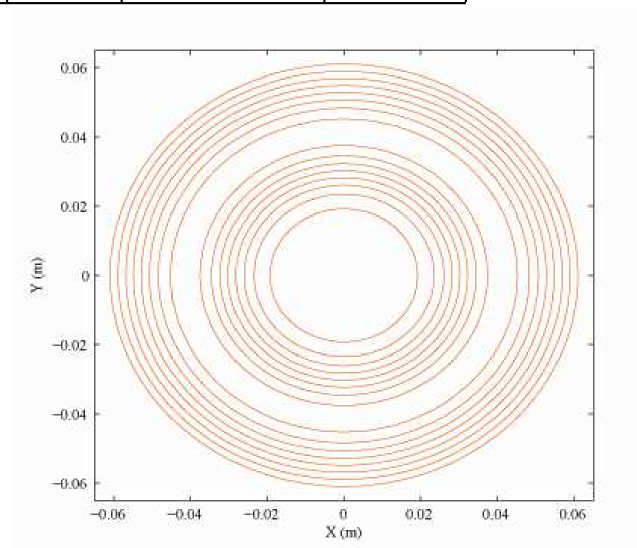


Figure 3: z-gradient coil optimized by simulated annealing method.

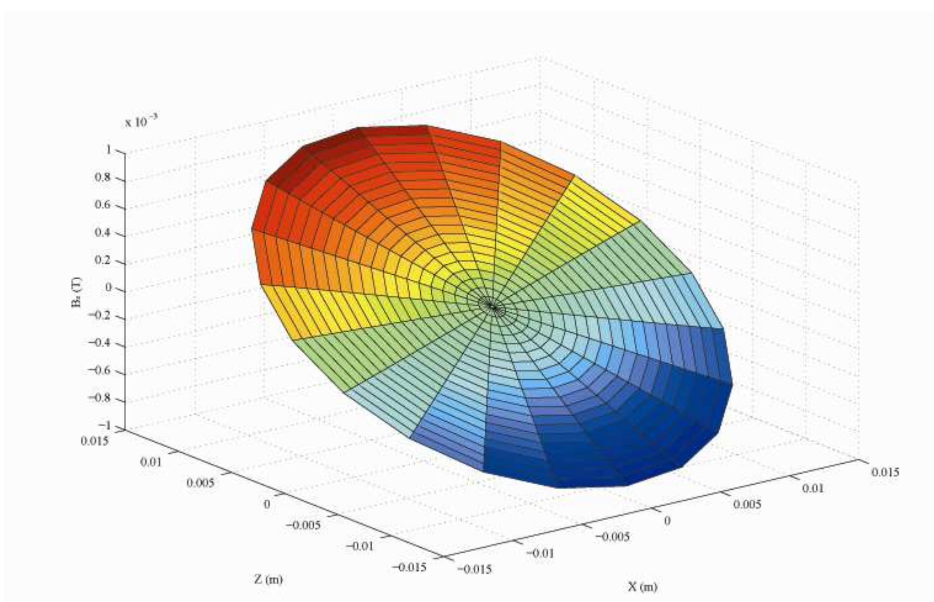


Figure 4: Magnetic field produced by the final coil set.

In traditional target field method only the first term of coefficient C is used to design z-gradient coil, the results have been displayed in the first line of Table 1. However, during the process of optimization, we will use coefficient C_2, C_3, ..., C_N as the variable optimized in SA method. The results have been shown in the second line of Table 1. The inhomogeneity in the DSV region is reduced from 13.59% to 4.91%, the efficiency of the gradient coil is increased from 11.56 mT/m/A to 18.07 mT/m/A, and the minimum distance is allowed to be adjusted from 0.95 mm to 2.02 mm.

Figure 2 is the gradient coil designed by traditional target field method, the current distribution often focus to a location which makes it hard to build. Figure 3 illustrates the final gradient wire distribution optimized by simulated annealing method, and the coil can produce a high quality gradient field, as shown in Figure 4.

3. CONCLUSIONS

A new approach to optimize the longitudinal gradient performance has been proposed in this paper. Those high order coefficients are used as the variables optimized in simulated annealing method. Using the optimization method, the coil efficiency, uniformity and minimum distance have been considerably improved. In view of these test results, it appears that this new method is worthy and convenient to be used in the application of shim and gradient coils design.

ACKNOWLEDGMENT

This research is subsidized by National Key Technologies Research and Development Programme of China (Grant No. 001BA210A03) and Key Project of 2005 Beijing Science and Technology Programme (Grant No. D0205002000021).

REFERENCES

1. Turner, R., "A target field approach to optimal coil design," *J. Phys. D: Appl. Phys.*, L147, 1986.
2. Turner, R., "Turner R. Gradient coil design: A review of methods," *Magn. Reson. Imaging*, Vol. 11, 903–920, 1993.
3. Lemdisaov, R. A. and R. A. Ludwig, "Stream function method for gradient coil design," *Concepts Magn. Reson. Part B*, Vol. 26B, 67–80, 2005.
4. Morrone, T., "Optimization gradient coils and shim coils for magnetic resonance scanning system," U.S. Patent 5760582, 1998.
5. Crozier, S. and D. M. Doddrell, "Gradient coil design by simulated annealing," *J. Magn. Reson.*, Vol. A103, 354–357, 1993.
6. Crozier, S., L. K. Forbes, and D. M. Doddrell, "The design of transverse gradient coil of restricted length by simulated annealing," *J. Magn. Reson.*, Vol. A107, 126–128, 1994.

A New Method for Shimming a Magnetic Field in NMR System

Zhe Jin[1], **Xin Tang**[2], **Feng Qi**[1], **Donglin Zu**[2], **and Weimin Wang**[1]

[1]Key Laboratory for Quantum Information and Measurements, Ministry of Education
School of Electronics Engineering & Computer Science
Peking University, Beijing 100871, China
[2]Institute of Heavy Ion Physics, and Beijing Key Laboratory of Medical Physics and Engineering
Physical School, Peking University, Beijing 100871, China

Abstract— A new approach has been presented in the paper to gain a uniform magnetic field in NMR system. First, the adopted shimming piece is modeled as a magnetic dipole moment to calculate its magnetization effect on the background field within region of interest. Then, sequential quadratic programming method is utilized to determine the ideal solution for shimming work. Finally, the ideal solution is discretized, and quantization error control technique is prepared for special cases. This new method helps to reduce the inhomogeneity in the region of interest from 56.8 ppm to 14.21 ppm, within one hour in practical shimming work.

In NMR (Nuclear Magnetic Resonance) system, it is very important to gain a uniform magnetic field to improve the performance of the whole assembly. A process named shimming is applied in the generation of magnetic field to adjust the homogeneity to an acceptable level. Nowadays, there are two main methods being used in shimming process: active shimming and passive shimming [1, 2]. Passive shimming saves the trouble of applying expensive electrical current source to the system, while its effect on background field is hard to predict. A lot of work has been done to find a solution with global optimization technique, like SA (Simulated Annealing method), GA (genetic algorithm), however, these methods are usually limited in the practical shimming work because of their inefficient, in terms of time or computational source required. Here a new passive shimming process design is put forward to meet the need of creating a controlled B_0 field with minimal inhomogeneity in practical shimming work.

1. METHODS

1.1. Shimming Process

Generally, the whole shimming process consists of three stages. First, the magnetic field is measured in a plurality of locations within region of interest to get an initial B_0 map; second, based on the comparison of target field and B_0 map, a shimming model is developed to form a B' field which substantially cancels the inhomogeneities within initial magnetic field. In this design, cylindrical permanent magnetic pieces of specific sizes are adopted to create the compensating field. A computer program based on SQP (Sequential Quadratic Programming) method is executed to determine an amount of shimming piece required at each target location. The details will be discussed in the following sections; third, quantization error control technique is put forward as an option. If necessary, it will be utilized to get a field with even better homogeneity at the price of adding three more different kinds of permanent magnetic pieces to the shimming system. After the whole shimming process is carried out, a near-homogenous with controlled B_0 and minimal quantization error is achieved by employing least amount of shimming elements at strategic locations.

1.2. Magnetic Dipole Moment Model

As mentioned in Section 1.1, because the size of shimming piece adopted is much smaller than the size of ROI (Region of Interest), each shimming piece is modeled as a magnetic dipole moment in our model to calculate the magnetization effect it brings to the field. And as a consequence of the solenoidal property of the magnetic induction, B can be written by taking the curl of the vector potential $\boldsymbol{A}(\boldsymbol{r})$ in the form (1). $\boldsymbol{A}(\boldsymbol{r})$ is defined to describe the magnetic field at the observation point r, which is generated by the current density $\boldsymbol{J}(\boldsymbol{r'})$. The vacuum permeability $\mu_0{=}4\pi{\cdot}10^{-7}N/A^2$.

$$\boldsymbol{A}(\boldsymbol{r}) = \frac{\mu_0}{4\pi}\int\frac{\boldsymbol{J}(\boldsymbol{r'})dV'}{|\boldsymbol{r}-\boldsymbol{r'}|} \tag{1}$$

$$\boldsymbol{B}(\boldsymbol{r}) = \nabla\times\boldsymbol{A}(\boldsymbol{r}) = \frac{\mu_0}{4\pi}\int\nabla\frac{1}{|\boldsymbol{r}-\boldsymbol{r'}|}\times\boldsymbol{J}(\boldsymbol{r})d\tau' \tag{2}$$

We can write the $\boldsymbol{A(r)}$ in Taylor expansion series in the neighborhood of original points as:

$$\boldsymbol{A(r)} = \frac{\mu_0}{4\pi}\int \boldsymbol{J(r')}\left[\frac{1}{R} - r'\cdot\nabla\frac{1}{R} + \frac{1}{2!}\sum_{i,j} r'_i r'_j \frac{\partial^2}{\partial r_i \partial r_j}\frac{1}{R} + \dots\right] dV' \tag{3}$$

where the first term vanishes because of the continuity of constant current, which may be illustrated in equation (4):

$$\boldsymbol{A}^{(0)}(\boldsymbol{x}) = \frac{\mu_0}{4\pi R}\int \boldsymbol{J(x')}dV' = \oint \boldsymbol{I}dl = 0 \tag{4}$$

Thus, using the second term to approximate $A(r)$, we get:

$$\boldsymbol{A}^{(1)}(\boldsymbol{x}) = -\frac{\mu_0 \boldsymbol{I}}{4\pi}\oint x'\cdot\nabla\frac{1}{R}dl' \tag{5}$$

$$\boldsymbol{B}^{(1)} = \nabla\times A^{(1)} = -\frac{\mu_0}{4\pi}(\mathbf{m}\cdot\nabla)\frac{\mathbf{R}}{\mathbf{R^3}} \tag{6}$$

$$\mathbf{m} = \frac{1}{2}\int \boldsymbol{x'}\times dl' \tag{7}$$

where $\mathbf{m}$ is the magnetic moment, $R = \sqrt{(x^2+y^2+z^2)}$ is distance to original point. In the Cartesian coordinate system, the z component of $\boldsymbol{B}^{(1)}$, which is our concern, can be written as:

$$B_z = \frac{\mu_0 m}{4\pi}\left(\frac{3\left(z-z'\right)^2 - R^2}{R^5}\right) = C\cdot\left(\frac{3\left(z-z'\right)^2 - R^2}{R^5}\right) \tag{8}$$

where the value of constant C is decided by the property of shimming piece.

1.3. Sequential Quadratic Programming Method

SQP method represents the state of art in nonlinear programming methods. Based on the work of Schittkowski, Biggs, and Han [3–5], the method is capable of performing accurately, efficiently over a large number of nonlinear programming problems. In our model, we want to achieve appropriate homogeneity and desired center field by putting the least passive shims at strategic locations, then, SQP method is chosen to minimize a ABS function, which may be defined as:

$$\min\sum_{i=1}^{n}|X_i| \tag{9}$$

The requirements of homogeneity and center field are put into following constraints:

$$-\boldsymbol{\delta}_0 \le \mathbf{B}\boldsymbol{X} \le \boldsymbol{\delta}_0 \tag{10}$$

$$B_{\text{cent}} - \boldsymbol{\delta} \le \mathbf{A}\boldsymbol{X} - \mathbf{B}\boldsymbol{X} \le B_{\text{cent}} + \boldsymbol{\delta} \tag{11}$$

where X_i is a state variable of passive shims, which indicates the amount of passive shims at location i; $\mathbf{B}$ is B_{cent} contribution made by $\boldsymbol{X}$; $\boldsymbol{\delta}_0$ is the maximum value B_{cent} allows to be changed; B_{cent} is the original magnetic flux intensity at the center of ROI (Region of Interest); $\boldsymbol{\delta}$ is the constraint bound vector in terms of field (Gauss, Tesla) or inhomogeneity (ppm); $\mathbf{A}$ is the shim strength matrix of passive shims at each shim location, in terms of field contribution to each shimming points.

1.4. Quantization Error Control (QEC)

SQP method provides us with the solution $\boldsymbol{X}$, which is a vector of real number indicating the amount of shimming element put at predetermined locations. However, in practical shimming work, it is impossible for us to get pieces of all the sizes demanded by $\boldsymbol{X}$. So we have to discrete the ideal solution, usually by replacing the real number in $\boldsymbol{X}$ with nearest integral number. Certain quantization error is added into our final solution because of this approximation, which may be not tolerable in some cases. In our work, specific quantization error control technique is put forward to solve the problem.

The quantization error control technique can be called "plate in plate" technique as well. The main shimming plate has dozens of holes of same size for accepting sub shimming plates. Each sub shimming plate has five holes of different sizes decided by its shimming strength. And then a combination of shimming pieces from a selection of full strength piece, 10% strength piece, 20% strength piece, 50% strength piece, 90% strength piece is placed in sub shimming plate to approximate nominal shimming amount. A computer program or even manual calculation may suffice to select the shimming piece combination. As an example, suppose the shimming piece of plate limit is 10, Table 1 lists available combinations for the convenience of shimming work.

Table 1: Available combinations of shimming pieces for shimming work.

Total strength	Full strength piece	10% strength piece	20% strength piece	50% strength piece	90% strength piece
10	10	0	0	0	0
9.9	9	0	0	0	1
9.8	9	1	1	1	0
9.7	9	0	1	1	0
......	......	......	......	......	......
5.6	5	1	0	1	0
5.5	5	0	0	1	0
5.4	5	0	2	0	0
......	......	......	......	......	......
0.3	0	1	1	0	0
0.2	0	0	1	0	0
0.1	0	1	0	0	0

Omitted portions of table are indicated by "......", which may be easy to fill up. With this method, we are capable of gaining a good approximation by employing only five different kinds of pieces in practical shimming work. It also makes it easy to locate the combination requiring the minimum number of pieces to produce the demanded shimming strength.

2. RESULTS AND DISCUSSION

Shimming plates for Micro-NMR system are fixed on the two pole pieces of the permanent magnet assembly which is shown together with Gauss-meter in Figure 1. The exact ROI shape and its sampling mode are shown in Figure 2, where the color bar indicates the strength of magnetic induction of the original field.

Starting with the original B_0 field, traditional shimming method, the new shimming method without quantization error control technique and the new shimming method with quantization error control technique are applied into practical work. Table 2 shows the difference of the three.

Figure 1.

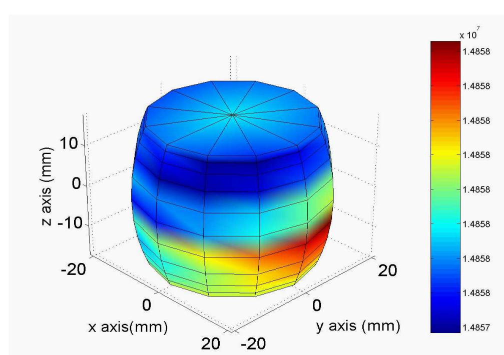


Figure 2.

Table 2: Comparison between the traditional shimming method and the new method.

	Initial inhomogeneity (ppm)	Final inhomogeneity (ppm)	Time Taken (min)	Number of shimming piece	Types of shimming piece
Tradition method	56.8	35	1920 (unpredictable)	145	Unpredictable
New method without QEC	56.8	21.05	60	61	1
New method with QEC	56.8	14.21	80	63.7	5

It is clearly illustrated that the new methods greatly outperform the traditional method in terms of efficiency, accuracy, convenience and predictability. At the price of adding some shimming time and shimming pieces, which is an acceptable cost, the QEC technique improves the field homogeneity from 21.05 ppm to 14.21 ppm.

3. CONCLUSIONS

A new approach to gain a uniform magnetic field has been proposed in this paper. The shimming amount at predetermined locations is decided with sequential quadratic programming method. With this approach, a controlled B_0 magnetic field with minimal inhomogeneity is achieved in a short time. It is proved efficient and effective in application of shimming a magnetic field.

ACKNOWLEDGMENT

This research is subsidized by National Key Technologies Research and Development Programme of China (Grant No. 001BA210A03) and Key Project of 2005 Beijing Science and Technology Programme (Grant No. D0205002000021).

REFERENCES

1. Rimkunas, et al., "Method of shimming a magnetic field," US patent 6714109 B1, 1990.
2. Huang, et al., "Methods and apparatus for passive shimming of magnets," US patent 6778054 B1, 1990.
3. Biggs, M. C., "Constrained minimization using recursive quadratic programming," *Towards Global Optimization* (L. C. W. Dixon and G. P. Szergo, eds.), North-Holland, 341–349, 1975.
4. Han, S. P., "A globally convergent method for nonlinear programming," *J. Optimization Theory and Applications*, Vol. 22, 297, 1977.
5. Schittkowski, K., "NLQPL: A fortran-subroutine solving constrained nonlinear programming problems," *Annals of Operations Research*, Vol. 5, 485–500, 1985.

A New Eddy-current Compensation Method in MRI

C. Ma and X. H. Jiang
Department of Electrical Engineering, Tsinghua University, Beijing 100084, China

Abstract— We present here a new eddy-current compensation method. A mathematical model of eddy-current field is introduced to calculate first order eddy-current distorted non-uniform sampling trajectory in k-space. Based on the computed k-space trajectory, compensated image is reconstructed using the gridding method. In addition, the model of eddy-current field can also be used to calculate zero order eddy-current induced additional phase to correct the phase of MRI signal before gridding reconstruction, as long as such phase correction is able to further improve qualities of reconstructed image. Initial experiment on a 0.3T permanent MRI system demonstrates that the proposed method is able to reconstruct FSE (fast spin-echo) images of similar qualities as those produced by conventional gradient pre-emphasis method.

1. INTRODUCTION

Time-varying gradient fields are essential for signal localization in MRI. Meanwhile, however, eddy-current is induced in conducting structures around the gradient coils according to the *Faraday's law of induction*. In permanent MRI system, eddy-current is more severe than that in superconducting MRI system due to the ferromagnetic pole pieces which are attached on the main magnets to shim the main field. Eddy-current causes significant artifacts and distortions, and may seriously degrade the overall performance of MRI magnet.

The most effective way to reduce eddy-current is to use so called active shielding gradient coils which are designed to have minimum fringe fields. Turner and Bowley [1] first developed a method to design active gradient coils in a cylindrical geometry; Yoda [2] expanded the target field method of Turner from cylindrical to planar geometries. But these hardware design methods can not be 100% effective in practice; gradient pre-emphasis method is widely used to further compensate effect-current [3]. Duyn, Yang, and etc. [4] proposed a different way to correct eddy-current effects by measuring actual k-space trajectories in MRI, but this method is difficult to correct images acquired with conventional sequences, such as SE and FSE sequences.

In this paper, a new eddy-current compensation method is presented. Initial experiment on a 0.3T permanent MRI system demonstrates that the proposed method is able to reconstruct FSE (fast spin-echo) images of similar qualities as those produced by conventional gradient pre-emphasis method.

2. THEORY

Suppose $\bar{B}_e(\bar{r}, t)$ is the eddy-current field induced by gradient field $\bar{G}(t)$, and then the spatial dependence of eddy-current is described by a Taylor expansion:

$$\bar{B}_e(\bar{r}, t) = b_0(t) + \bar{g}(t) \cdot \bar{r} \tag{1}$$

where $b_0(t)$ is the zero order eddy-current which is spatially invariant like the main static field; $\bar{g}(t)$ is the first order eddy current, the spatial dependence of which is similar to gradient field. High order eddy-current field is neglected. Let $\bar{G}(t)$ be the input of a linear system, $b_0(t)$ and $\bar{g}(t)$ be the outputs of the system, and $e_{b0}(t)$ and $e_g(t)$ be the step response function of the system; then $b_0(t)$ and $g(t)$ are given by:

$$b_0(t) = -\frac{\mathrm{d}G(t)}{\mathrm{dt}} \otimes e_{b0}(t),\ \ g(t) = -\frac{\mathrm{d}G(t)}{\mathrm{dt}} \otimes e_g(t) \tag{2}$$

where "$\otimes$" denotes convolution; and $e_{b0}(t)$ and $e_g(t)$ are given by a sum of decaying exponential functions:

$$e_{b0}(t) = \sum_n \alpha_n e^{-t/\tau_n},\ \ e_g(t) = \sum_m \beta_m e^{-t/\xi_m} \tag{3}$$

When eddy-current exists, the MRI signal in 2D case can be written as:

$$S(\overline{k_e}) = \iint \rho(\bar{r}) e^{-j2\pi k_e \cdot r} e^{-j\theta} d^2 r \tag{4}$$

$$\theta = \gamma \int_0^t b_0(\tau) d\tau \tag{5}$$

$$k_e = \frac{\gamma}{2} \left(\int_0^t G(t) \cdot r d\tau + \int_0^t g(t) \cdot r d\tau \right) \tag{6}$$

where $\rho(\bar{r})$ is the number of proton spins per unit volume in two dimensions; $S(\overline{k_e})$ is MRI signal; γ is the gyromagnetic ratio; θ is the additional phase caused by zero order eddy-current; and $\overline{k_e}$ is the first order eddy-current distorted k-space trajectory.

Once given parameters of eddy-current model, θ and $\overline{k_e}$ can be respectively calculated by Eq. (5) and Eq. (6). In the first step of the proposed eddy-current compensation method, based on the computed k-space trajectory $\overline{k_e}$, compensated image is reconstructed using the gridding method which is developed to reconstruct correct images from k-space data of arbitrary k-space trajectories. In addition, zero order eddy-current induced additional phase θ can be used to correct the phase of MRI signal before gridding reconstruction, as long as such phase correction is able to further improve qualities of reconstructed image.

In the process of gridding reconstruction, the MRI signal $S(\overline{k_e})$ with non-uniform sampling is first convoluted with a gridding kernel $C(\bar{k})$, and then is resampled to a uniform grid to produce resampled data $S'(\bar{k})$; secondly, a 2D inverse FFT reconstructs intermediate image $I'(\bar{r})$ from $S'(\bar{k})$; finally, correct image $I(\bar{r})$ is generated by applying devolution in space domain to $I'(\bar{r})$. Calculation process described above is expressed by Eq. (7) to Eq. (9):

$$S'(\overline{k_n}) = \sum_k S(\overline{k_{e,j}}) C(\overline{k_n} - \overline{k_e}) w(\overline{k_{e,j}}) \tag{7}$$

$$I'(\bar{r}) = FFT^{-1}\left(S'(\overline{k_n})\right) \tag{8}$$

$$I(\bar{r}) = I'(\bar{r})/c(\bar{r}) \tag{9}$$

where $\overline{k_n}$ denotes the coordinate of resampling grid in k-space; $\overline{k_{e,j}}$ denotes the coordinate of sampling data; $w(\overline{k_{e,j}})$ is sampling density weighting factor; $c(\bar{r})$ is the Fourier pair of gridding kernel $C(\bar{k})$.

3. METHOD

3.1. Eddy-current Measurement

The pulse sequence to measure eddy-current field is given in Figure 1. A small phantom is placed in different places in the imaging volume. The time derivation of the phase of the FID signal indicates the decay of the eddy-current field. A group of FID signals with different delay time are detected to measure the decay in a time range of 0∼1 s. Eddy-current field decay curves at a group of points are collected; then curve-fitting is used to determine the parameters of eddy-current model.

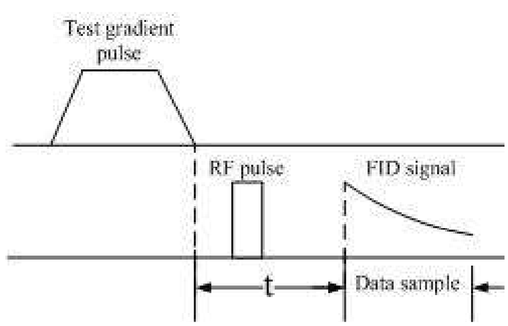


Figure 1: Eddy-current measurement pulse sequence diagram.

The eddy-current measurement experiment was performed on a 0.3T vertical permanent magnet interfaced to an MR Research Systems console (MR Research Systems, UK). The small phantom used in experiments was a column, 16 mm in diameter and 25 mm in length, which was full of specific solution with 3.6 g/L NaCl and 1.25 g/L $CuSO_4$. Duration time of gradient pulse was 200 ms; ramp time of gradient pulse was 0.5 ms; and measurement time range was 0∼1000 ms. For each direction of gradient field, measurement was performed at 6–7 points in that direction with a separation of 40 mm. The gradient pre-emphasis device of gradient amplifier in the channel of test gradient was disabled during the measurement.

3.2. θ and k_e Computation

A computation program coded in the C language was developed to calculate zero order eddy-current induced additional phase θ and the first order eddy-current distorted k-space trajectory $\overline{k_e}$. Besides computed parameters of eddy-current model, waveforms, amplitudes and timing diagrams of a given MRI sequence were input into the program in the form of parameter files; eddy-current field induced by each gradient lobe of the sequence was calculated by Eq. (2) and (3); then θ and $\overline{k_e}$ were calculated by Eq. (5) and (6). In addition, in the initial program, RF pulse was assumed ideally homogeneous, and the function of 180° refocusing RF pulse was simply assumed to be changing the sign of accumulated phase of MRI signal. It typically takes 15 seconds to compute the θ and $\overline{k_e}$ of a given sequence.

3.3. Gridding Reconstruction Parameters

The initial method used 2X grid to reduce aliased artifacts. Since the first order eddy-current distorted k-space trajectories has no specific geometrical shape, areas of Voronoi diagram which were first recommended by Rasche et al., [5] were used as sampling density weighting factors. The Voronoi diagram area computation program was coded by Matlab script language using integrated routine of Matlab to generating Voronoi diagram of given k-space trajectory.

Two often recommended gridding kernels were compared. The first one, recommended by Beatty et al. [6], is a Kaiser-Bessel kernel with parameters as follows: width = 4 and $\beta = 18.6389$; the second one, recommended by Sarty et al. [7], is a Gaussian kernel with parameters as follows: width = 5 and $b = 0.5993$. Experiment reported that the two kernel reconstructed images with trivial difference, so the former one was chosen for smaller width of gridding kernel and thus less computation time.

3.4. Simulation and MRI Experiment

The simulated images were computed by a 2D MRI simulator, which was developed to simulate eddy-current effects on MRI images by integrating eddy-current model into to the common Bloch equation [8]. Computed eddy-current parameters from measurement were used in the simulation program. A proton-density weighting image was selected as a virtual object for simulation. The image was downloaded from the website: http://www.bic.mni.mcgill.ca/brainweb/, which is an online interface to a 3D MRI simulated brain database [9]. Parameters of the downloaded image were slice thickness = 5 mm, noise (calculated relative to the brightest tissue) = 3%, intensity non-uniformity ("RF") = 40%. The sequence for simulation was an FSE sequence. Key parameters of the sequence were echo-train-length = 4, 256 × 384 matrix size, slice direction = y direction, frequency-encoding direction = x direction, phase encoding direction = z direction;

The imaging experiment was performed on the same MRI system as that used in eddy-current measurement. Images were acquired by an FSE sequence with the same parameters as those in simulation. Some extra acquiring parameters were TE = 35 ms, TR = 2000 ms, center echo

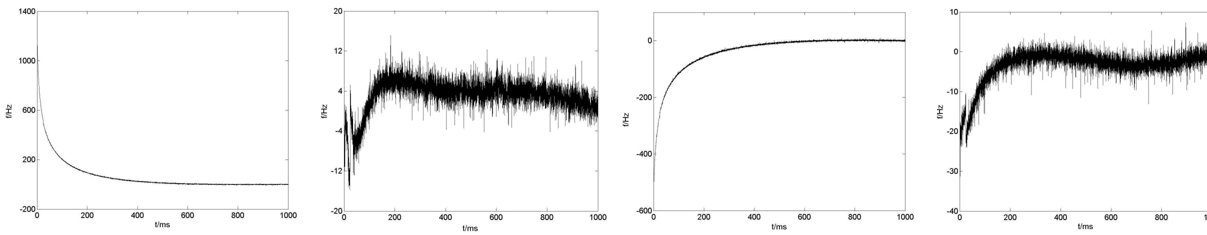
Table 1: Eddy-current parameters.

x direction		y direction		z direction	
τ/ms	α	τ/ms	α	τ/ms	α
245.43	−0.0414	75.40	0.0556	15.33	0.0358
59.44	−0.0319			2.35	0.0204
8.83	−0.0259				
ξ/ms	β	ξ/ms	β	ξ/ms	β
206.14	0.0203	249.70	0.0213	162.43	0.0225
49.74	0.0149	54.80	0.0209	38.76	0.0135
12.74	0.0203	12.81	0.0179	11.94	0.0167
2.18	0.0079	2.89	0.0083	2.10	0.0079

index = 2, number of average = 2, FOV = 250×500 mm, slice thickness = 10 mm.

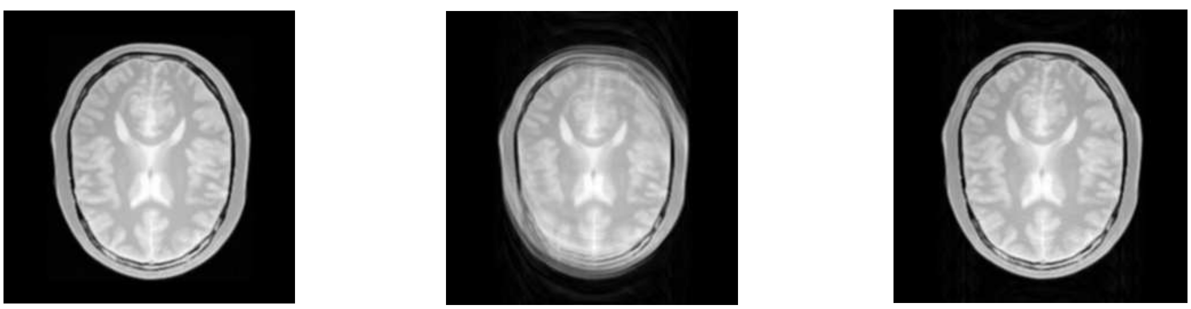
4. RESULTS

Computed parameters of eddy-current field are shown in Table 1. In addition, Figure 2 shows measured eddy-current field curves and the differences between measured curves and the model predicted curves at points of $z = 100$ mm and $z = -60$ mm, demonstrating that the eddy-current model with computed parameters describes the real eddy-current field in good precision.



(a) (b) (c) (d)

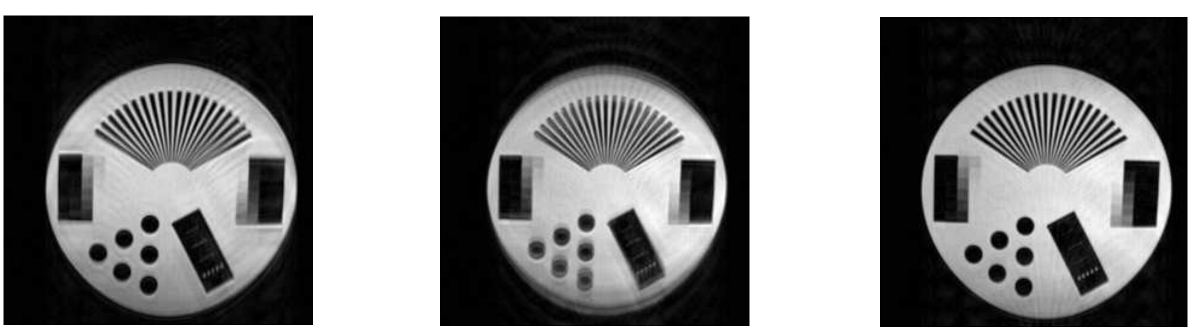
Figure 2: Eddy-current field curves from measurement and the differences between measured curves and the model predicted curves: (a) measured eddy-current field curve at $z = 100$ mm; (b) the difference curve at $z = 100$ mm; (c) measured curve at $z = -60$ mm; (d) the difference curve at $z = -60$ mm.



(a) (b) (c)

Figure 3: Performance of proposed eddy-current compensation method on simulated MRI images. (a) Simulated image when no eddy-current exists. (b) Simulated image when eddy-current exists reconstructed by conventional FFT. (c) Simulated image when eddy-current exists reconstructed by eddy-current compensation method in this paper.

Figure 3 illustrates the performance of proposed eddy-current compensation method on simulated MRI images. Simulation results are tailored into a matrix of 256×256 for display. Simulation



(a) (b) (c)

Figure 4: Performance of proposed eddy-current compensation method on experiment phantom images. (a) Phantom image when gradient pre-emphasis device was activated. (b) Phantom image reconstructed by conventional FFT when gradient pre-emphasis devices were disabled. (c) Phantom image reconstructed by eddy-current compensation method in this paper when gradient pre-emphasis devices were disabled.

results in Figure 3 demonstrate that the new eddy-current compensation method effectively reduces the artifacts and distortion caused by eddy-current.

Experiment results in Figure 4 further strengthen the effectiveness of the proposed method. The image (c) reconstructed by proposed method when gradient pre-emphasis devices were disabled has similar qualities as the image (a) produced by conventional method.

5. CONCLUSION AND DISCUSSION

We presented a new eddy-current compensation method. Initial experiments demonstrate that the presented method is able to reconstruct the images obtained when gradient pre-emphasis devices were disabled of similar, if not slightly better, qualities than those produced by conventional gradient pre-emphasis method.

The major advantage of the presented method is that no extra gradient pre-emphasis device is needed, and thus could get rid of the time consuming pre-emphasis adjustment work and reduce costs. On the other hand, the major disadvantage of this method is that extra computation time is required. Despite the time of computing areas of Voronoi diagrams which can be pre-computed and stored, the presented method requires $O(mN\log mN + mNw)$ arithmetic operations versed by $O(N\log N)$ of conventional reconstruction, where N is the total number of pixels of the image; m is the overgridding factor of gridding reconstruction; and w is the width of the convolution kernel. However with the development of high performance computation technology, the extra computation is becoming to be tolerable.

A possible merit of the presented method is that optimization process can be integrated into the method to self-adjust parameters of the eddy-current model which would make real-time eddy-current compensation possible. The major problems will be how to design appropriate image quality evaluation function, and how to implement the optimization process fast.

REFERENCES

1. Turner, R., "Gradient coil design: a review of methods," *Magnetic Resonance Imaging*, Vol. 11, 903–920, 1993.
2. Yoda, K., "Analytical design method of self-shielded planar coils," *Journal of Applied Physics*, Vol. 67, No. 9, 4349–4353, 1990.
3. Bernstein, M., et al., *Handbook of MRI Pulse Sequence*, Elsevier Press, New York, 2004.
4. Duyn, J., et al., "Simple correction method for k-space trajectory deviations in MRI," *Journal of Magnetic Resonance*, Vol. 132, No. 53, 1998.
5. Rasche, V., et al., "Resampling of Data Between Arbitrary Grids Using Convolution Interpolation," *IEEE Trans. Med. Imag.*, Vol. 18, No. 5, 385–392, 1999.
6. Beatty, P., et al., "Rapid Gridding Reconsttruction With a Minimal Oversampling Ratio," *IEEE Trans. Med. Imag.*, Vol. 24, No. 6, 799–808, 2005.
7. Sarty, G., et al., "Direct reconstruction of non-cartesian k-space data using a nonuniform fast fourier transform," *Mag. Reson. Med.*, Vol. 45, 908–915, 2001.
8. Ma, C., et al., "The computer simulation of eddy currents effects in MRI," (in Chinese), *Journal of Tsinghua University Science and Technology*, (accepted).
9. Cocosco, C., et al., "BrainWeb: online interface to a 3D MRI simulated brain database," *NeuroImage*, Vol. 5, No. 4, Part 2/4, 1997.

Calculating Efficient Noise Resistance of RF Coils for Low Field MRI Systems

Y. Li and X. H. Jiang
Department of Electronical Engineering, Tsinghua University, Beijing 100871, China

Abstract— An analytical method to calculate efficient noise resistance of RF coils in low field MRI systems is presented. Coil self-resistance and sample resistance are evaluated respectively. Using this method, efficient noise at 12.72 MHz of solenoid coil and square-shaped surface coil are calculated. The results show that the coil self-resistance and sample resistance of surface coils are comparable. Thus, both parts of efficient resistance should be calculated when optimizing signal-to-noise ratio to design geometries of RF coils.

1. INTRODUCTION

The problems of studying efficient noise resistance of RF coils arise at the development and evaluation of RF coils in MRI systems [1]. In general, the efficient noise resistance consists of coil self-resistance, sample resistance and dielectric loss in sample [2–4]. Since dielectric loss can be reduced by using distributed capacitance and coil self-resistance may be neglected at high resonance frequency, noise from coils is contributed only from sample resistance in high field MRI [4, 5]. However, dominant noise from coils is different in low field MRI [1, 4, 6]. This paper considers both self- and sample resistance of RF coils in low field MRI systems. The dominant noises of solenoid coil and surface coil are studied as examples. Moreover, an optimized thickness of conductor is found to reduce coil self-resistance.

2. ANALYTICAL METHODS

Neglecting propagation and dielectric loss within the sample, the efficient noise resistance of coils includes coil self-resistance and sample resistance, which is associated with induced eddy currents in sample. Coil self-resistance is evaluated by current density in conductor at resonance frequency. The current density is given by

$$J = -\sigma \cdot \left(\frac{\partial A}{\partial t} + \nabla \cdot \varphi \right) \tag{1}$$

In the quasi-static approximation, the current density $\dot{J}$ can be calculated by the surface integral of the conductor

$$\dot{J} = -j\sigma\mu f \iint\limits_S \dot{J} \ln r ds - \sigma \frac{\partial \varphi}{\partial z} \tag{2}$$

where σ and μ are the conductivity and the permeability of the conductor respectively, f is the resonance frequency, r respects to the distance between ds, S is the whole region of $\dot{J}$, and z is current direction. This integral equation is solved using numerical method which changes integral equations to matrix equations [7]. The cross section of conductor is discretized on a $M \times N$ lattice. According to the works by J. Li et al. [7], the current density in centre region of conductor decreases not so rapidly as that of edge region of conductor. We formed the lattice by a non-uniform rectilinear grid with longer steps in the centre region to reduce matrix scale. Assuming the current density of each cell on grid is constant, we obtain the following $M \times N + 1$ linear equations

$$\begin{gathered} \dot{J}_i + j\sigma\mu f \iint\limits_S \dot{J}_i \ln r ds + \dot{J}_0 = 0 \\ \sum_{i=1}^{M \times N} s_i \dot{J}_i = \dot{I} \end{gathered} \tag{3}$$

where $\dot{J}_i$ and s_i are current density and area of the ith cell respectively, $\dot{J}_0$ is a constant in this problem and $\dot{I}$ is the current in conductor. After the discretization of conductor and calculation of the surface integral, the integral equation (2) can be represented in the matrix-vector form as follows

$$[A][J] = [I] \tag{4}$$

where A is coefficient matrix, $J = [J_1,\, J_2,\, \ldots J_i,\, \ldots J_{M\times N},\, J_0]$ is current density vector and $I = [0,\, 0,\, \ldots 0,\, \dot{I}]$. We find the solutions of current density of conductor in the form

$$[J] = [A]^{-1}[I] \tag{5}$$

The self-resistance R_a of unit length conductor can be express as

$$R_a = \frac{\sum_{i=1}^{M\times N} s_i \left|\dot{J}_i\right|^2}{\sigma I^2} \tag{6}$$

Sample resistance is expressed by the power absorbed by the sample for a unit current in the coil, calculated from the induced eddy-current

$$R_{sample} = P = \sigma \int_V |E(r)|^2 dv \tag{7}$$

where $E(r)$ is the electric field generated by a unit current flowing in coil, σ is the sample conductivity, and the integral extends over whole volume of sample [4, 8].

3. RESULTS

Volume coil and surface coils are wildly used in low field MRI systems [5]. A MATLAB program is employed to calculate the efficient noise resistance of a two-circle solenoid coil with 24.8 cm diameter and 9.2 cm length, as well as square-shaped surface coils at 12.72 MHz as examples. The conductor is copper with 0.18 mm thickness and 12 mm width, whose geometries are showed in Fig. 1. The cross section of conductor is discretized on a 150-by-10 grid. The lattice is formed by a non-uniform rectilinear grid in width direction with a space step of 0.002 mm in region A and 0.392 mm in region B. In thickness direction, a constant space step of 0.018 mm is chosen while the penetration depth is 0.023 mm. The sample is a 22 cm diameter and 26 cm length cylinder ($\sigma = 0.8$ S/m and $\mu = \mu_0$ [4]) for solenoid coil, as well as a 40 cm length, 48 cm width and 28 cm thickness cube for surface coil.

Using method described above, the self-resistance and sample resistance of solenoid coil are 0.177 Ω and 1.362 Ω respectively, which means sample resistance of volume coils is dominant. The contour of cross section current density is shown in Fig. 2. The result shows current density decreases rapidly in the edge region (about 5-fold of penetration depth) related to skin effect and using long space step in centre region reduces matrix scale and calculation time.

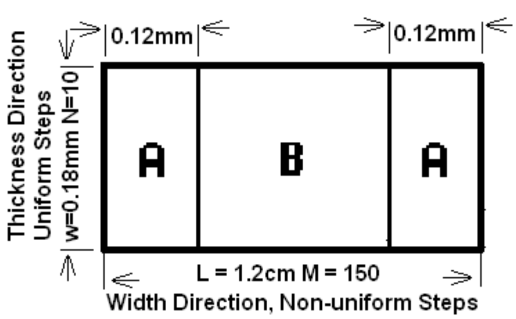


Figure 1: Cross section geometries.

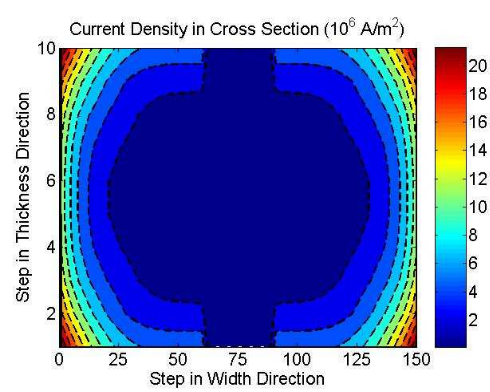


Figure 2: Contours of current density.

The self-resistance and sample resistance of square-shaped surface coils with length from 20 cm to 80 cm are calculated. Self-resistance of coils increases in proportion to coil length while sample resistance increases more rapidly. Thus the dominant noise depends on coil geometry. The sample resistance to self-resistance ratio of each coil is shown in Fig. 3(a). The curve in Fig. 3(a) shows the sample resistance exceeds self-resistance when length of coil is over 28 cm. The normalized signal-to-noise ratio (SNR) of each coil at the centre of sample is also calculated, as shown in Fig. 3(b). When we obtain the max SNR with a 52 cm length square-shaped coil, the corresponding sample resistance is 3-fold to self-resistance. Compared with volume coils such as solenoid coils, self-resistance of surface coil is much close to sample resistance.

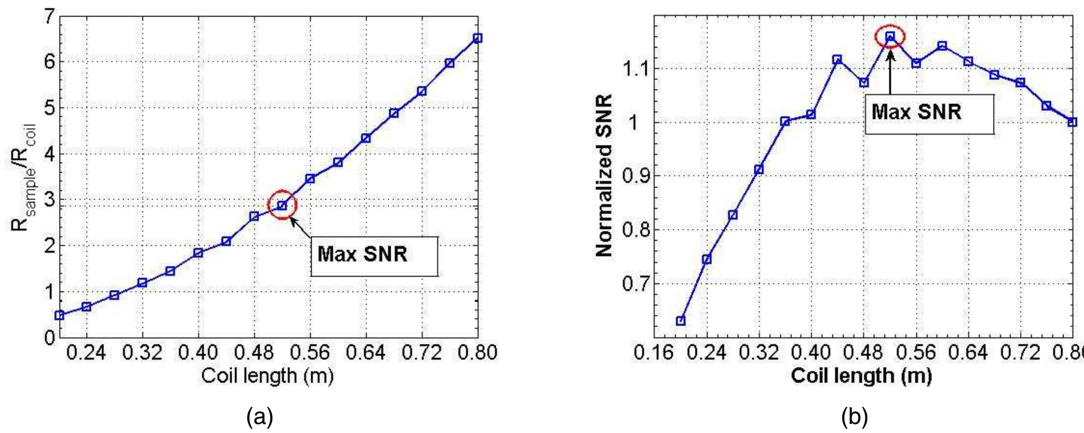


Figure 3: Sample resistance to coil resistance ratio (a) and normalized SNR at the centre of sample (b) versus coil length.

It is well known that a general method to reduce self-resistance is increasing conductor thickness. However, due to the skin effect the self-resistance is not exactly in inverse proportion to conductor thickness. Fig. 4 gives the normalized resistance response of conductor thickness from penetration depth to 8-fold of penetration depth at 12.72 MHz. The curve shows self-resistance decreases slowly when conductor is thicker than 4-fold of penetration depth, which means the proper thickness in this situation is approximately 0.09 mm.

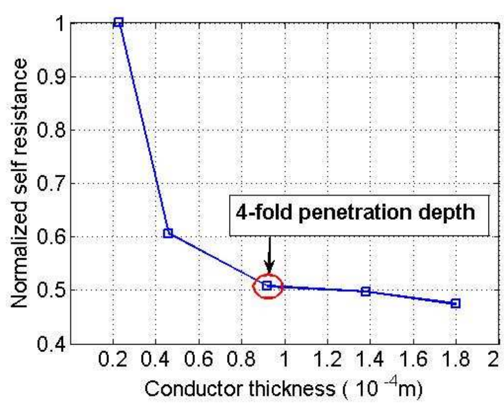


Figure 4: Normalized self-resistance of copper with 1.2 cm width at 12.72 MHz versus conductor thickness.

4. CONCLUSION

In this paper, an approach to calculate self-resistance and sample resistance of RF coils is introduced. The calculation has proved that the dominant noise in low field MRI systems depends on coil structure. For volume coils, sample resistance is approximately 10-fold greater than self-resistance while self-resistance and sample resistance are comparable for surface coils. It means self-resistance should be well considered when we optimize SNR for RF surface coils design. In addition, the effect of increasing conductor thickness is limited by skin effect to reduce self-resistance. We has presented the proper conductor thickness is approximately 4-fold of penetration depth.

REFERENCES

1. Darrasse, L. and J. C. Ginefri, "Perspectives with cryogenic RF probes in biomedical MRI," *Biochimie*, Vol. 85, 915–937, 2003.
2. Hoult, D. I. and R. E. Richards, "The signal-to-noise ratio of the nuclear magnetic resonance experiment," *Journal of Magnetic Resonance*, Vol. 24, 71–85, 1976.
3. Wang, J., A. Reykowski, and J. Dickas, "Calculation of the signal-to-noise ratio for simple surface coils and arrays of coils," *IEEE Transaction on Biomedical Engineering*, Vol. 42, No. 9, 908–917, 1995.
4. Chu, X. and X. Jiang, "SNR-optimized approach for the design of RF surface coils for permanent MRI systems," *Journal of Tsinghua University (Sci & Tech)*, Vol. 45, No. 3, 351–354, 2005.
5. Zhang, H., X. Song, S. Bao, and D. Zu, "MRI radiofrequency coil technology," *China Journal of Medical Imaging Technology*, Vol. 21, No. 9, 1440–1441, 2005.
6. Redpath, T. W., "Signal-to-noise ratio in MRI," *The British Journal of Radiology*, Vol. 71, 704–707, 1998.
7. Li, J., Z. Niu, D. Fang, and Y. Shi, "Numerical analysis of the skin effect in long rectangle conductor," *Journal of Information Engineering University*, Vol. 7, No. 2, 167–171, 2006.
8. Chu, X., X. Jiang, and J. Jiang, "An inverse method to design RF coils with arbitrary conductor patterns for MRI systems," *Proceedings of the CSEE*, Vol. 25, No. 13, 139–143, 2005.

On the Radio-frequency Power Requirements of Human MRI

L. Tang[2] and T. S. Ibrahim[1,2]

[1]Departments of Radiology and Bioengineering, University of Pittsburgh
Pittsburgh, Pennsylvania, USA
[2]School of Electrical and Computer Engineering and Bioengineering Center
University of Oklahoma, Norman, Oklahoma, USA

Abstract— In high and ultrahigh field magnetic resonance imaging (MRI) research, computational electromagnetic techniques are now taking an important role in the design and evaluation of MRI radiofrequency (RF) coils. This paper focuses on the RF power requirements and specific absorption rate (SAR) associated with the MRI operation at different field strength. This paper also presents new techniques for achieving high-quality transmit field homogeneity simultaneously with lower total RF power deposition. The studies are done utilizing the finite difference time domain (FDTD) method and the validation of the methods is performed using ultra high field MRI volume coils.

1. INTRODUCTION

Since magnetic resonance imaging (MRI) technique has been in clinical and research use over the last 30 years, operation at higher magnetic field strength has been a constant goal for the advancement of this diagnostic tool. Although it faces some difficulties such as technical complexity and an increased financial burden, operation at high field MRI is greatly desirable as a result of the associated higher signal-to-noise ratio, contrast-to-noise ratio, and shorter scanning time. Operation at high field MRI and therefore increased frequencies is also associated with complicated interactions of the electromagnetic waves with the tissue since the operating wavelength becomes comparable to or less than the dimensions of the load (human head/body) and RF coil. This can potentially cause severe operational problems such as the presence of inhomogeneous excitation and reception, increased power absorption, and higher local specific absorption rate (SAR).

In human MRI, the total RF power deposition and SAR have been characterized by many researchers [1–4]. For example, at low magnetic field MRI where the wavelength is relatively large compared to the load and RF coil dimensions, quasistatic field approximations were used in the design and assessing the performance of RF coils [5]. Conversely at high or ultrahigh (≥ 7 Tesla) magnetic fields for designing and evaluating RF coils, the significant interactions of the electromagnetic waves with the load invalidate the use of quasistatic approximations and require the application of full wave techniques [6–8]. In this work, a full wave computational electromagnetic method, namely the finite difference time domain (FDTD) technique is implemented in a rigorous fashion by treating the coil and the load as a single system [3] to predict the RF power requirements and SARs of human MRI at high and ultrahigh fields. This computational model is then utilized to design new techniques that can achieve high-quality transmit field homogeneity simultaneously with total RF power deposition lower than that achieved with the standard quadrature excitation [5] for 7 and 9.4 Tesla human MRI.

2. METHODS

2.1. Simulation Model

The model we used is a 16-element TEM resonator [9], which is based on multi-conductor transmission line theory [10], loaded with an anatomically detailed human head mesh [11] as shown in Figure 1. By using the FDTD [12] technique, both the RF coil and the load were modeled as a single system [13] and bounded using perfect matched layers (PML) [14]. In such a modeling approach, the electromagnetic effects on the load due to the coil and on the coil due to the load are included.

From the multi-conductor transmission line theory [10], 9 modes at 9 different frequencies exist in a 16-element TEM coil, where the second mode (mode 1) produces a linearly polarized field (when coil is empty) that can be utilized for imaging. Similar to experiment, the coil was tuned while loaded with the human head mesh by adjusting the gap between the tuning stubs (coil elements).

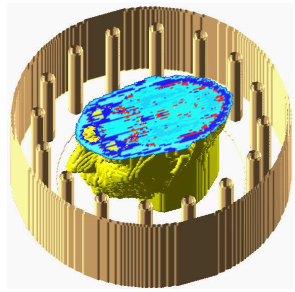


Figure 1: Axial cut of 3D FDTD model system of anatomically detailed human head mesh loaded within a 16-element TEM resonator.

2.2. Power and SAR Calculations

The power input from the RF coil contains the absorbed power by the tissue and the radiated power. It is expressed by Equation (1) derived from the Maxwell equations.

$$P_{\text{input}} = P_{\text{abs}} + P_{\text{rad}} = \frac{\sigma}{2}\iiint_v \left|\vec{E}\right|^2 dv + \frac{1}{2}\iint_s \left(\vec{E}\times\vec{H}^*\right) ds \tag{1}$$

where P_{abs} and P_{rad} are the absorbed and radiated power, σ is conductivity, and $\vec{E}$ and $\vec{H}$ are the electric and magnetic fields respectively. $\iiint_v$ is volume integration over the loaded object, and $\iint_s$ is the closed surface integration that includes the whole system including the coil and the load. Because the RF power absorbed in tissue is related to the flip angle (linearly proportional to one of the circularly polarized components of the transverse magnetic field, known as the B_1^+ field) which directly affects the induced MRI signal, we deliberately only considered the absorbed power in the human head mesh. In numerical simulation, the continuous integration changes to the summation of each FDTD grid as represented by Equation (2).

$$\text{Power} = \sum_i\sum_j\sum_k\left[\frac{1}{2}\sigma_{(i,j,k)}\times\left(E^2_{x_{(i,j,k)}}+E^2_{y_{(i,j,k)}}+E^2_{z_{(i,j,k)}}\right)\right] \tag{2}$$

where $\sigma_{(i,j,k)}$ (S/m) is the conductivity of the FDTD cell at the (i,j,k) location; E_x, E_y and E_z (V/m) are the magnitudes of the electric field components in the x, y, and z directions, respectively; and the summation is performed over the whole volume of the head mesh.

The RF energy during MRI will induce the thermoregulatory imbalance and therefore it is important to monitor the distribution of the energy inside the head. The SAR is intended to indicate the energy absorbed into a tissue of given density by the radio transmitter. In our numerical modeling, it was calculated using the following equation:

$$\text{SAR}_{(i,j,k)} = \frac{1}{2}\frac{\sigma_{(i,j,k)}\left(E^2_{x_{(i,j,k)}}+E^2_{y_{(i,j,k)}}+E^2_{z_{(i,j,k)}}\right)}{\rho_{(i,j,k)}} \tag{3}$$

where $\rho_{(i,j,k)}$ is the tissue density at location (i,j,k). SAR has the unit watt/kilogram.

2.3. Improving the Homogeneity of the B_1^+ Field Distributions

To improve the inhomogeneous excitation (i.e., inhomogeneous B_1^+ field distribution), at 7 and 9.4 Tesla, we utilized an optimization algorithm, based on gradient and genetic algorithms, combined with multi-element/phased-array excitation in order to improve the homogeneity of the B_1^+ field simultaneously with lowering (compared to the standard quadrature excitation) the total RF power deposition in the whole human head.

2.4. Validation of the Electric Field Distributions

To test the validity of the simulation model of the coil-load system, the simulation results were compared to the experimental results obtained utilizing infrared imaging [15]. Figure 2 shows the infrared measured [15] and FDTD calculated [15] square of transverse electric field distribution of an axial slice inside the 16-strut TEM resonator loaded with a spherical saline phantom and tuned to 340 MHz which corresponds to the Larmor frequency of ^{1}H at 8 Tesla.

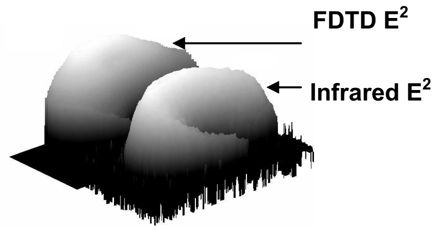


Figure 2: Comparison of infrared images and FDTD calculated electric field distribution at 8 Tesla.

3. RESULTS AND DISSCUSSION

3.1. The Power Absorption Dependency on the External Static Magnetic Field Strength

Using the designed FDTD model, the RF power absorption could be numerically analyzed. Calculations were done at five magnetic field strengths that vary between 4 Tesla and 9.4 Tesla before and after optimization of the homogeneity of the B_1^+ field distribution. The RF power (watts) was scaled to achieve a constant value of 1.174 μ Tesla for the average B_1^+ field intensity in selected slices. The results (Figure 3) interestingly show that higher magnetic fields are not necessary associated with more RF power deposition. For example in Axial Slice 2 and Coronal Slice, compared to 8 Tesla, lower RF power absorption was observed at 9.4 Tesla.

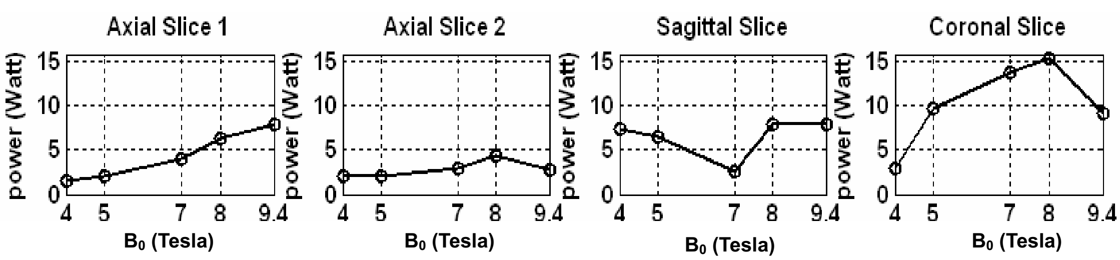


Figure 3: Plots of total power absorbed in the human head as a function of B_0 (external static field) field strength for two axial slices (Axial Slice 1 is a lower brain slice compared to Axial Slice 2) , one sagittal and one coronal slices in order to achieve the same average B_1^+ field intensity and homogeneity (as denoted by the ration of the maximum B_1^+ field intensity divided by the minimum B_1^+ field intensity within the slice of interest).

3.2. Distributions of SAR

The Food and Drug Administration has strict limitations on the peak SAR values and on the continuous MRI examination time. Accurate SAR and total RF power absorption analysis is a must in evaluating new coil designs and excitation techniques. The proposed excitation method demonstrate that a significant improvement in the overall (as denoted by standard deviation) B_1^+

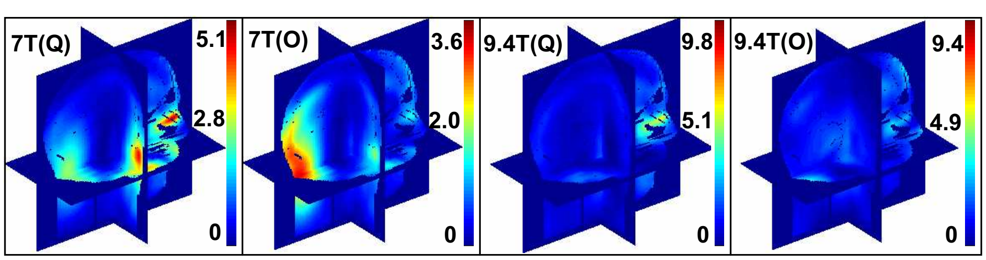


Figure 4: Plots of SAR distributions using quadrature excitation (Q) and optimized (O) B_1^+ field excitation at 7 Tesla (7T) and 9.4 Tesla (9.4T). The corresponding color bars represent the SAR values (watts/kg) for every 10 gm of tissue.

field homogeneity could be achieved while obtaining lower RF power absorptions compared to the standard quadrature excitation. Figure 4 shows some simulation results for SAR distributions at 7 Tesla (7T) and 9.4 Tesla (9.4T) using quadrature excitation (Q) and using optimized excitation (O) over 3-D brain regions. For example at 7 Tesla, compared to the standard quadrature excitation, with optimization of the B_1^+ field over the whole 3-D human brain region:

a. the homogeneity of B_1^+ field distribution was improved by 144%,
b. the peak SAR value was decreased by 29%,
c. and the total RF absorbed power was decreased by 15%.

4. CONCLUSIONS

To obtain high quality images in high and ultra high field human MRI applications, computational electromagnetic techniques, such as FDTD, are taking a significant role in designing the needed RF coils and excitation approaches. Utilizing FDTD method, this work demonstrates that homogenous excitation can be safely achieved at 7 and 9.4 Tesla MRI for human head applications.

REFERENCES

1. Hoult, D. I., "Sensitivity and power deposition in a high-field imaging experiment," *J. Magn. Reson. Imaging*, Vol. 12, No. 1, 46–67, 2000.
2. Collins, C. M. and M. B. Smith, "Signal-to-noise ratio and absorbed power as functions of main magnetic field strength, and definition of "90 degrees" RF pulse for the head in the birdcage coil," *Magn. Reson. Med.*, Vol. 45, No. 4, 684–691, 2001.
3. Ibrahim, T. S., "A numerical analysis of radio-frequency power requirements in magnetic resonance imaging experiment," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 52, No. 8, 1999–2003, 2004.
4. Ibrahim, T. S., R. Lee, B. A. Baertlein, and P.-M. L. Robitaille, "B1 field homogeneity and SAR calculations in the high pass birdcage coil," *Physics in Medicine and Biology*, Vol. 46, 609–619, 2001.
5. Tropp, J., "The theory of the bird-cage resonator," *J. Magn. Reson.*, Vol. 82, 51–62, 1989.
6. Ibrahim, T. S., R. Lee, B. A. Baertlein, A. Kangarlu, and P-M. L. Robitaille, "Application of finite difference time domain method for the design of birdcage RF head coils using multiport excitations," *Magnetic Resonance Imaging*, Vol. 18, 733–742, 2000.
7. Vaughan, J. T., M. Garwood, C. M. Collins, W. Liu, L. DelaBarre, G. Adriany, P. Andersen, H. Merkle, R. Goebel, M. B. Smith, and K. Ugurbil, "7T vs. 4T: RF power, homogeneity, and signal-to-noise comparison in head images," *Magn. Reson. Med.*, Vol. 46, 24–30, Jul. 2001.
8. Wei, Q., F. Liu, L. Xia, and S. Crozier, "An object-oriented designed finite-difference time-domain simulator for electromagnetic analysis and design in MRI–applications to high field analyses," *J. Magn. Reson.*, Vol. 172, 222–30, 2005.
9. Vaughan, J. T., H. P. Hetherington, J. G. Harrison, J. O. Out, J. W. Pan, P. J. Noa, and G. M. Pohost, "High-frequency coils for clinical nuclear magnetic resonance imaging and spectroscopy," *Phy. Med. IX*, 147–153, 1993.
10. Baertlein, B. A., O. Ozbay, T. S. Ibrahim, R. Lee, A. Kangarlu, and P-M. L. Robitaille, "Theoretical model for a MRI radio frequency resonator," *IEEE Trans, Biomed. Eng.*, Vol. 47, 535–546, 2000.
11. Ibrahim, T. S., R. Lee, A. M., Abduljalil, B. A. Baertlein, and P.-M. L. Robitaille, "Effect of RF coil excitation on field inhomogeneity at ultra high fields: A field optimizaed TEM resonator," *Magn. Reson. Imag.*, Vol. 19, 1339–1347, 2001.
12. Yee, K. S., "Numerical solutions of the initial boundary value problems involving Maxwell's equations in isotropic media," *IEEE Trans. Ant. Prop.*, Vol. 14, 302–317, 1966.
13. Chen, J., Z. Feng, and J. M. Jin, "Numerical simulation of SAR and B1-field inhomogeneity of shielded RF coils loaded with the human head," *IEEE Tran. Bio. Eng.*, Vol. 45, No. 5, May 1998.
14. Berenger, J. P., "A perfectly matched layer for the absorption of electromagnetic waves," *J. Computational Physics*, Vol. 114, 185–200, 1994.
15. Ibrahim, T. S. and R. Lee, "Evaluation of MRI RF probes utilizing infrared sensors," *IEEE Tran. Bio. Eng.*, Vol. 53, No. 5, May 2006.

Comparison of Maximum Induced Current and Electric Field from Transcranial Direct Current and Magnetic Stimulations of a Human Head Model

Mai Lu[1], **T. Thorlin**[2], **Shoogo Ueno**[3], **and Mikael Persson**[4]

[1]Department of Signals and systems, Chalmers University of Technology, SE-412 96, Gothenburg, Sweden
[2]Department of Neurology, Institute for Clinical Neuroscience, Sahlgrens University Hospital
Bla Straket 7, SE 413 45 Gothenburg, Sweden
[3]Department of Biomedical Engineering, Graduate School of Medicine
University of Tokyo, Tokyo 113-0033, Japan
[4]Department of Signals and Systems, Chalmers University of Technology, SE-412 96, Gothenburg, Sweden

Abstract— As important non-invasive techniques in brain stimulation, transcranial direct current stimulation (tDCS) and transcranial magnetic stimulation (TMS) have been studied and compared in this paper by employing impedance method and a 3D human head model. The quantitative analysis of distributions of current density and electric field by tDCS and TMS have been presented. Results are compared and potential applications are discussed.

1. INTRODUCTION

Transcranial magnetic stimulation (TMS) is a technique for stimulating the brain. Transcranial magnetic stimulation uses powerful rapidly changing magnetic fields to induce electric fields in the brain by electromagnetic induction without the need for surgery or external electrodes. As a non-invasive method to stimulate brain, TMS has attracted considerable interest as an important tool for studying the functional organization of the human brain as well as a therapeutic tool in numerous clinical trials to improve a variety of psychiatric diseases [1, 2].

Recently another non-invasive method transcranial direct current stimulation has been developed. The tDCS method involves application of low intensity direct current stimulation of cortex through large surface scalp electrodes. The principle of how tDCS works in the brain is roughly the same as that of TMS. They both seek to make neurons in the prefrontal cortex more excitable. tDCS has already been shown to improve implicit motor learning, and visuo-motor coordination. Moreover, with regard to neurologic diseases e.g., epilepsy, depression, migraine, and Parkinson's disease, it offers new interesting therapeutic option [3, 4].

Although tDCS shows promise in fighting psychiatric disorders, we need to know a lot more before it can be accepted as an effective treatment method. To date the spatial distribution of the current density and electric field within the volume of the human brain for tDCS is largely unknown. Also there has been no comparison between the tDSC and TMS methods. This is main motivation of the present work. By employing the impedance method, and a 3D human head model, this paper provides a quantitative analysis and a comparison of current density and electric field distributions in the human head model by tDCS and TMS.

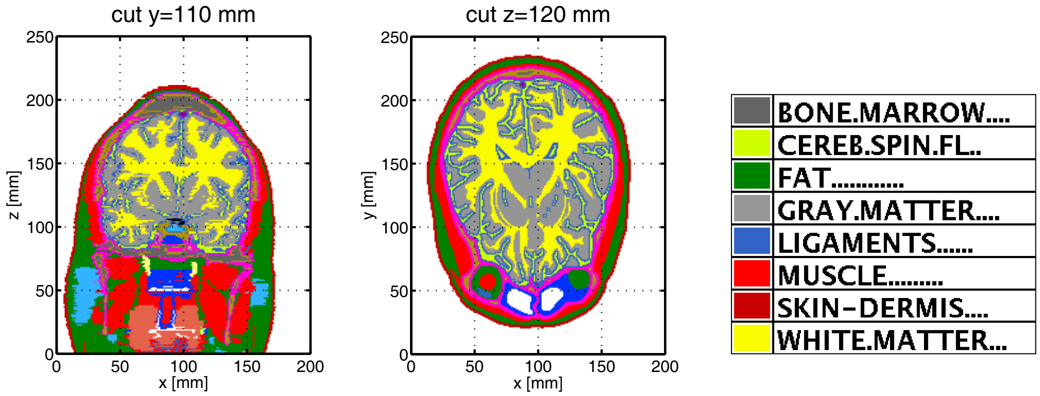
2. HUMAN HEAD MODEL

In this paper we use a 3D human head model obtained from Brooks Air Force Laboratory, USA. The model which has 24 different tissues is based on anatomical slices from a male cadaver and it is originally obtained from the Visible Human Project. The electrical properties are modeled using the 4-Cole-Cole model [5]. Typical sliced layer in the head model and tissue colour palette for part of the tissues are shown in Fig. 1. Conductivities of some important tissues used in the present paper are given in Table 1.

3. 3-D IMPEDANCE METHOD

The human head model is described using a uniform 3D Cartesian grid and is composed of small cubic voxels. Assuming that, in each cell, the electric conductivities are isotropic and constant in all direction, the model is represented as a 3D network of impedances. The impedances for various directions can be written as

$$Z_m^{i,j,k} = \frac{\Delta m}{\Delta n \Delta p(\sigma + j\omega\epsilon)} \tag{1}$$



(a) (b) (c)

Figure 1: Typical sliced layers in human head model. (a) cross section at $y = 110\,\text{mm}$, (b) cross section at $z = 120\,\text{mm}$, (c) Tissue color Palette.

Table 1: Tissue properties used in the calculations.

Tissue	Conductivity $\sigma[S/m]$ (TMS case)	Conductivity $\sigma[S/m]$ (tDCS case)
BLOOD	7.00e-01	7.00e-01
BONE.MARROW	2.52e-3	5.00e-04
CEREB.SPIN.FL	2.00e+00	2.0e+00
FAT	2.34e-02	1.00e-02
LIGAMENTS	3.85e-01	2.5e-01
NERVE.(spine)	3.21e-02	6.00e-03
GRAY.MATTER	1.07e-01	2.00e-02
WHITE.MATTER	6.55e-02	2.00e-02
SKIN-DERMIS	2.01e-4	2.00e-04

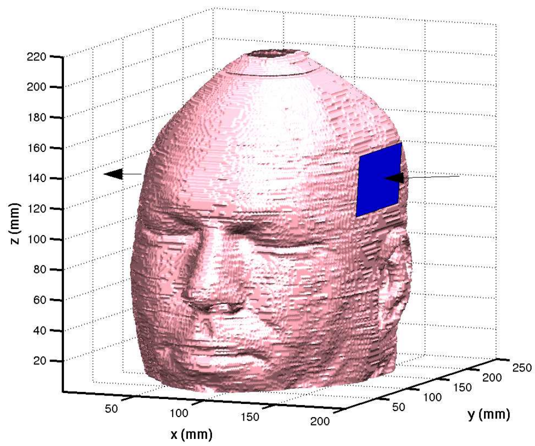
where i, j, k indicate the cell index; Δm, Δn, and Δp are the size of the voxels in m, n, p directions. σ and ϵ are the conductivity and the electrical permittivity for the voxel(i, j, k). Kirchoff voltage law around each loop in this network generates a system of equations for the loop currents. In the case of TMS, these loop currents are driven by Faraday induction from the magnetic field of the applicator. In the tDCS case, the currents are injected at the electrodes and then distributed according to the Kirchoff laws. This system of equations is solved numerically using a standard iterative method. The net induced currents within the head are then calculated from these known loop currents. The induced electric field is in turn calculated from the net induced currents using the Ohm's law. Details implementation of the impedance method can be found in[6–8].

4. RESULTS AND DISCUSSIONS

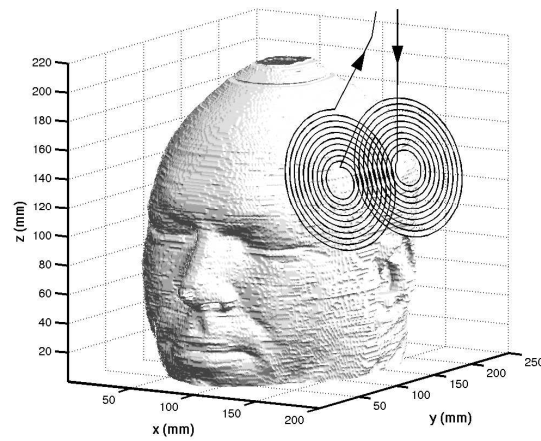
For the tDCS calculation, we use a pair of large surface scalp electrodes with an area of 35cm^2 that are placed on the 3D head model in the bilateral position, as shown in Fig. 2(a). A direct current of 1 mA is injected from right electrode and extracted the same current on the left side. For the TMS calculation, we design a TMS coil with figure of eight shaped and place it near the left side of head model (Fig. 2(b)). It consists of two circular coils with inner radius of 10 mm, and outer radii 50 mm, and the number of wire turns in each wing is $n_r = 10$. A typical clinic application current (sine wave with current amplitude $I = 7.66\,\text{kA}$, and working frequency $f = 3.6\,\text{kHz}$ with repetition of 20 Hz) was implemented in the calculations.

By employing the impedance method as described in Section 3, the current density **J** and electric field **E** are calculated, and the results for tDCS and TMS cases are shown in Table 2 and Table 3, respectively.

For tDCS case, we know that fat (at $x = 69\,\text{mm}$, $y = 114\,\text{mm}$, $z = 202\,\text{mm}$) and skin-dermis (at $x = 112\,\text{mm}$, $y = 137\,\text{mm}$, $z = 208\,\text{mm}$) under the right patch exhibit maximum values



(a)



(b)

Figure 2: tDCS patch and TMS coil placed on the human head model. (a) tDCS, (b) TMS.

Table 2: Maximum values of current density **J**, electric field **E** and their 95%-CI in different tissues for tDCS case.

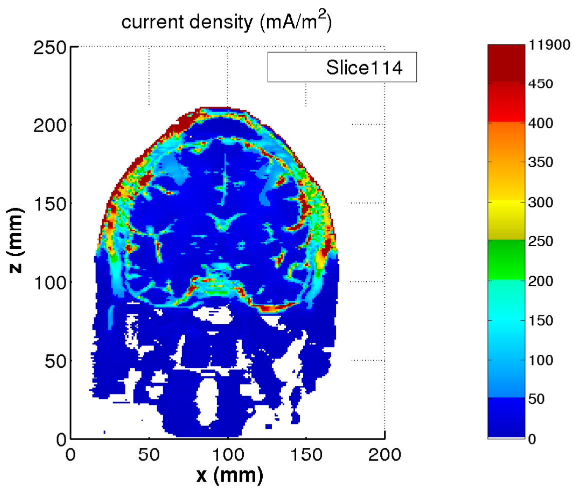
Tissue Name	$\lvert\mathbf{J}\rvert$ [mA/m^2]	95%-CI [mA/m^2]	$\lvert\mathbf{E}\rvert$ [mv/m]	95%-CI [mv/m]	(x, y, z) (unit: mm)
BLOOD..........	5.6e+02	4.0e+01	8.0e+02	5.7e+01	(104, 73,174)
BODY.FLUID.....	2.4e+01	2.7e+00	1.6e+01	1.8e+00	(126, 74, 13)
BONE.MARROW....	1.1e+03	1.3e+02	2.1e+06	2.7e+05	(72,126,197)
CARTILAGE......	1.1e+02	6.3e+00	7.2e+02	4.2e+01	(103, 20, 85)
CEREB.SPIN.FL..	2.5e+03	1.6e+02	1.2e+03	8.0e+01	(144,132,163)
EYE.(lens).....	1.1e+02	1.2e+01	3.8e+02	4.1e+01	(128, 48,110)
FAT............	1.2e+04	7.6e+02	1.2e+06	7.6e+04	(69,114,202)
GLANDS.........	3.4e+02	6.6e+00	6.9e+02	1.3e+01	(90,109,100)
LIGAMENTS......	8.6e+03	1.1e+03	3.4e+04	4.3e+03	(71,114,200)
MUSCOUS.MEMB...	6.2e+02	9.0e+01	1.6e+06	2.2e+05	(98, 69,113)
MUSCLE.........	7.8e+03	1.5e+03	3.9e+04	7.6e+03	(71,125,202)
NERVE.(spine)..	3.7e+02	4.4e+01	6.2e+04	7.4e+03	(89,106,106)
GRAY.MATTER....	9.9e+02	8.9e+01	5.0e+04	4.4e+03	(133,137,163)
WHITE.MATTER...	3.5e+02	4.0e+01	1.7e+04	2.0e+03	(88,133,139)
CEREBELLUM.....	2.2e+02	2.0e+01	5.4e+03	4.9e+02	(64,147, 91)
SKIN-DERMIS....	1.0e+04	5.8e+02	5.2e+07	2.9e+06	(112,137,208)

of current density: $|\mathbf{J}| = \mathbf{1.2} \times \mathbf{10^4}\,\text{mA/m}^2$, and $|\mathbf{J}| = \mathbf{1.0} \times \mathbf{10^4}\,\text{mA/m}^2$, respectively. Skin-dermis (same position as above) and bone marrow (at $x = 72\,\text{mm}$, $y = 126\,\text{mm}$, $z = 197\,\text{mm}$) under the left patch exhibits maximum values electric field: $|\mathbf{E}| = \mathbf{5.2} \times \mathbf{10^7}\,\text{mV/m}$, and $|\mathbf{E}| = \mathbf{2.1} \times \mathbf{10^6}\,\text{mV/m}$ respectively. This means that the current density and the electric field in tDCS are mostly distributed in the skull. The distributions of current density $|\mathbf{J}|$ and electric field $|\mathbf{E}|$ at the cross sections have been illustrated in Fig. 3. From Table 2, we know that brain tissues, such as spine nerve has maximum value of current density $|\mathbf{J}| = \mathbf{370}\,\text{mA/m}^2$ and electric field $|\mathbf{E}| = \mathbf{6.2} \times \mathbf{10^4}\,\text{mV/m}$ at the position $(89, 106, 106)$. As we know this position is located in the central and deep part of the brain. Since $|\mathbf{E}| = \mathbf{6.2} \times \mathbf{10^4}\,\text{mV/m}$ can be comparable with 100 mV/mm (the effective fields for electrically induced effects in living cells [9]), we believe tDCS can also play a role in deep brain stimulation.

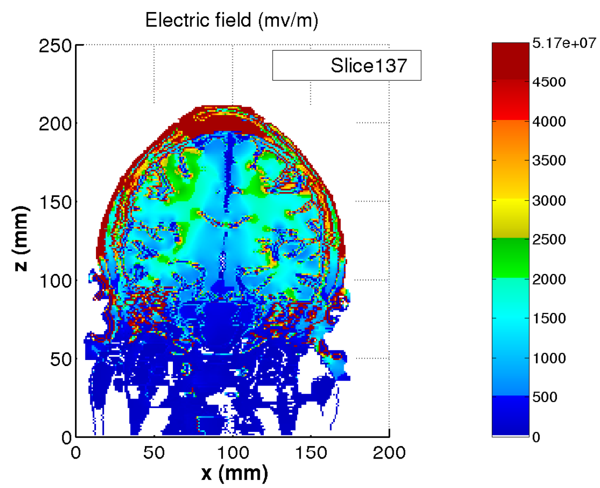
For TMS case, cerebro spinal fluid (at $x = 149\,\text{mm}$, $y = 113\,\text{mm}$, $z = 149\,\text{mm}$), and ligaments (at $x = 149\,\text{mm}$, $y = 113\,\text{mm}$, $z = 147\,\text{mm}$) at the coil side exhibit maximum values of current density

Table 3: Maximum values of current density **J**, electric field **E** and their 95%-CI in different tissues for TMS case.

Tissue Name	$\|\mathbf{J}\|$ [mA/m^2]	95%-CI [mA/m^2]	$\|\mathbf{E}\|$ [mv/m]	95%-CI [mv/m]	(x, y, z) (unit: mm)
BLOOD..........	3.8e+04	2.0e+03	5.5e+04	2.9e+03	(98, 56,155)
BODY.FLUID.....	1.2e+04	1.3e+03	7.8e+03	8.8e+02	(126, 74, 13)
BONE.MARROW....	2.7e+04	2.9e+03	1.1e+07	1.2e+06	(156,112,144)
CARTILAGE......	9.8e+03	1.0e+03	5.6e+04	5.9e+03	(164,137, 82)
CEREB.SPIN.FL..	4.1e+05	3.9e+04	2.0e+05	2.0e+04	(149,113,149)
EYE.(lens).....	3.6e+04	4.2e+03	1.1e+05	1.3e+04	(135, 48,107)
FAT............	9.7e+04	8.7e+03	4.2e+06	3.7e+05	(158,102,152)
GLANDS.........	5.8e+04	5.9e+03	1.1e+05	1.1e+04	(147, 58,104)
LIGAMENTS......	2.9e+05	3.0e+04	7.5e+05	7.8e+04	(149,113,147)
MUSCOUS.MEMB...	4.4e+04	6.1e+03	4.2e+07	5.8e+06	(106, 70,111)
MUSCLE.........	1.8e+05	6.3e+03	5.3e+05	1.9e+04	(158,103,151)
NERVE.(spine)..	2.2e+04	2.4e+03	6.8e+05	7.5e+04	(124, 65,106)
GRAY.MATTER....	1.3e+05	1.1e+04	1.2e+06	1.0e+05	(151,110,138)
WHITE.MATTER...	2.9e+04	1.7e+03	4.4e+05	2.6e+04	(150,115,130)
CEREBELLUM.....	2.2e+04	1.9e+03	1.7e+05	1.5e+04	(140,159, 82)
SKIN-DERMIS....	3.7e+04	4.8e+03	1.8e+08	2.4e+07	(156, 64,100)



(a)

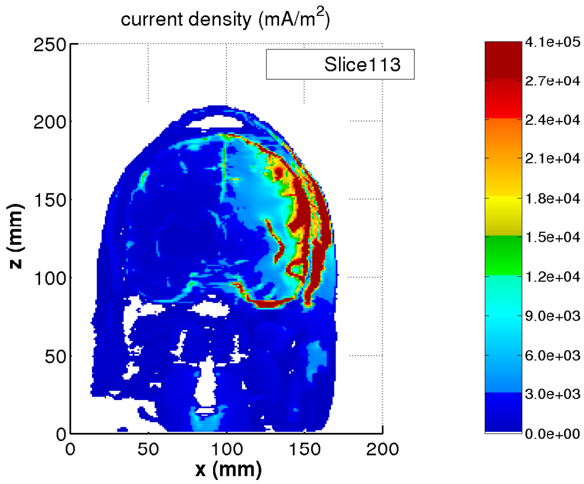


(b)

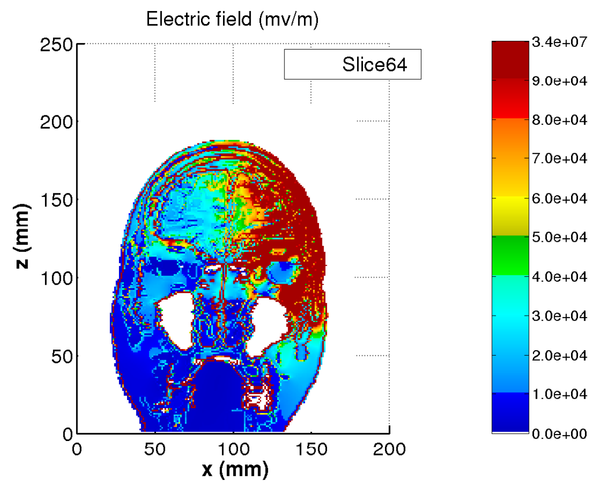
Figure 3: Current density ($|\mathbf{J}|$) and electrical field ($|\mathbf{E}|$) distributions for tDCS case. (a) $|\mathbf{J}|$ at the cross section $y = 114$ mm, (b) $|\mathbf{E}|$ at the cross section $y = 137$ mm.

$|\mathbf{J}| = \mathbf{4.1} \times \mathbf{10^5}$ mA/m^2, and $|\mathbf{J}| = \mathbf{2.9} \times \mathbf{10^5}$ mA/m^2, respectively. Electric field exhibits maximum values $|\mathbf{E}| = \mathbf{1.8} \times \mathbf{10^8}$ mV/m in skin (at $x = 156$ mm, $y = 64$ mm, $z = 100$ mm) at the coil side, and $|\mathbf{E}| = \mathbf{4.2} \times \mathbf{10^7}$ mV/m in muscous membranes (at $x = 106$ mm, $y = 70$ mm, $z = 111$ mm) in the deep part of the brain. This means during TMS implementation, current density will mostly distributed in the brain area, because the applied magnetic field can easily penetrate the skull. While electric field will be distributed in the scalp as well as in the brain. Fig. 4 illustrates the distribution of current density $|\mathbf{J}|$ and electric field $|\mathbf{E}|$ in cross sections in TMS case.

In the head model with 1 mm resolution, the thickness of the tissues is often close to the grid size, and severe staircasing will locally perturb the calculation. A statistical analysis was thus carried out to estimate the maximum $|\mathbf{J}|$ and $|\mathbf{E}|$ with a 95% confidence interval (95%-CI). Table 2 and Table 3 summarise 95%-CI in various tissues for tDCS case and TMS case, respectively. For example, the relative errors of $|\mathbf{J}|$ and $|\mathbf{E}|$ in skin for tDCS case are 5.8% and 5.6%, respectively, while in the brain tissue i.e., white matter, the estimated relative errors are 11.4% for $|\mathbf{J}|$ and 11.7%



(a)



(b)

Figure 4: Current density ($|\mathbf{J}|$) and electrical field ($|\mathbf{E}|$) distributions for TMS case. (a) $|\mathbf{J}|$ at the cross section $y = 113$ mm, (b) $|\mathbf{E}|$ at the cross section $y = 64$ mm.

for $|\mathbf{E}|$, which also reveals a fact that tDCS causes large current density and electric field levels in the skull.

Theoretical studies show that a nerve is activated by the first derivative of the component of an induced electric field along the nerve, the so called activating function, during magnetic stimulation [10, 11]. It is important to investigate the activating functions in human brain by tDCS, and TMS, and the precise study will be reported soon.

REFERENCES

1. Post, R. M., T. A. Kimbrell, U. D. McCann, R. T. Dunn, E. A. Osuch, A. M. Speer, and S. R. B. Weiss, "Repetitive transcranial magnetic stimulation as a neuropsychiatric tool: present status and future potential," *Journal of ECT*, Vol. 15, No. 1, 39–59, 1999.
2. Ueno, S., "Biomagnetic approaches to studying the brain," *IEEE Engineering in Medicine and Biology*, Vol. 18, No. 3, 108–120, 1999.
3. Nitsche, M. A., A. Schauenburg, N. Lang, D. Liebetanz, C. Exner, W. Paulus, and F. Tergau, "Facilitation of implicit motor learning by weak transcranial direct current stimulation of the primary motor cortex in the human," *J. Cong. Neurosci*, Vol. 15, No. 4, 619–626, 2003.
4. Liebetanz, D., F. Fregni, K. K. Monte-Silva, M. B. Oliveira, et al., "After-effects of transcranial direct current stimulation (tDCS) on cortical spreading depression," *Neuroscience Letters*, Vol. 398, No. 1–2, 85–90, 2006.
5. Cole, K. S. and R. H. Cole, "Dispersion and absorption in dielectrics: alternating current characteristics," *J. Chem. Phys.*, Vol. 9, No. 9, 341–351, 1941.
6. Orcutt, N. and O. P. Gandhi, "A 3-D impedance method to calculate power deposition in biological bodies subjected to time varying magnetic fields," *IEEE Trans. on Biomedical Engineering*, Vol. 35, No. 8, 577–583, 1988.
7. Nadeem, M., T. Thorlin, O. P. Gandhi, and M. Persson, "Computation of electric and magnetic stimulation in human head using the 3D impedance method," *IEEE Trans. on Biomedical Engineering*, Vol. 50, No. 7, 900–907, 2003.
8. Nadeem, M., Y. Hamnerius, K. H. Mild, and M. Persson, "Magnetic field from spot welding equipment — Is the basic restriction exceeded?," *Bioelectromagnetics*, Vol. 25, No. 4, 278–284, 2003.
9. Robinson, K. R. and M. A. Messerli, "Electric embryos: The embryonic epithelium as a generator of developmental information," *Frontiers in Neurobiology 2: Nerve Growth and Guidance*, 131–141, C. D. MCCaig, Ed. London, U.K., Portland, 1996.
10. Roth, B. J. and P. J. Basser, "A model of the stimulation of a nerve fiber by electromagnetic induction," *IEEE Trans. Biomed.Eng.*, Vol. 37, No. 4, 588–597, 1990.
11. Liu, R. and S. Ueno, "Calculating the activating function of nerve excitation in inhomogeneous volume conductor during magnetic stimulation using the finite element method," *IEEE Trans. Magn.*, Vol. 36, No. 4, 1796–1799, 2000.

Magnetic Field Produced by Compound Action Potential of Degenerated Human Nerve

T. Hayami[1] **and K. Iramina**[2]

[1]Digital Medicine Initiative, Kyushu University, Japan
[2]Graduate School of Information Science and Electrical Engineering, Kyushu University, Japan

Abstract— In this study, we discussed about a new diagnostic technique of peripheral nervous disturbance using biomagnetic measurement. The magnetic field around a limb after nerve stimulation was simulated for the cases when some of the nerve fibers composing the nerve are lacked. As a result, disappearance of a pair of peaks of in- and out-flux was observed. The magnetic field seems to be reflective on the change of the distribution of the number of nerve fibers inside a nerve.

1. INTRODUCTION

The nerve conduction velocity test is a general method to diagnose diabetic neuropathy or any other malfunctions of peripheral nerves. Usually in this test, the conduction velocity of skin potential generated by the compound action potential of the nerve is measured. However as the most part of the skin potential is contributed to by thick nerve fibers, degeneration of thin nerve fibers is difficult to find with this test [1].

Measurement of magnetic field as the substitute of skin potential has an advantage to analyze the property of compound action potential. The action potential of a nerve fiber is possible to generate a pair of positive and negative peaks, whose location reflects the depth and the thickness of the fiber. On the contrary, the magnetic field of it has two pairs of peaks of in- and outmagnetic flux [2]. As each location reflects the depth and the thickness of the fiber, the more implications about morphology seemed to be obtained. To investigate the relationship between compound action potential and the magnetic field of its produce, the magnetic field around a limb after electric current stimulation was simulated.

2. METHODS

A human nerve was modeled as a straight cylindrical tube of 2 mm in diameter. The nerve was composed of myelinated nerve fibers with different diameters (1–14 μm). The fibers of each thickness were distributed uniformly inside the nerve arrayed in parallel inside the cylinder. Each fiber was assumed as a straight chain of a node of Ranvier, connected with an axoplasm electrical resistivity (Fig. 1). The parameters of membrane characteristics were referred to the experimental data of human nerve measurement in vitro [3], corrected with Q_{10} values [4, 5] into 37°C. The relationship between the diameter of a fiber and the diameter of the axon and internodal distance was followed to a precedent nerve fiber model [6] proposed according to a morphometric study of human sural nerve [7].

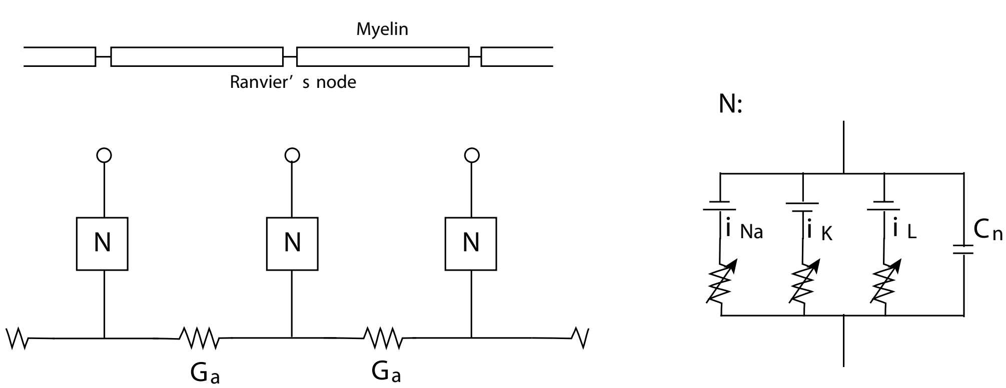


Figure 1: The nerve fiber model used as components of a nerve. G_a: Axoplasm conductance; i_{Na}, i_K, i_L: Nodal currents of potassium, sodium, and leakage; C_n: Nodal capacitance.

Electric current stimulation was given as a rectangle current pulse with duration of 100 μsec at a single point 10 mm off from the core of the nerve. As the fibers were assumed to be immersed in the liquid of uniform conductivity, the potential given to every node was determined as inversely proportional value to the distance of the node from the stimulation point [8].

Although the primary current source of the magnetic field around a limb is every axonal current of nerve fibers inside the stimulated nerve, a plenty of computer resources are required to calculate the current for some hundreds or thousands of nerve fibers. To reduce the amount of calculation, excitation rate e was defined as

$$e = S/\pi r^2 \tag{1}$$

where

$$S = R^2\left(\theta_1 - \frac{\sin(2\theta_1)}{2}\right) + r^2\left(\theta_2 - \frac{\sin(2\theta_2)}{2}\right) \tag{2}$$

$$\theta_1 = \cos^{-1}\left(\frac{l^2+R^2-r^2}{2lR}\right), \quad \theta_2 = \cos^{-1}\left(\frac{l^2-R^2-r^2}{2lr}\right) \tag{3}$$

In these equations, R is the distance from the stimulation point to the farthest excitable nerve fiber for the stimulation pulse. r is the diameter of the nerve. l is the distance between the stimulation point and the core of the nerve cylinder. On the orthogonal plane to the nerve including the stimulation point (Fig. 2), e denotes the rate of the common area of two circles to the area of cross-section circle of the nerve.

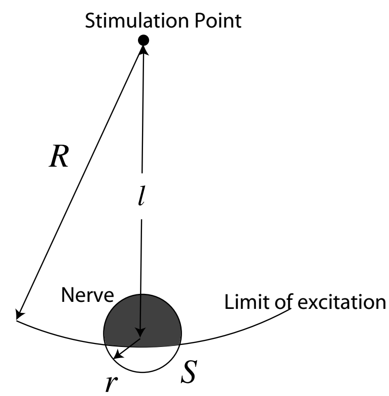


Figure 2: The orthogonal plane to the nerve including the stimulation point to make definition of excitation rate e.

Using e, the total axonal current I_c inside the nerve may be described as

$$I_c = \int I_f(k)e(k)n(k)dk \tag{4}$$

where $I_f(k)$ is the axonal current of an active nerve fiber with diameter k μm, $e(k)$ is the excitation rate for the diameter, and $n(k)$ denotes the number of fibers inside the nerve. The distribution of the number of nerve fiber $n(k)$ assumed as Fig. 3.

The limb was assumed as a cylinder of 100 mm in diameter. The nerve was laid at the depth of 10 mm from the surface. The magnetic field on the surface of the cylinder induced by the total axonal current I_c was calculated with boundary element method.

3. RESULT

The magnetic field after 5 msec from stimulation is shown in Fig. 4. The magnetic field after 5 msec from stimulation is shown in Fig. 4. For the case of (a), (d), (e), (f) and (h), two pairs of in- and out-flux were observed obviously, however in (b), (c) and (g), a pair of in- and out- flux seems to be disappeared. Although a preceding experimental study showed that the magnetic field consists of two pairs of in- and out-magnetic flux [9], this result can be read that the pattern of magnetic field around limbs is possible to be reflective on the degeneration of nerve for some cases when the disappearance of nerve fiber has a relationship to the diameter of the fiber.

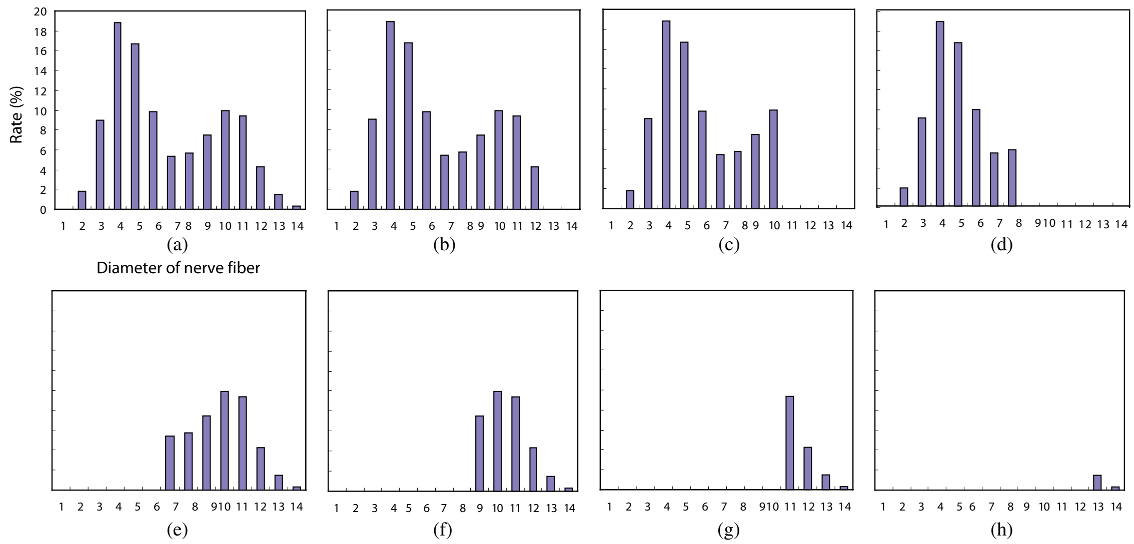


Figure 3: The distribution of the number of nerve fiber $n(k)$ tested in this simulation, (a) General distribution [7], (b)–(d) Maximum diameters are limited to 12-8 µm respectively, (e)–(h) Minimum diameters are limited to 6-12 µm respectively.

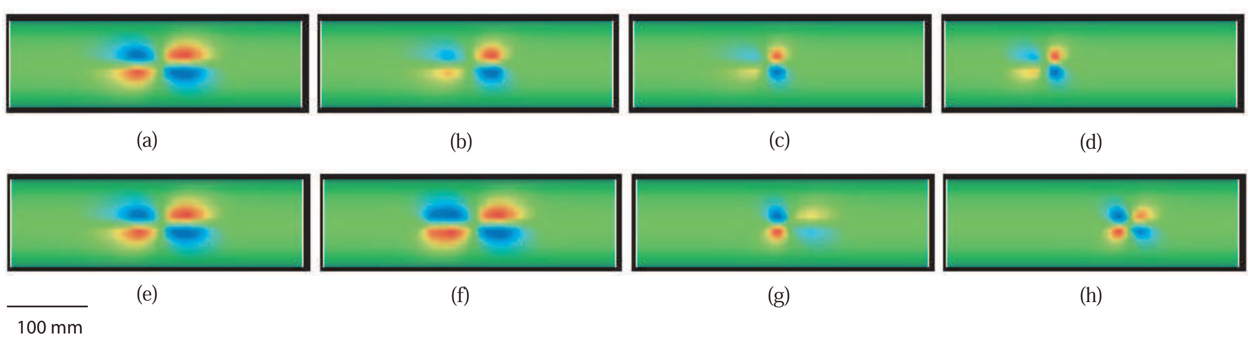


Figure 4: The normal component of the magnetic field around the cylinder at 5 msec from the stimulation.

REFERENCES

1. Veen, B. K., R. L. Schellens, D. F. Stegeman, R. Schoonhoven, and A. A. Gabreels-Festen, “Conduction velocity distributions compared to fiber size distributions in normal human sural nerve,” *Muscle & Nerve*, Vol. 18, 1121–1127, 1995.
2. Wikswo, J. P. , “Biomagnetic sources and their models,” *Advances in Biomagnetism*, 1–18, 1990.
3. Schwarz, J. R., G. Reid, and H. Bostock, “Action potentials and membrane currents in the human node of Ranvier,” *Pflügers Arch.*, Vol. 430, 283–292, 1995.
4. Schwarz, J. R. and G. Eikhof, “Na currents and action potentials in rat myelinated nerve fibers at 20 and 37°C,” *Pflügers Arch.*, Vol. 409, 569–577, 1987.
5. Frankenhaeuser, B. and L. E. Moore, “The effect of temperature on the sodium and potassium permeability changes in myelinated nerve fibres of Xenopus laevis,” *J. Physiol.*, Vol. 169, 431–437, 1963.
6. Wesselink, W. A., J. Holsheimer, and H. B. K. Boom, “A model of the electrical behaviour of myelinated sensory nerve fibres based on human data,” *Med. Biol. Eng. Comput.*, Vol. 37, 228–235, 1999.
7. Behse, F. “Morphometric studies on the human sural nerve,” *Acta Neurol. Scand. Suppl.*, Vol. 82, 1–38, 1990.
8. McNeal, D. R., “Analysis of a model for excitation of myelinated nerve,” *IEEE Trans. BME*, Vol. 23, 329–337, 1976.
9. Hoshiyama, M., R. Kakigi, and O. Nagata, “Peripheral nerve conduction recorded by a micro gradiometer system (micro-SQUID) in humans,” *Neurosci. Lett.*, Vol. 272, 199–202, 1999.

EM Methods for MIC Modeling and Design: An Overview

Mustapha C. E. Yagoub[1] **and Mohamed L. Tounsi**[2]

[1]SITE, University of Ottawa, 800 King Edward, Ottawa, Ontario, Canada K1N 6N5
[2]Faculty of Electronics and Informatics, USTHB, P.O. Box 32, El-Alia, 16111, Algiers, Algeria

Abstract— Higher integration and smaller layout size, two major trends in today's RF/microwave industry, lead to more prominent electromagnetic high order effects. In this paper, the authors present a survey of existing electromagnetic techniques to compute such effects in microwave integrated circuits, highlighting their specific advantages/disadvantages for a circuit designer.

1. INTRODUCTION

Since their introduction in the 1950s, microwave integrated circuits (MICs) have played an important role in advancing the RF/microwave technologies. This progress would not have been possible without the advances of solid-state devices and planar transmission lines (TLs). These structures are the backbone of MICs, and represent an important research topic for many engineers. Along with the advances of MICs and planar TLs, numerous electromagnetic (EM) methods for RF/microwave passive structures have been developed for accurate MIC analysis and design. It is no longer efficient, or even feasible, to tune a circuit once fabricated. Therefore, accurate characterization methods that can include high frequency EM effects are needed to design the structures. These methods have in turn helped further investigation and development of new planar lines. Not only have planar lines fulfilled their most fundamental objective of delivering signals, but they can also be exploited to create various RF/microwave devices, such as wideband hybrid junctions, by appropriately combining them. Furthermore, any EM numerical method needs to be as efficient as possible in terms of both CPU time and storage requirements, although recent advances in computers impose less severe restrictions on the method. Therefore, numerical methods are chosen on the basis of trade-offs between accuracy, speed, storage requirement and versatility and are often structure-dependent. In this paper, the authors present a brief overview of existing EM numerical methods used in MIC modeling and design, highlighting their specificities for circuit designers.

2. EM NUMERICAL METHODS FOR MICROWAVE INTEGRATED CIRCUITS

Numerical solution of EM problems started in the mid-1960s with the availability of modern high-speed computers. Since then, considerable effort has been expended on solving complex EM-related problems for which closed form analytical solutions are either intractable or do not exist. Based on Maxwell's equations, each numerical method has its own unique advantages and disadvantages for specific needs. These methods, in fact, provide a foundation for the derivation of current and future analysis methods. They not only represent some of the most useful and commonly used techniques for analyzing planar lines, but also serve as means to present the fundamentals of applying EM theory to the analysis of boundary-value problems. There are two approaches in analyzing a TL: quasi-static and full-wave. The first produces line parameters for the TEM mode only. On the other hand, the dynamic approach can produce the line parameters not only for the TEM but also for the hybrid modes, whose parameters are frequency dependent. TEM mode parameters obtained by the static approach are theoretically only valid at dc. However, a number of millimeter-wave circuits up to W-band have successfully been designed using only static results. Nevertheless, at microwave and millimeter-wave frequencies, a dynamic approach remains the most appropriate for more accurate determination of the line parameters.

2.1. Variational Methods

In EM problems, solutions are usually obtained by directly solving appropriate differential or integral equations while variational methods operate by seeking a functional that gives the maximum/minimum of a desired quantity [1, 2]. Its main advantage is that it produces stationary formulas, which yield results insensitive to the first-order errors. There are three kinds of variational methods, depending on the technique used to obtain approximate solutions expressed in a variational form: the direct method based on the classical Rayleigh-Ritz or simply Ritz procedure, the indirect method such as Galerkin and least squares, and the semi-direct method based on separation of variables. Applications of variational methods include analysis of TLs to obtain characteristic impedances, effective dielectric constants, and losses, analysis of discontinuities, determination of resonant frequencies of resonators, and determination of impedances of antennas and obstacles in waveguides.

2.2. Spectral Domain Method (SDA)

The spectral-domain analysis (SDA) is extensively used in analyzing planar transmission lines, resonators, and scattering problems [3–5]. It is basically a Fourier transformed version of the integral equation method; but compared to the conventional space-domain integral equation method, the SDA has several advantages: (i) its formulation results in a system of coupled algebraic equations instead of coupled integral equations, (ii) closed-form expressions can easily be obtained for the Green's functions, (iii) incorporation of physical conditions of analyzed structures via the so-called basis functions is achieved with stationary solutions. These features make the SDA numerically simpler and more efficient than the conventional integral equation method.

2.3. Mode-Matching Method (MMM)

The mode-matching method is a useful technique for structures consisting of two or more separate regions [6, 7]. Based on matching the EM field at the boundaries of the different regions, it lends itself naturally to boundary-value problems. It has widely been used for scattering and transmission problems, as well as TL analysis. Scattering problems include discontinuities in waveguides and transmission lines, and obstacles in a medium. Transmission problems include analysis of filters, impedance transformers, and power dividers. TL analyses include determination of the line propagation constant and characteristic impedance, such as those of microstrip lines and coplanar strips. Coplanar strips with finite strip metallization thickness have been analyzed.

2.4. Finite Difference Method (FDM)

The finite difference method has been applied to solve many EM-related problems such as TL problems and waveguides [8–10]. Finite difference solution basically involves three steps: (i) dividing the solution region into a grid of nodes (ii) approximating the given differential equation by finite difference equivalent that relates the dependent variable at a point in the solution region to its values at the neighboring points (iii) solving the difference equations subject to the prescribed boundary conditions and/or initial conditions.

2.5. Method of Moments (MOM)

The use of MOM [11, 12] has been successfully applied to a wide variety of EM problems of practical interest such as radiation due to thin-wire elements and arrays, scattering problems, microstrip and lossy structures, propagation over an inhomogeneous earth, and antenna beam pattern, to mention a few. The literature on MOM is already so large as to prohibit a comprehensive bibliography. A partial bibliography is provided by Adams [13]. The procedure for applying MOM usually involves four steps: (i) derivation of the appropriate integral equation (IE), (ii) conversion (discretization) of the IE into a matrix equation using basis (or expansions) functions and weighting (or testing) functions, (iii) evaluation of the matrix elements, and (iv) solving the matrix equation and obtaining the output parameters.

2.6. Finite Element Method (FEM)

Although the FDM and the MOM are conceptually simpler and easier to program than the finite element method (FEM), FEM is a more powerful and versatile numerical technique for handling problems involving complex geometries and inhomogeneous media. The systematic generality of the method makes it possible to construct general-purpose computer programs for solving a wide range of problems in different fields and with little modifications [14]. The finite element analysis involves basically four steps [15]: (i) discretizing the solution region into a finite number of subregions or *elements*, (ii) deriving governing equations for a typical element, (iii) assembling all elements in the solution region and, (iv) solving the obtained system of equations.

2.7. Transmission-Line Matrix Method (TLM)

The TLM is a numerical technique for solving field problems using equivalent circuit representation. It is based on the equivalence between Maxwell's equations and voltages/currents equations on a mesh of continuous two-wire transmission lines [16, 17]. A major advantage of the TLM as compared with other numerical techniques, is the ease with which even the most complicated structures can be analyzed. Its flexibility and versatility reside in the fact that the mesh incorporates the EM field properties and their interaction with the boundaries and material media. Hence, the EM problem needs not be formulated for every new structure. Another advantage of using the TLM is that there are no problems with convergence, stability or spurious solutions. The method is limited only by the amount of memory storage required by the mesh. Also, being an explicit numerical solution, the TLM method is suitable for nonlinear or inhomogeneous problems since any variation of material properties may be updated at each time step. Note that the TLM method is a physical discretization approach, compared to the FDM and FEM which are mathematical discretization approaches.

In the TLM, a field discretization involves replacing a continuous system by an array of lumped elements and dividing the solution region into a rectangular mesh of transmission lines. Junctions are formed where the lines cross forming impedance discontinuities. A comparison between the TL and Maxwell's equations allows equivalences to be drawn between voltages/currents on the TL and EM fields in the solution region. Thus, the TLM involves two basic steps [18]: (i) replacing the field problem by the equivalent network and deriving analogy between the field and network quantities. (ii) solving the equivalent network by iterative methods.

2.8. Method of Lines (MOL)

The method of lines was introduced into the EM community in the 1980s [19, 20]. It is regarded as a special FDM but more effective with respect to accuracy and computational time than the regular DFM. Originally, MOL was developed for problems with closed solution domain, but absorbing boundary conditions appropriate for MOL have been investigated [21, 22]. The MOL is a differential-difference approach of solving elliptic, parabolic, and hyperbolic partial differential equations and thus, involves discretizing a given differential equation in one or two dimensions while using analytical solution in the remaining direction. The MOL has the merits of both the FDM and the analytical method; it does not yield spurious modes nor does it have the problem of "relative convergence." Besides, the MOL has the following properties that justify its use: (i) Computational efficiency: the semi analytical character of the formulation leads to a simple and compact algorithm that yields accurate results with less computational effort than other techniques. (ii) Numerical stability: by separating discretization of space and time, it is easy to establish stability and convergence. (iii) Reduced programming effort: by making use of the state-of-the-art well documented and reliable ordinary differential equations solvers, programming effort can be substantially reduced. (iv) Reduced computational time: since only a small amount of discretization lines are necessary in the computation, there is no need to solve a large system of equations.

To apply MOL usually involves the following five basic steps: (i) partitioning the solution region into layers (ii) discretization of the differential equation, (iii) transformation to obtain decoupled ordinary differential equations, (iv) inverse transformation and introduction of the boundary conditions, (v) solution of the equations.

2.9. Artificial Neural Networks (ANN)

The recent exploitation of iteratively refined surrogates of fine, accurate or high-fidelity models, and the implementation of space mapping (SM) methodologies address this issue of neural-based approaches for solving EM-related problems [23]. ANN modeling is one of the most recent trends in microwave CAD. Fast, accurate and reliable neural network models can be trained from measured or simulated data. Once developed, these neural models can be used in place of CPU-intensive physics/EM models of devices to speed up microwave design. A full review of ANN applications in RF/microwave modeling and design is found in [24–26].

3. COMPARISON BETWEEN EM NUMERICAL METHODS

As mentioned earlier, there are two groups of numerical methods, namely, differential and integral methods. The major advantage of the differential methods is their adaptability to various complex structures. However, larger is the complexity of these structures, dense should be the grid. On the other hand, integral methods are well adapted to the programming on microcomputers. Each numerical method has advantages and disadvantages. Various aspects of numerical methods are qualitatively compared in Table 1, but in practice, there is no clear-cut boundary assigned between such aspects, since an experienced designer can often improve a numerical processing. Model intended to characterize planar microwave structures must satisfy the schedule of the analyzed structure and obey to certain severe constraints whose effectiveness depends on these models. Among these constraints: short CPU time, compactness memory and a good precision. The choice of an optimal method must carry out a good compromise between these requirements.

Table 1: Comparison of the main numerical methods of passive millimeter-wave structures. *E: excellent, G: good, L: large, Ma: Marginal, M: moderate, S: Small.*

EM Method	Storage requirement	CPU time	Generality	Preprocessing
Finite difference (FDM)	L	L	E	-
Finite element (FEM)	L	M/L	E	S
Transmission line matrix (TLM)	M/L	M/L	E	S
Integral equation	S/M	S/M	G	M
Transverse resonance method	S/M	S/M	Ma	M
Method of lines (MOL)	M	S	G	M/L
Spectral domain (SDA)	S	S	Ma	L
Neural networks (ANN)	S	S	M	L

REFERENCES

1. Bhat, B. and S. K. Koul, "Unified approach to solve a class of strip and microstrip-like transmission lines," *IEEE Trans. Microwave Theory Tech.*, Vol. 30, No. 5, 679–686, 1982.
2. Yamashita, E. and R. Mittra, "Variational method for the analysis of microstrip lines," *IEEE Trans. Microwave Theory Tech.*, Vol. 16, No. 4, 251–256, 1968.
3. Itoh, T. and R. Mittra, "Spectral-domain approach for calculating the dispersion characteristic of microstrip line," *IEEE Trans. Microwave Theory Tech.*, Vol. 21, No. 7, 498–499, 1973.
4. He, M. and X. Xu, "Closed-form solutions for analysis of cylindrically conformal microstrip antennas with arbitrary radii," *IEEE Trans. Antennas Propag.*, Vol. 53, No. 1, 518–525, 2005.
5. Mirshekar-Syahkal, D., *Spectral Domain Method for Microwave Integrated Circuits*, Research Studies Press Ltd., Somerset, England, 1990.
6. Wexler, A., "Solution of waveguide discontinuities by modal analysis," *IEEE Trans. Microwave Theory Tech.*, Vol. 9, No. 4, 508–517, 1967.
7. Bao, G. and W. Zhang, "An improved mode-matching method for large cavities," *Antennas Wireless Propag. Lett.*, Vol. 4, 393–396, 2005.
8. Sadiku, M. N. O., "Finite difference solution of electrodynamic problems," *Int. Jour. Elect. Engr. Educ.*, Vol. 28, No. 4, 107–122, 1991.
9. Zscheile, H., F. J. Schmuckles, and W. Heinrich, "Finite-difference formulation accounting for field singularities," *IEEE Trans. Microwave Theory Tech.*, Vol. 54, No. 5, 2000–2010, 2006.
10. Gwarek, W. K., "Analysis of an arbitrarily-shaped planar circuit — a time domain approach," *IEEE Trans. Microwave Theory Tech.*, Vol. 33, No. 10, 1067–1072, 1985.
11. Harrington, R. F., "Matrix methods for field problems," *Proc. IEEE*, Vol. 55, No. 2, 136–149, 1967.
12. Ney, M. M., "Method of moments as applied to electromagnetics problems," *IEEE Trans. Microwave Theory Tech.*, Vol. 33, No. 10, 972–980, 1985.
13. Adams, A. T., "An introduction to the method of moments," Syracuse Univ., Report TR-73-217, Vol. 1, 1974.
14. Desai, C. S. and J. F. Abel, *Introduction to the Finite Element Method: A Numerical Approach for Engineering Analysis*, Van Nostrand Reinhold, New York, 1972.
15. Lu, C. and B. Shanker, "Solving boundary value problems using the generalized (partition of unity) finite element method," *IEEE Antennas Propag. Int. Symp.*, 125–128, July 2005.
16. Hoefer, W. J. R., "The transmission-line matrix method-theory and applications," *IEEE Trans. Microwave Theory Tech.*, Vol. 33, No. 10, 882–893, 1985.
17. Christopoulos, C., *The Transmission-Line Modeling Method (TLM)*, IEEE Press, New York, 1995.
18. Cabeceira, A. C. L., A. Grande, I. Barba, and J. Represa, "A time-domain modeling for EM wave propagation in bi-isotropic media based on the TLM method," *IEEE Trans. Microwave Theory Tech.*, Vol. 54, No. 6, 2780–2789, 2006.
19. Pregla, R. and W. Pascher, "The method of lines," T. Itoh(ed.), *Numerical Techniques for Microwave and Millimeter-wave Passive Structures*, 381–446, John Wiley, New York, 1989.
20. Dreher, A. and T. Rother, "New aspects of the method of lines," *IEEE Micro. Guided Wave Lett.*, Vol. 11, 408–410, 1995.

21. Pregla, R. and L. Vietzorreck, "Combination of the source method with absorbing boundary conditions in the method of lines," *IEEE Micro. Guided Wave Lett.*, Vol. 5, No. 7, 227–229, 1995.
22. Wu, K. and X. Jiang, "The use of absorbing boundary conditions in the method of lines," *IEEE Micro. Guided Wave Lett.*, Vol. 6, No. 5, 212–214, 1996.
23. Bandler, J. W., R. M. Biernacki, S. H. Chen, P. A. Grobelny, and R. H. Hemmers, "Space mapping technique for electromagnetic optimization," *IEEE Trans. Microwave Theory Tech.*, Vol. 42, No. 12, 2536–2544, 1994.
24. Rayas-Sanchez, J. E., "EM-based optimization of microwave circuits using artificial neural networks: the state-of-the-art," *IEEE Trans. Microwave Theory Tech.*, Vol. 52, No. 1, 420–435, 2004.
25. Zhang, Q. J. and K. C. Gupta, *Neural Networks for RF and Microwave Design*, Artech House, Norwood, 2000.
26. Rahouyi, E. B., J. Hinojosa, and J. Garrigos, "Neuro-fuzzy modeling techniques for microwave components," *IEEE Microwave Wireless Components Lett.*, Vol. 16, No. 2, 72–74, 2006.

CAD of Left-handed Transmission Line Bandpass Filters

L. Zhu, V. K. Devabhaktuni, and C. Wang

Department of ECE, Concordia University
1455 de Maisonneuve West, Montreal H3G1M8, Canada

Abstract— Owing to the negative and nonlinear nature of phase constant (β) versus frequency, left-handed transmission lines (LH-TLs) can present higher values of β as compared to right-handed ones. As such, left-handed (LH) structures are promising in terms of size minimization of microwave circuits. In this paper, a simulation study leading to computer aided design (CAD) of LH-TL bandpass filters is presented. Motivated by a recent work, in which, a bandpass filter is realized by coupling several LH units, we propose a new structure consisting of a single unit with composite right/left-handed (CRLH) characteristics. The structure offers several advantages over traditional filters, e. g., compact dimensions and low loss. A new CAD algorithm for automated filter design, which can be valuable to designers, is demonstrated through an example.

1. INTRODUCTION

Modern microwave circuits and systems require a large number of high-quality passive components. Multiple factors including cost and size can place stringent constraints on the choice of materials and on the complexity of technologies to be used in the fabrication of commercial products. Ever-increasing demand for "lower cost" and "higher performance" has led to the constant exploration of new materials [1] and/or new structures [2]. In the microwave community, theoretical study of left-handed materials (LHM), which simultaneously exhibit negative permeability and permittivity, has been ongoing [3]. Hypothesis predicting the existence of left-handed (LH) mediums was experimentally verified in 2001 [4] and several researchers have applied LH techniques to design RF/microwave circuits [5]. Among other innovative structures, LHM based bandpass filters have been studied [6, 7]. Typically, LH structures are promising both in terms of "reduced size" and "wider bandwidth" as compared to traditional right-handed RF/microwave structures.

In this paper, a brief overview and simulation study of left-handed transmission line (LH-TL) bandpass filters is presented. Motivated by a recent work, we propose a compact bandpass filter structure exploiting composite right/left handed transmission line (CRLH-TL) concepts. As can be seen in Section 2, this structure offers considerable reduction of size and dramatic increase in bandwidth. Physical dimensions of the filter, i. e., design variables, can be adjusted to meet the given specifications. Based on simulations, a new CAD algorithm for fully-automated design of the filter is presented in Section 3. The algorithm begins with the initial physical dimensions. Enriched by the tuning knowledge base compiled as part of our study, the physical dimensions are then adjusted iteratively. At the end of each iteration, the algorithm checks if the specifications have been met, and continues or terminates accordingly. Finally, Section 4 contains concluding remarks.

2. LH-TL BANDPASS FILTER STRUCTURES

2.1. Overview

The electrical length of a transmission medium can be calculated as a product of its physical length and phase constant (β). For a given electrical length, if a high-β transmission medium can be

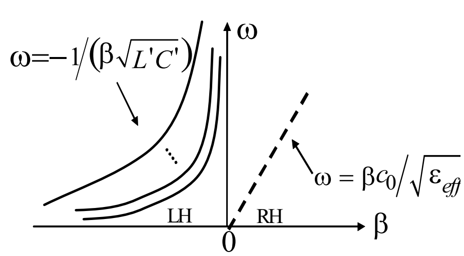


Figure 1: Phase constant β of LH-TL and RH-TL as functions of ω for the lossless case. Here, the dashed line shows the linear nature of β of RH-TL and solid lines show the nonlinear nature of β of LH-TL for different values of L' and C'.

used, the "real-estate needs" of the resulting component are bound to be smaller than traditional mediums. The wave number (γ) of a LH-TL in terms of per-unit-length impedance (Z') and admittance (Y') is given by

$$\gamma(\omega) = \alpha(\omega) + j\beta(\omega) = \sqrt{Z'Y'} = \sqrt{(R' + j\omega L')^{-1}\,(G' + j\omega C')^{-1}}, \tag{1}$$

where R', L', C' and G' are per-unit-length quantities. For the lossless case, $\beta(\omega) = -1/\omega\sqrt{L'C'}$. As can be seen in Fig. 1, higher values of $|\beta|$ can be attained using LH-TL at certain frequencies and for certain $L'C'$ values. Microwave components realized using LH-TL structures can hence lead to effective size minimization.

2.2. Coupling LH Bandpass Filter

In this sub-section, a recent LH microstrip bandpass filter [8] is reviewed. Two inter-digital capacitors in series shown in Fig. 2(a) are considered as a single unit of LH-TL. The outer arms of these capacitors whose edges are grounded by vias act as inductors. The substrate height is 0.38 mm and ε_r is 9.8. EM simulations performed using Zeland's *IE3D* and presented in Fig. 2(b) show a highpass behavior. A LH bandpass filter is then realized by coupling two afore-mentioned units as shown in Fig. 2(c). The simulation results are presented in Fig. 2(d). The filter has a centre frequency of 2.24 GHz and a 3 dB bandwidth of 350 MHz. It has a flat pass-band and low return-loss. However, the structure is observed to have a limited bandwidth i. e., in the MHz range. If the bandwidth specification were to be in the GHz range, this structure would not be suitable.

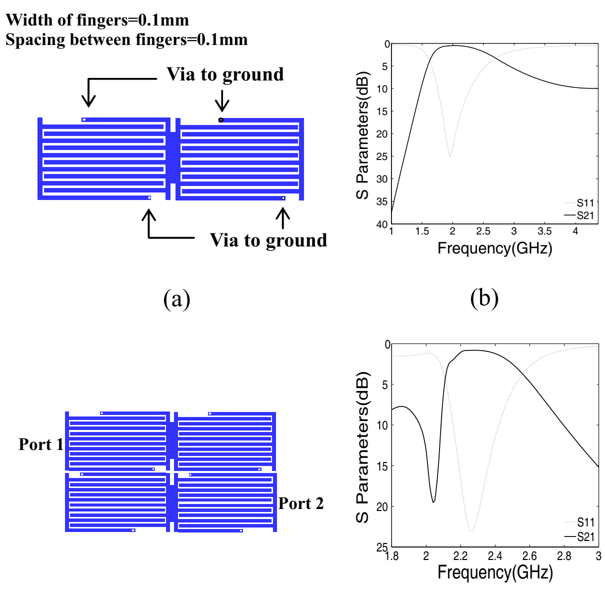


Figure 2: (a) Layout of a single unit of the LH-TL, (b) simulated S-parameters of the unit, (c) layout of the coupling LH filter, and (d) its simulated S-parameters.

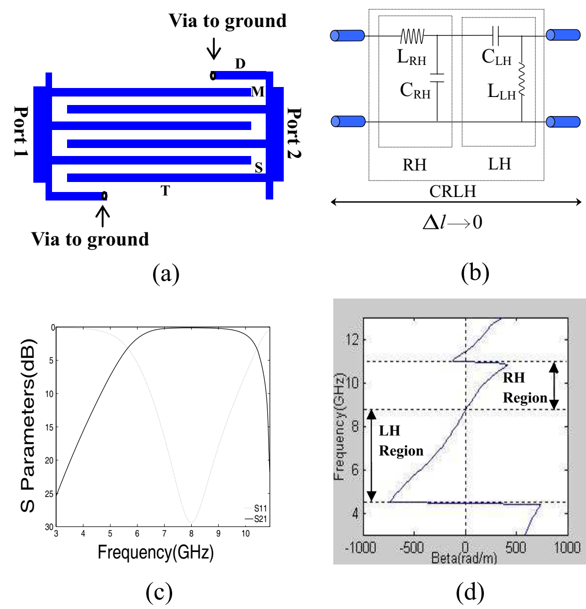


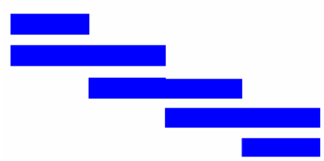
Figure 3: (a) Layout of the CRLH-TL filter with $D = 1.1$ mm, $T = 3.925$ mm, $M = 0.1$ mm, $N = 6$ and $S = 0.1$ mm, (b) infinitesimal LC model of the lossless CRLH-TL, (c) simulated S-parameters of the CRLH filter, and (d) ω-β diagram resulting from $\varphi(S_{21})$.

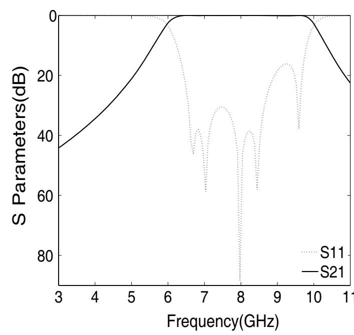
2.3. CRLH Bandpass Filter

Any LH-TL can be treated as a CRLH-TL, due to unavoidable parasitic series inductance and shunt capacitance that lead to a right-handed (RH) contribution, which increases with frequency ω [5]. Motivated by this idea, a compact section of Fig. 2(a) shown in Fig. 3(a) is considered. It has been observed that by adjusting its physical dimensions, i. e., length (D) of outer arms, length (T), width (M), number of fingers (N) and gap between fingers (S), this compact structure can be made to realize bandpass characteristics. The substrate height is 0.254 mm and ε_r is 2.2. The EM simulation results of the structure in Fig. 3(a), whose lumped model is shown in Fig. 3(b), are presented in Fig. 3(c). These results show bandpass characteristics with wide bandwidth and low loss. Fig. 3(d) confirms that the structure is indeed a CRLH-TL. As can be inferred from Table 1 and Fig. 4, this simple yet interesting structure leads to considerable real-estate savings as compared to traditional ones.

Table 1: Size comparison.

Filter	Size
Traditional Microstrip Hairpin Filter	$15.5\times6.5\,\mathrm{mm}^2$
Traditional Coupled Microstrip RH-TL Filter	$28.5\times0.61\,\mathrm{mm}^2$
CRLH Bandpass Filter	$4.23\times1.5\,\mathrm{mm}^2$





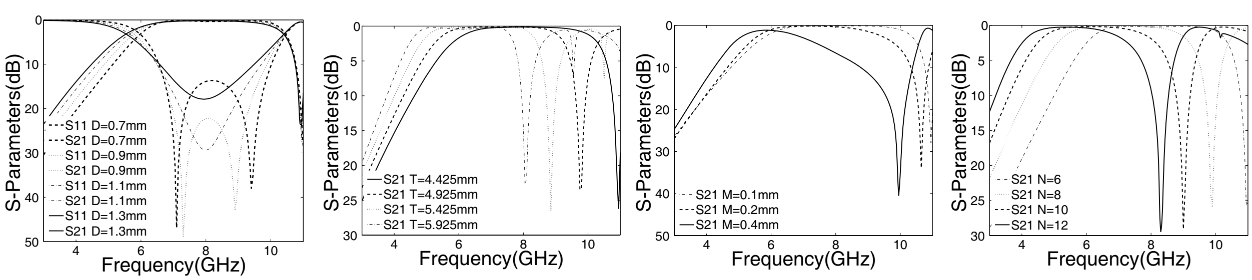
(a) (b)

Figure 4: (a) Layout of the traditional coupled microstrip RH-TL filter and (b) its simulated S-parameters.

3. CAD METHODOLOGY FOR THE CRLH FILTER

3.1. Analysis of Simulation Results

As it is the case with design of other modern RF/microwave circuits, CAD of CRLH filters is of potential interest to designers. As a first step toward developing a fully-automated CAD tool, a simulation study has been carried out. The design parameters have been varied (or swept) and their effect on various design specifications has been examined. Some of the results of the study are shown in Fig. 5. In Fig. 5, each design parameter is varied, while keeping the other parameter values



(a) (b) (c) (d)

Figure 5: (a)–(d) Simulated S-parameters of Fig. 3(a) with different values of D, T, M and N.

constant. Consequently, the overall effect of each design parameter on the filter's specifications is clearly brought out.

Design/tuning of RF/microwave circuits is multi-dimensional and hence complex. Several strategies have been employed in the CAD area to address this challenge [9]. In this work, sensitivity information has been analyzed and compiled into a knowledge base shown in Table 2. Such knowledge identifies the design parameter(s), which highly influence each of the filters specifications. For instance, increasing N results in a negative shift in centre frequency. In Table 2, NC stands for "negligible change".

Table 2: Knowledge base for the CAD methodology.

Tuning Action	**Centre Frequency**	**0.5 dB Bandwidth**	**Edge Frequencies**	
Change D ↑ / ↓	NC	NC	NC	NC
Change T ↑ / ↓	↓ / ↑ (slight change)	NC	↓ / ↑	↓ / ↑
Change M ↑ / ↓	↓ / ↑	↓ / ↑	NC	↓ / ↑
Change N ↑ / ↓	↓ / ↑ (slight change)	NC	↓ / ↑	↓ / ↑
Note: Maximum pass-band attenuation and minimum stop-band attenuation are observed when "impedance matching" is satisfied.				

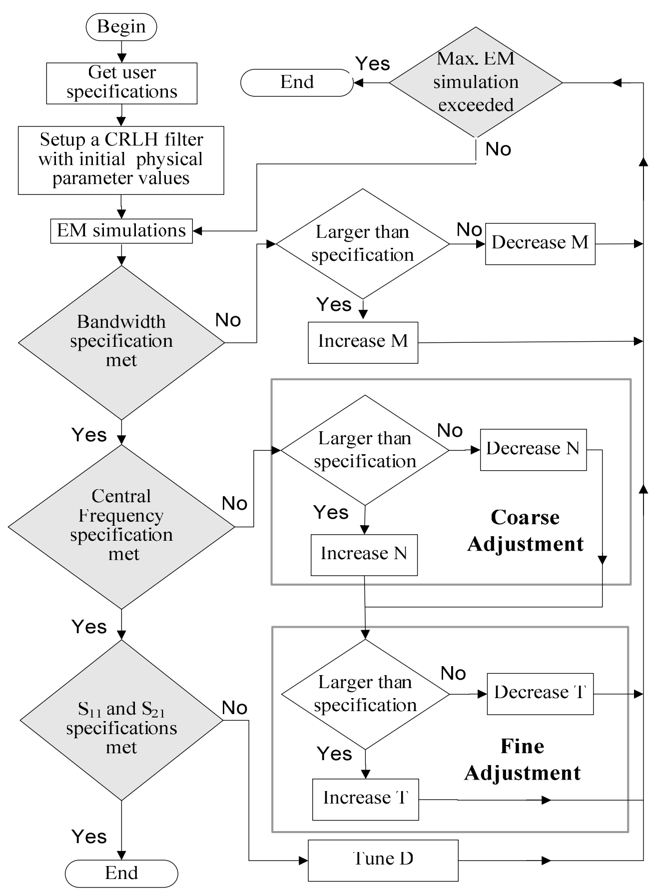


Figure 6: Flow-chart of the routine that forms a basis for the proposed CAD methodology.

3.2. Design Algorithm

Based on the above knowledge base, a CAD algorithm shown in Fig. 6 has been developed for automated design of CRLH filers.

3.3. Design Example

In this subsection, an example of filter design using the proposed CAD methodology is presented. Given user-specifications are $A_{\max} = 0.1\,\text{dB}$, $A_{\min} = 25\,\text{dB}$, $f_c = 6\,\text{GHz}$ and 3 dB bandwidth $BW = 2.5\text{GHz}$.

Step 1: An initial IE3D simulation is performed with initial design parameter values. The width of the passband is almost zero as may be seen in Fig. 7(a).

Step 2: Based on the flow-chart, the width of the passband can be increased by decreasing M, and this is a relatively simpler 1-dimensional design problem. Gradually decreasing M from 0.8 mm to 0.2 mm has resulted in a BW, which is closer to the given specification (see Fig. 7(b)).

Step 3: The centre frequency (currently 7.5 GHz) can be "improved", i. e., decreased, by increasing N. Changing N from 6 to 8 has resulted in $f_c = 6.25\,\text{GHz}$, which is closer to the given specification (see Fig. 7(c)). This change is categorized as a coarse adjustment.

Step 4: A fine adjustment is performed by increasing T from 4 mm to 4.2 mm such that the specification $f_c = 6\,\text{GHz}$ has been precisely met (see Fig. 7(d)).

Step 5: In order to meet the desired values of attenuation, i. e., S_{11} and S_{21}, D has been adjusted from 2 mm to 0.49 mm for impedance matching. As seen in Fig. 7(e), all the specifications have been met.

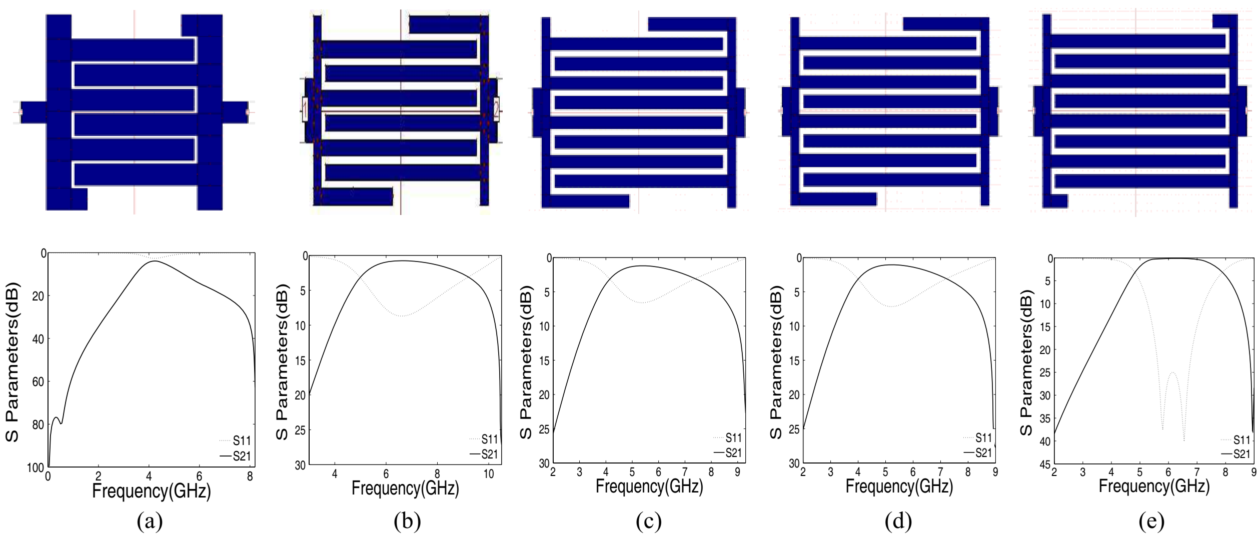


Figure 7: (a) Step 1 with $D = 2\,\text{mm}$, $T = 4\,\text{mm}$, $M = 0.8\,\text{mm}$, $N = 6$, (b) Step 2 with $D = 2\,\text{mm}$, $T = 4\,\text{mm}$, $M = 0.2\,\text{mm}$, $N = 6$, (c) Step 3 with $D = 2\,\text{mm}$, $T = 4\,\text{mm}$, $M = 0.2\,\text{mm}$, $N = 8$, (d) Step 4 with $D = 2\text{mm}$, $T = 4.2\,\text{mm}$, $M = 0.2\,\text{mm}$, $N = 8$, and (e) Step 5 with $D = 0.49\text{mm}$, $T = 4.2\,\text{mm}$, $M = 0.2\,\text{mm}$, $N = 8$.

4. CONCLUSIONS

A simulation study of the recent LH-TL bandpass filters has been presented. Results of the study have been systematically incorporated into a CAD algorithm for automated design of CRLH filter starting from user-specifications. The proposed algorithm has been illustrated through a practical example. Future work will strive to develop a CAD tool and to study time-domain applications of the CRLH filter, e. g., LH delay-lines.

REFERENCES

1. Gunasekaran, M., "A simplified low-cost materials approach to shielding in EMC applications," *Proc. Int. Conf. EM Compatibility*, 58–61, Washington, DC, August 1990.
2. Barnwell, P. and J. Wood, "Fabrication of low cost microwave circuits and structures using an advanced thick film technology," *Proc. IEMT/IMC Symp.*, 327–332, Tokyo, Japan, April 1998.
3. Veselago, V. G., "The electrodynamics of substance with simultaneously negative values of ε and μ," *Soviet Physics Uspekhi*, Vol. 10, 509–514, 1968.
4. Shelby, R. A., D. R. Smith, and S. Schultz, "Experimental verification of a negative index of refraction," *Science*, Vol. 292, 77–79, 2001.

5. Caloz, C. and T. Itoh, "Transmission line approach of left-handed (LH) materials and microstrip implementation of an artificial LH transmission line," *IEEE Trans. Antennas Propagation*, Vol. 52, 1159–1166, 2004.
6. Wei, T. and Z. R. Hu, "Left-handed multilayer super compact bandpass filter," *Proc. APMC*, Suzhou, China, Dec. 2005.
7. Bonache, J., L. Gill, J. Garcia, and F. Martin, "Novel microstrip bandpass filters based on complementary split-ring resonators," *IEEE TMTT*, Vol. 54, 265–271, 2006.
8. Zhu, L., Q. Zhu, C. Chen, and J. Zhang, "Bandpass filter with micro-strip LH transmission line structure," *Proc. APMC*, New Delhi, India, Dec. 2004.
9. Yamini, A. H. and V. K. Devabhaktuni, "CAD of dual-mode elliptic filters exploiting segmentation," *Proc. EuMC*, 560–563, Manchester, UK, Sept. 2006.

Optimal Shape Design of the Electromagnetic Devices in a Level Set Based Implicit Moving Boundary Framework

Jijun Xiao
Institute of Electrical Engineering, Chinese Academy of Sciences, China

Liu Yang
China Agricultural University, China

Abstract— In this paper we present a new framework to approach the problem of Electromagnetic shape optimization. We use a level set method with an implicit moving interface model. Such level set models are flexible in handling complex shape changes. Furthermore, by using a simple Hamilton-jacobi equation, the movement of the implicit interface is driven by a transformation of objective and the constraints into a speed function that defines the level set propagation. The numerical result shows the validity of the present technique.

1. INTRODUCTION

In the shape design of electromagnetic devices using the sensitivity analysis based on the finite element method, the choice of design variables has an effect on success of optimal design. Boundary-based design variables has been a major choice for the shape optimization, In essence, the design domain is represented by a set of design variables directly control the boundary, for example, through the control points of B-splines. This technique, however, has several drawbacks including the need of expensive re-parameterization in case of topological changes.

The level set method, which has several advantages, was instigated by Osher and Sethian for numerically tracking fronts and free boundaries [1], and recently introduced in the field of shape optimization. First, its main feature is to enable an accurate description of the boundaries on a fixed mesh. Therefore it leads to fast numerical algorithms. Second, its range of application is very wide, since the front velocity can be derived from the classical shape sensitivity. Finally, it can handle some kinds of topology changes, and can split and merge as necessary in the course of deformation without the need for re-parameterization. Because of these advantages, level set methods are becoming a powerful mathematical tool to deal with shape and topology optimization problems. This paper employs the level set method for the shape optimization of electromagnetic devices. the boundary of the structure is embedded in a scalar function of a higher dimensional. The dynamic change of the boundary is governed by a partial differential equation of Hamilton-jacobi type, thus, the shape optimization is described as a solution of the Hamilton-jacobi equation. To show validity of presented technique, a magnet shape design for generating uniform magnetic field is tested.

2. SHAPE OPTIMIZATION PROBLEM

In this paper we use a magnetostatic structure to describe the problem of shape optimization in Electromagnetic Devices. The shape design problems of electromagnetic devices always result in designing the material interface boundary between different materials as depicted in Fig. 1. Consider two different reluctivity in two dimensional space, where Ω^+ and Ω^- denote the two regions, Γ is the boundary of Ω^+, γ is the interface of Ω^+ and Ω^-. Let $\Gamma = \Gamma^0 + \Gamma^1$ be fixed and interface γ be varied, the general problem of magnetostatic shape optimization is specified as

$$Minimize \qquad F = \int_{\Omega_+} f\,(A_1) m_f d\Omega + \int_{\Omega_+} g\,(H_1) m_g d\Omega + \int_{\gamma} h\,(A_1) m_h d\gamma \tag{1}$$

$$Subject\ to \quad -\nabla \times (v\nabla \times A - M) + J = 0, \quad \text{in } \Omega = \Omega^+ \cup \Omega^- \tag{2}$$

$$A = c, \quad \text{on } \Gamma^0 \; {\partial A}/{\partial n} = 0 \text{ on } \Gamma^1 \tag{3}$$

where A and H are the magnetic vector potential and magnetic field intensity, ν is the reluctivity, M is the permanent magnetization and J is the current density, The symbols m_f, m_g, and m_h denote the characteristic functions that represent parts of the analysis space Ω and γ.

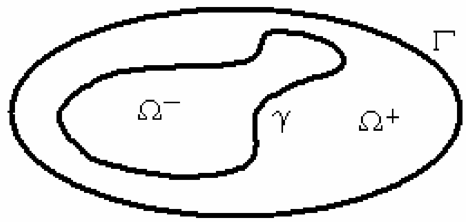


Figure 1: Shape optimization problem.

3. THE LEVEL-SET MODEL OF SHAPE OPTIMIZATION

The key feature of the level set approach is to represent domains and their boundaries not via parameterizations, but as level sets of a continuous function Φ, the so-called level set function [1].

$$\begin{aligned} &\Phi(x,\,t) > 0 \qquad x \in \Omega^+ \\ &\Phi(x,\,t) < 0 \qquad x \in \Omega^- \\ &\Phi(x,\,t) = 0 \qquad x \in \gamma \end{aligned} \tag{4}$$

The implicit function $\Phi(x)$ is used to represent the boundary and to optimize it, as it was originally developed for curve and surface evolution. The change of the implicit function $\Phi(x)$ is governed by the simple Hamilton-jacobi equation

$$\frac{\partial \Phi\left(x,\,t\right)}{\partial t} + \nabla \Phi\left(x,\,t\right) \cdot V\left(x\right) = 0 \tag{5}$$

where $V(x)$ defines the "velocity" of each point on the boundary. Since the tangential components of V would vanish, it can be written as

$$\frac{\partial \Phi\left(x,\,t\right)}{\partial t} + \left|\nabla \phi\left(x,\,t\right)\right| V_n\left(x\right) = 0 \tag{6}$$

where V_n defines the normal velocity of each point on the boundary. In this dynamic level set model, the shape optimization process can be viewed as follows. Let V_n be the normal movement of a point on a surface driven by the objective of the optimization, such that it can be expressed in terms of the position and the geometry of the surface at that point. Then, the optimal structural boundary is expressed as a solution of a partial differential equation, When the steady state of this equation is achieved (i.e., $V_n = 0$), the optimality condition is also achieved and, hence, an optimal shape of the structure is obtained. This is the well-known gradient descent method, where the velocity V_n, which is to be chosen plays an important role.

4. VELOCITY CHOICE

The velocity controls the scheme of the optimization in the level set model and an appropriate choice is thus essential.. our approach presented here is to bridge the well-established methods of shape sensitivity analysis with the powerful methods of level sets to fulfill our goal of general shape optimization within the implicit boundary framework.. Sensitivity is expressed as the total derivative of the objective function to design parameters. it can be developed using continuum approach. The material derivative concept of continuum mechanics and the adjoint variable method are used to express the relation between the change of shape and that of objective function. the general sensitivity formula in a linear magnetostatic system is expressed as [2]

$$\frac{dF}{dp} = \int\limits_{\gamma} \left\{ B\left(\lambda^{**}\right)^T \left[\left(\nu^* - \nu^{**}\right) B\left(A^*\right) + \left(M^{**} - M^*\right)\right] + \lambda^{**}\left(J^{**} - J^*\right) \right\} n^T \frac{\partial X}{\partial p} d\gamma \tag{7}$$

B is the magnetic flux density vector, p the shape design parameter, n the normal vector on the interface γ, X the position on the interface, * and ** mean respectively region Ω^+ and Ω^-, λ is the adjoint variable which can be described in the adjiont system as [3]:

$$-\nabla \times \left(\nu \nabla \times \lambda - g_1 m_g\right) + f_1 m_f = 0 \qquad \text{in } \Omega = \Omega^+ \cup \Omega^- \tag{8}$$

$$\lambda = 0, \quad \text{on } \Gamma^0, \qquad \frac{\partial \lambda}{\partial n} = 0 \quad \text{on } \Gamma^1 \tag{9}$$

$$n \times \left(\nu^{**} \nabla \times \lambda^{**} - \nu^* \nabla \times \lambda^*\right) = -h_1 m_h \qquad \text{on } \gamma \tag{10}$$

where $g_1 = {}^{\partial g}/_{\partial H}$, $f_1 = {}^{\partial f}/_{\partial A}$ and $h_1 = {}^{\partial h}/_{\partial A}$, discrete the equation, we can get the normal velocity at any point on the interface as

$$V_n = B\left(\lambda^{**}\right)^T \left[\left(\nu^* - \nu^{**}\right) B\left(A^*\right) + \left(M^{**} - M^*\right)\right] + \lambda^{**}\left(J^{**} - J^*\right) \tag{11}$$

5. NUMERICAL PROCESS

In this section, we will give some useful hints for the numerical implementation of such an algorithm.

5.1. Initialization

The first step in this iterative process is to define an initial shape of the object. This can be done by choosing the domain Ω_0 and deduce the level set $\phi(0, x)$, The choice of this level set is quite free, we define the level set as the oriented distance function [1]:

$$\phi(0, x) = \begin{cases} -dist(x, \partial\Omega_0) \text{ if } x \in \Omega_0 \\ +dist(x, \partial\Omega_0) \text{ if x} \notin \Omega_0 \end{cases} \tag{12}$$

As the interface evolves, φ will generally drift away from its initialized value as signed distance, in order to keep φ as signed distance function, we can solve a Hamilton-jacobi equation for its steady state [1],

$$\frac{\partial\phi}{\partial t} = sign(\phi)\left(1 - |\nabla\phi|\right) \tag{13}$$

where $sign(\phi)$ is a sign function taken as 1 in Ω^+, -1 in Ω^-, and 0 on the interface.

5.2. Field Calculation

At each step, we have to compute a direct field A and adjiont field λ, both problem can be solved by finite element method. So the normal velocity at any point on the interface can be deduced by using Equation (11). In the level set formulation, we need the normal velocity V_n in a neighborhood of the interface γ, the most natural way to extend the normal velocity off the interface is to let the V_n be constant along the normal to γ such that

$$\nabla V_n \cdot \nabla\phi = 0 \tag{14}$$

This lead to the following Hamilton-jacobi Equation [1]

$$\frac{\partial V_n}{\partial t} + sign(\phi)\frac{\nabla\phi}{|\nabla\phi|} \cdot \nabla V_n = 0 \tag{15}$$

5.3. Discrete Computation Schemes [1]

The discrete solution to the Hamilton-jacobi equation is computed using finite differences over discrete time steps $\Delta t = h_t$ and on the discrete grid $\Delta x = h_x$, based on the notion of weak solutions and entropy limits, a so called "up-wind scheme" is proposed to solve with first order update equation. Higher order schemes can also be obtain for discrete approximation.

5.4. Algorithm

Let us briefly summarize the algorithm:

1. Choose the initial domain Ω_0 and define the level set function through the oriented distance function;
2. Iteration until convergence, for $k \geq 0$:
   (a) computation of the direct field A and adjiont field λ in Ω_k by the finite element method, according the Equation (11),compute the normal velocity V_n on the interface.
   (b) extend the normal velocity off the interface by solving the Hamilton-jacobi Equation (15)
   (c) Deformation of the shape by solving the Hamilton-jacobi Equation (6), the new shape Ω_{k+1} is characterized by the level set function ϕ_{k+1} solution of (6) after a time step Δt_k, the time step Δt_k is chosen such that $F(\Omega_{k+1}) < F(\Omega_k)$.
3. From time to time, for stability reasons, we also reinitialize the level set function φ by solving (13).

6. NUMERICAL EXAMPLE

Numerical example presented here for electromagnetic shape optimization have been studied in the relevant literature [4], the initial magnet shape is show in Fig. 2. The aim of this design is a uniform magnetic flux density in the air gap (region B in Fig. 2) The objective function is defined as

$$F = \sum_{i=1}^{nn} (B_i - B_{0i})^2 \tag{16}$$

where B_i is the calculated magnetic flux density at the i-th point, B_{oi} is the target value, nn is the number of measure points.

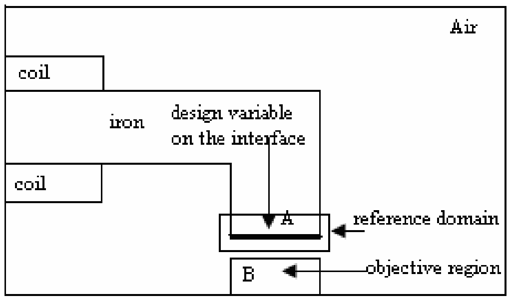


Figure 2: Initial shape of a magnet design model.

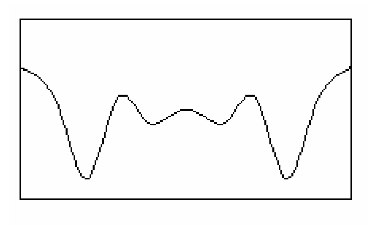


Figure 3: Design shape in surface A in Fig. 2.

The design variable is the interface A of iron and air, Choose the reference domain Ω_0 involve the interface. The reference domain is divided into square cells of 81×41, the interface is embedded in the reference domain by the level set function. As the optimization process, the interface A will be changed as motion defined by the Hamilton-jacobi equation until a termination as illustrated in Fig. 3. The convergence trend of the objective function is illustrated in Fig. 4.

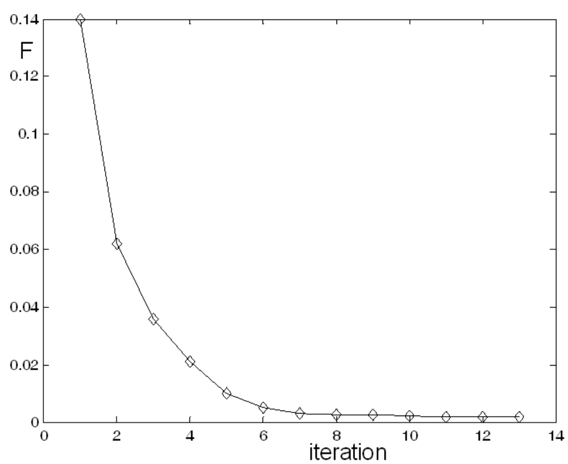


Figure 4: Convergence trends of object function values.

7. CONCLUSION

We have presented a level set based method for the shape optimization of electromagnetic devices. This method leads naturally to a dynamic framework of a Hamilton-jacobi equation governing motions of the level set with flexibility of handling shape changes. We have established a relationship between the velocity field in the Hamilton-jacobi equation to the shape sensitivity analysis. This relationship justifies a proper choice of the velocity field for an optimization process. The numerical result shows the validity of the present technique.

ACKNOWLEDGMENT

This work was Supported by National Natural Science Foundation of China (NSFC) (No. 50377043). Author Xiao Jijun also thanks Professor Zhang Yiming for fruitful discussions.

REFERENCES

1. Osher, S. and R. Fedkiw, *Level Set Methods and Dynamic Implicit Surfaces*, Springer-Verlag, New York, 2003.
2. Park, II-H., J. L. Colomb, and S.-Y. Hahn, "Implementation of continuum sensitivity analysis with existing finite element code," *IEEE Transactions on Magnets*, Vol. 29, No. 2, 1787–1790, 1993.
3. Kim, D.-H. and K. S. Ship, "Applying continuum design sensitivity analysis combined with standard EM software to shape optimization in magnetostatic problems," *IEEE Transactions on Magnets*, Vol. 40, No. 2, 1156–1159, 2004.
4. Kim, C.-H. and H.-B. Lee, "B-spline parametrization of finite element models for optimal design of electromagnetic devices," *IEEE Transactions on Magnets*, Vol. 35, No. 5, 3763–3765, 1999.

Optimal Model for Wiggly Coupled Microstrips in Directional Coupler and Schiffman Phase Shifter

Ming-Wei Zhou, Li Li, and Qiu-Yan Yin
East China Normal University, China

Abstract— A fast way to optimize wiggly coupled microstrips to reach high directivity is proposed, which makes it easy to fabricate a multi-section high directivity wideband coupler and Schiffman wideband phase shifter using wiggly coupled microstrip.

1. INTRODUCTION

High isolation Directional couplers and phase shifters are important modules in microwave test system, electrically scanned array and many other RF applications. Microstrip is a usual structure for its easy fabrication and low cost. But as well known, the inhomogeneous dielectric material of coupled microstrip causes the different odd and even phase velocity, which makes directivity and isolation of coupler deteriorate quickly. Several methods have been reported to overcome or utilize the difference of the velocity [1–6]. Some can get high isolation only on one frequency point [3–5], which is useless in wideband application. Some can ultilize the difference of the velocity to get high isolation called codirectional coupler [6]. But they can only be realized at very high frequency (e.g., millimeter-wave). An effective way to get a high isolation on wideband in microwave frequency is to use a wiggly coupled microstrip [1, 2]. Traditionally, the wiggly coupled microstrip is designed mainly based on semi-empiricism [2], so it's not widely used for its hard to get desired goal. Though Uysal derived design equations for the optimum dimensions of the wiggly coupler [7], the emulation of electromagnetism and physical experiment show it cannot get a satisfying result as Podell [1].In this paper, we will give some improved optimization program to make these designs more easily and quickly.

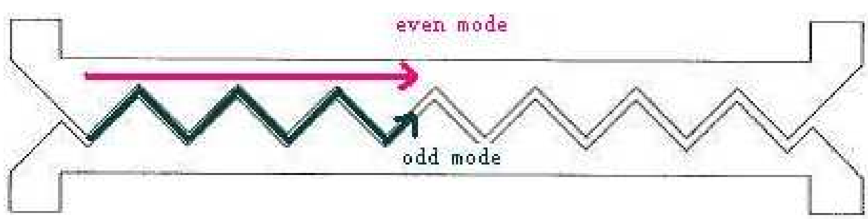


Figure 1: Classical wiggly coupler.

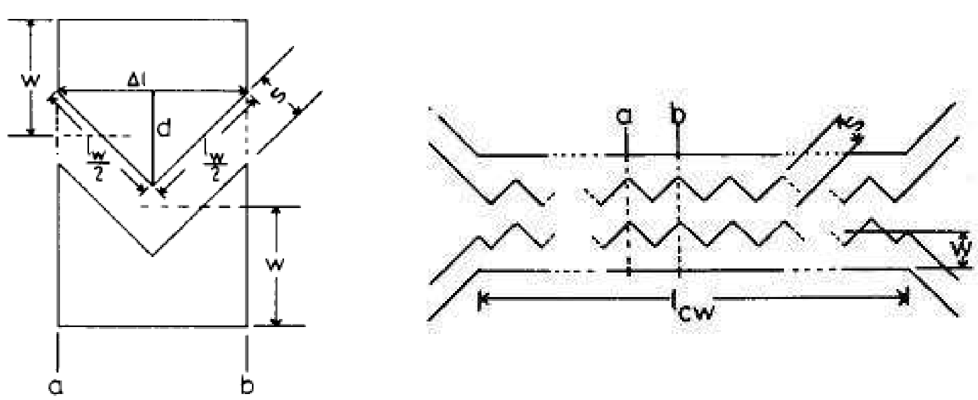


Figure 2: Uysal's calculation.

2. PRINCIPLE AND PROGRAMMING

For general coupled microstrip, the directivity of 90 degree microstrip deteriorate for different odd and even mode phase velocity, which are deviate from each other according following equations: $v_{po} = c/\sqrt{\varepsilon_{eo}}$, $v_{pe} = c/\sqrt{\varepsilon_{ee}}$, where c is velocity of light, ε_{eo} and ε_{ee} are effective dielectric constant. For v and $\sqrt{\varepsilon}$ are inversely proportional, the key of this paper is to find a way to equalize ε_{eo} and ε_{ee} fast and accurately. As a classical way to compensate the phase velocity distinctly, the wiggly microstrip add the length of odd mode in zigzag path, which can be thought as decreasing the odd phase velocity.

To confirm the wiggly scope and microstrip dimension, in Uysal's paper [7], two important equations are reported:

$$d = \frac{\Delta l}{2}\sqrt{(\frac{C'_{fo}}{C_{fo}})^2 - 1} \tag{1}$$

where d is the wiggle depth, C_{fo} is the capacitance between the coupled microstrip, C'_{fo} is the desired capacitance. Δl is the length of the wiggly teeth.

Another equation gives the shortened microstrip length for the increasing of even-mode capacitance:

$$l_{2c} = l_{cw}\sqrt{\frac{C_{pf} + C_{fe}}{C_{pf} + C'_{fe}}} \tag{2}$$

where l_{2c} is the exact length of the coupler, l_{cw} is the original length, C_{pf} is the capacitance of microstrip to ground and C_{fe} is the inner capacitance of coupled microstrip to ground. The detailed description appears in [7]. But electromagnetism emulation software (Em of Sonnet and Momentum2005a of Agilent) and practical test show these couplers doesn't give a good isolation as Podell's. It's caused by the impedance change with the increasing of the odd and even capacitance. Because of so many variables (such as microstrip length, width, wiggly teeth, height and space between lines) to be optimized, the computation requirement will be exponential increased as the increase of the numbers of the vairables. Ordinary PC and work station cannot provide such computation ability. So an easy optimizing program is suggested to replace the slow EM optimization partly. If optimizing program calculate the impedance as even mode effective dielectric constant, better result could be got. Detailed procedure shows as follows:

1. set center frequence, board dielectric constant, thickness and metal thickness t.
2. calculate desired odd and even mode impedance, which can be got through coupling coefficient.
3. Strip thickness correction for finite thin microstrip line thickness [8]
4. calculate ordinary single microstrip quasi-static effective dielectric constant [9].
5. calculate the effective dielectric constant with dispersion [10]
6. calculate the characteristic impedance with dispersion.

   $$Z_L(f_n) = Z_L(0) \cdot (\frac{R_{13}}{R_{14}})^{R_{17}}$$

   where the detailed description of $Z(0)$, R_{13}, R_{14} and R_{17} are reported in [10].
7. calculate even mode effective dielectric constant of coupled microstrip $\varepsilon_{eff,e}$ [11].
8. calculate the characteristic impedance of coupled microstrip [11].
9. set a initial coupled dimension of coupled strip width w and spacing s.
10. give a gradation optimizing grogram to find the proper w and s matching the desired odd and even mode impedance, where odd mode effective dielectric constant equals $\varepsilon_{eff,e}$ calculated above.
11. load Uysal's equations to get the exact wiggly teeth length l and height d.
12. em simulation and final adjustment

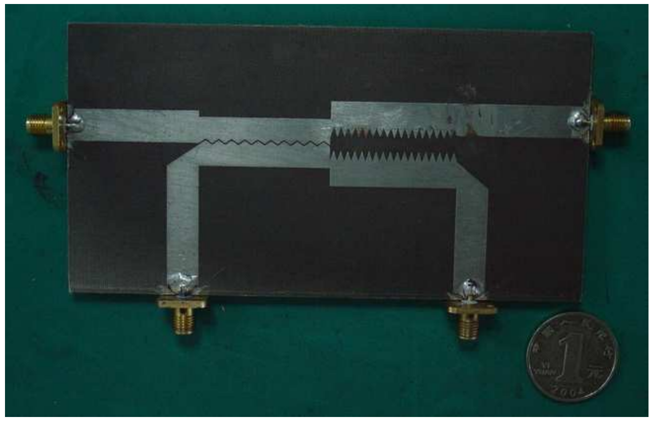


Figure 3: Photo of 1–2 GHz 10 dB coupler.

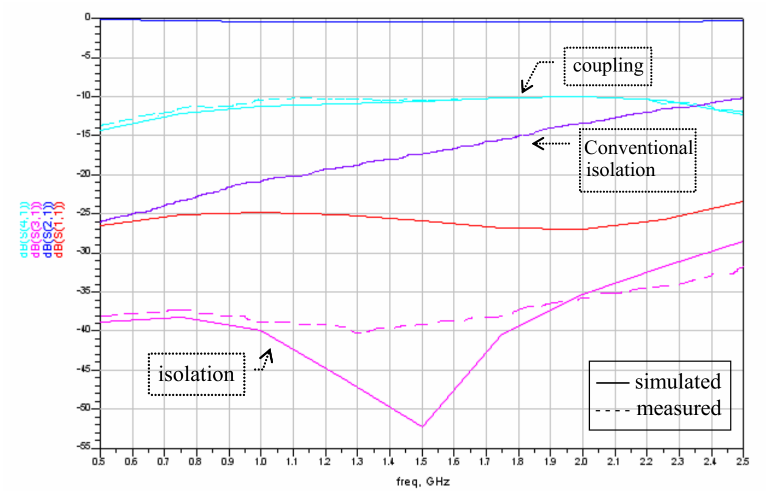


Figure 4: 10 dB coupler simulated and measured results.

3. SIMULATIONS AND MEASUREMENT

To demonstrate the validity of the proposed method, a 10 dB coupler in 1 GHz to 2 GHz and a Schiffman wideband phase shifter (using coupler microstrip) in 1 to 2 GHz are designed and fabricated by previous program.

The first coupler uses ordinary PTFE material with relative dielectric constant 2.6, 3 mm height and 35 um metal thickness.

The results show in Fig. 4 is better than ever got by Podell [1].

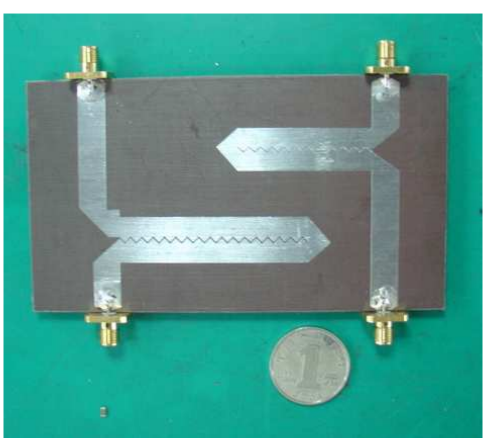


Figure 5: Double Schiffman phase shifter.

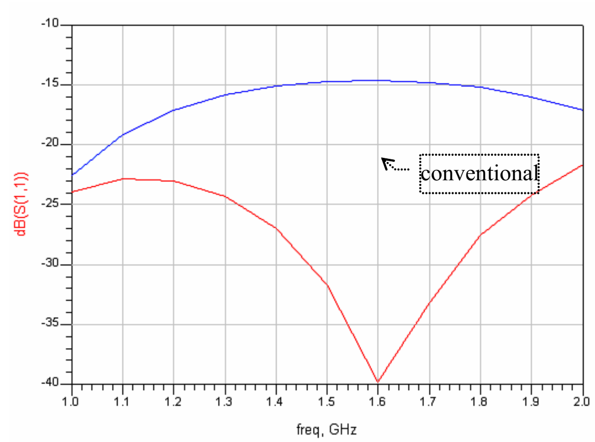


Figure 6: Return loss of wiggly microstrip phase shifter.

The second example is a wideband phase shifter utlize the special phase response of coupled transmission line element [12]. The difficulty for the coupled microstrip is also the different phase velocity causing the poor return loss [13]. Though it's reported wiggly microstrip doesn't give great improvement on return loss [13], the grogram given above offers a chance to change the condition.

4. CONCLUSION

The improvement of wiggly coupled microstrip makes it possible for the design of high directivity and good return loss. It is almost impossible to optimize the wiggly coupled microstrip for so many variables to be modified such as w, l, s, d, etc. The wiggly microstrip transmission line model offers a fast and accurate way to get the dimensions for actual purpose.

REFERENCES

1. Podell, A., "A high directivity microstrip coupler technique," *1970 MTT-S Dig.*, 33–36, May 1970.
2. Taylor, J. L. and D. D. Prigel, "Wiggly phase shifters and directional couplers for radio-frequency hybrid-microcircuit applications," *IEEE Trans. Parts, Hybirds and Packaging*, Vol. PHP-12, No. 4, Dec. 1976.
3. March, S. L., "Phase velocity compensation in parallel-coupled microstrip," *IEEE MTT-S Symp. Dig.*, 410–412, June 1982.
4. Chang, S.-F., J.-L. Chen, Y.-H. Jeng, and C.-T. Wu, "New high-directivity coupler design with coupled spurlines," *IEEE Microwave and Wireless Components Letter*, Vol. 14, No. 2, Feb. 2004.
5. Wang, S.-M., C.-H. Chen, and C.-Y. Chang, "A study of meandered microstrip coupler with high directivity," *2003 IEEE m-S Digest.*
6. Uysal, S. and J. Watkins, "Novel microstrip multifunction directional couplers and filters for microwave and millimeter-wave applications," *IEEE Trans on Microwave Theory and Techniques*, Vol. 39, No. 6, June 1991.
7. Uysal, S. and H. Aghvami, "Synthesis, design, and construction of ultra-wide-band nonuniform quadrature directional couplers in inhomogeneous media," *IEEE Trans. on Microwave Theory and Techniques*, Vol. 37, No. 6, June 1989.
8. Schneider, M. V., "Microstrip lines for microwave integrated circuits," *The Bell System Technical Journal*, Vol. 48, 1421–1444, May 1969.
9. Hammerstad, E. and O. Jensen, "Accurate models for microstrip computer-aided design," *Symposium on Microwave Theory and Techniques*, 407–409, June 1980.
10. Kirschning, M. and R. H. Jansen, "Accurate model for effective dielectric constant of microstrip with validity up to millimeter-wave frequencies," *Electronics Letters*, Vol. 8, No. 6, 272–273, Mar. 1982.
11. Kirschning, M. and R. H. Jansen, "Accurate wide-range design equations for the frequency-dependent characteristic of parallel coupled microstrip lines," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 32, No. 1, 83–90, Jan. 1984.
12. Schiffman, B. M., "A new class of broad-band microwave 90-degree phase shifters," *IRE Trans. Microwave Theory Tech.*, 232–237, Apr. 1958.
13. Schiek, B. and K. Jurgen, "A method for broad-band matching of microstrip differential phase shiftersr," *IEEE Trans. Microwave Theory Tech.*, Vol. 25, No. 8, 666–671, Aug. 1977.

A Simple Technique for Efficient Computation of Electromagnetic Coupling in Microwave Integrated Circuits

Dave McPhee, **Sashieka Seneviratne**, **and Mustapha C. E. Yagoub**

SITE, University of Ottawa, 800 King Edward, Ottawa K1N 6N5, Ontario, Canada

Abstract— Because of ever-higher operating frequencies and circuit integration, electromagnetic (EM) coupling effects are becoming increasingly important in microwave integrated circuits. Existing EM-based numerical methods can generate accurate models for passives but suffer from high computation time and memory requirements, while fast table look-up models and equivalent circuit models do not adequately include distributed EM coupling between adjacent components. Using simple network theory principles and de-embedding concepts, the proposed technique allows fast and efficient computation of such coupling, making the design more reliable.

1. INTRODUCTION

Parasitic electromagnetic (EM) coupling is becoming more of a concern in microwave integrated circuits (MIC) as packing densities of components are increased and frequencies are pushed higher [1]. Although such unintended coupling between components plays an important rule in the circuit performance, this quantity is very complex to evaluate. In the recent years, several techniques in bringing this forward into the circuit design space has been investigated [2–7]. Mainly based on the resolution of Maxwell's equations to quantify the EM field in a structure, these numerical methods generated advanced fullwave EM simulation tools [8, 9] that have demonstrated their efficiency in terms of accuracy, but still require huge computation time and memory space. This aspect is crucial when modern CAD tools lead to massive and highly repetitive computational tasks during simulation, optimization and statistical analysis. As such, development of fast and full EM representations with high-order coupling is crucial for modern circuit design. Furthermore, existing passive EM-based models suffer on an important lack at the circuit level. In fact, even if such device models are accurate, they are developed in a perfect shielded environment, i. e., excluding any external effects such as coupling from neighboring components.

This paper introduces a fast and accurate technique that efficiently integrates interactions between two circuit elements in microwave circuit simulators. Using basic network theory concepts and de-embedding techniques [10], this original approach allows efficient evaluation of parasitic coupling during design. Applications using EM and circuit commercial simulators are presented.

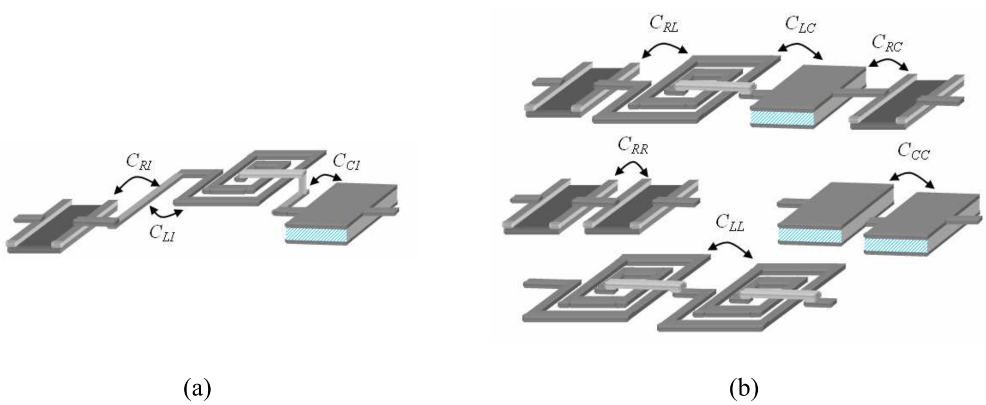


Figure 1: EM Couplings: (a) between Interconnect and Component, (b) Between Components. Subscripts R, L, C and I Refer to Resistors, Inductors, Capacitors, and Interconnects Respectively.

2. PROPOSED APPROACH

As shown in Figure 1, we defined different couplings according to the passives we considered. Then, we used network properties to incorporate such couplings into a circuit design cycle. For illustration,

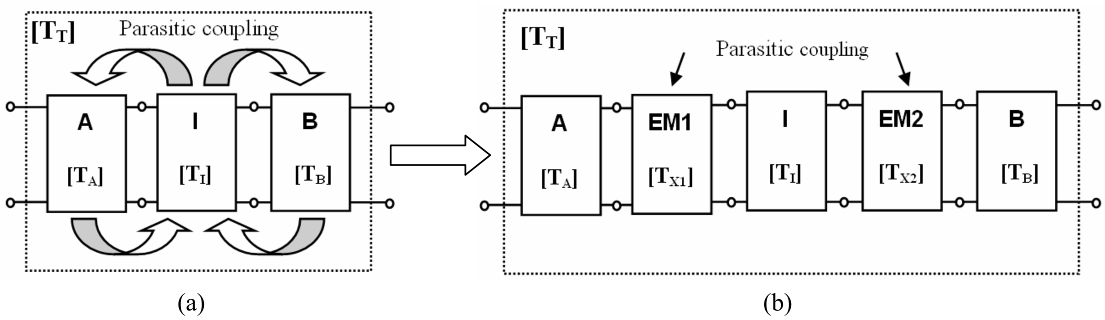


Figure 2: Parasitic coupling between series elements: (a) Original circuit. I is the interconnect of matrix $[\mathrm{T_I}]$ while $[\mathrm{T_A}]$ and $[\mathrm{T_B}]$ are the matrices of sub-networks A and B, respectively, (b) Modified topology, replacing the coupling by two-port networks of matrices $[\mathrm{T_{X1}}]$ and $[\mathrm{T_{X2}}]$ in series with the sub-networks A and B.

let us consider two devices A and B connected in series via an interconnect I (Figure 2(a)) and let $[\mathbf{T_T}]$ be the network response

$$[\mathbf{T_T}] = [\mathbf{T_A}] * [\mathbf{T_I}] * [\mathbf{T_B}] \tag{1}$$

Introducing coupling will lead to the modified circuit shown in Figure 2(b).

Thus, the procedure for characterizing the sub-networks EM1 and EM2 in the modified network was to compare the following T-matrices

$$[\mathbf{T_{T\ without\ coupling}}] = [\mathbf{T_A}] * [\mathbf{T_I}] * [\mathbf{T_B}] \tag{2}$$

$$[\mathbf{T_{T\ with\ coupling}}] = [\mathbf{T_A}] * [\mathbf{T_{X1}}] * [\mathbf{T_I}] * [\mathbf{T_{X2}}] * [\mathbf{T_B}] \tag{3}$$

The first matrix $[\mathbf{T_{T\ without\ coupling}}]$ could be obtained through individual simulation or measurements of the individual matrices $[\mathbf{T_A}]$, $[\mathbf{T_I}]$, and $[\mathbf{T_B}]$, while the second matrix $[\mathbf{T_{T\ with\ coupling}}]$ was obtained by evaluating the overall network response. A general procedure to extract the coupling between two components can be stated as:

Step 1 Simulate/measure the individual components including interconnects to obtain $[\mathbf{T_A}]$, $[\mathbf{T_I}]$, and $[\mathbf{T_B}]$.

Step 2 Use an EM simulator to obtain $[\mathbf{T_{AI}}]$ and $[\mathbf{T_{IB}}]$

$$[\mathbf{T_{AI}}] = [\mathbf{T_A}] * [\mathbf{T_I}] \tag{4}$$

$$[\mathbf{T_{IB}}] = [\mathbf{T_I}] * [\mathbf{T_B}] \tag{5}$$

Step 3 Repeat step 2 for all two-by-two combinations as defined in Figure 1.

Step 4 Measure all two-by-two combinations of the components to obtain $[\mathbf{T_{ACI}}]$ and $[\mathbf{T_{ICB}}]$ defined as

$$[\mathbf{T_{ACI}}] = [\mathbf{T_A}] * [\mathbf{T_{X1}}] * [\mathbf{T_I}] \tag{6}$$

$$[\mathbf{T_{ICB}}] = [\mathbf{T_I}] * [\mathbf{T_{X2}}] * [\mathbf{T_B}] \tag{7}$$

Step 5 Measure $[\mathbf{T_{T\,with\,coupling}}]$.

Step 6 Use (2) to get $[\mathbf{T_{T\,without\,coupling}}]$.

Step 7 Use (1)–(7) to compute $[\mathbf{T_{X1}}]$ and $[\mathbf{T_{X2}}]$.

This simple de-embedding technique was extended first to a three-port network (Figure 3) by implementing the connection matrix technique [11] into a symbolic computation tool [12], and then, generalizing to an N-port network [13]. Steps 1 to 4 were achieved *off-line* to create a library of T-matrices, while steps 5 to 7 depend on the device under test. Fundamentally, this library of all two-by-two component combinations was a set of neural models of passives generated by different EM simulators [8, 9] and then plugged into a circuit simulator [14]. The neural inputs were the passive geometrical/electrical parameters while the outputs are the S-parameters for a 1–80 GHz frequency range [5]. The library could be updated depending on the circuit element parameters.

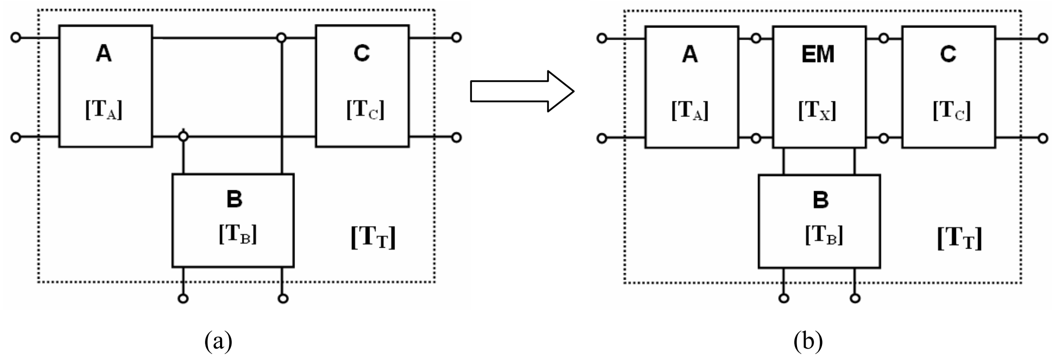


Figure 3: T-Junction network: (a) Original configuration with three sub-networks A, B, and C, (b) Modified circuit including the three-port coupling sub-network EM.

3. EXAMPLES

3.1. Series RLC Circuit

The first example is a simple series RLC circuit (Figure 4). As shown in Figure 5, our approach allowed obtaining very close results with original data generated in a 3D EM-simulator, i. e., HFSS [9]. The speedup in terms of computation time is significant, since our simulation was achieved in 2s while the required time in the 3D EM-simulator was 2860s.

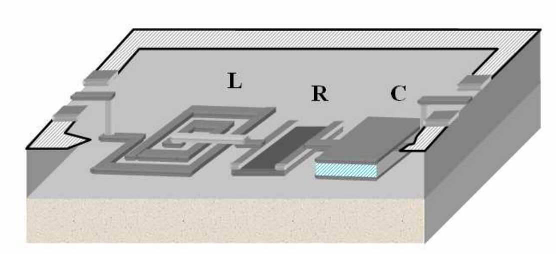


Figure 4: Series RLC Circuit.

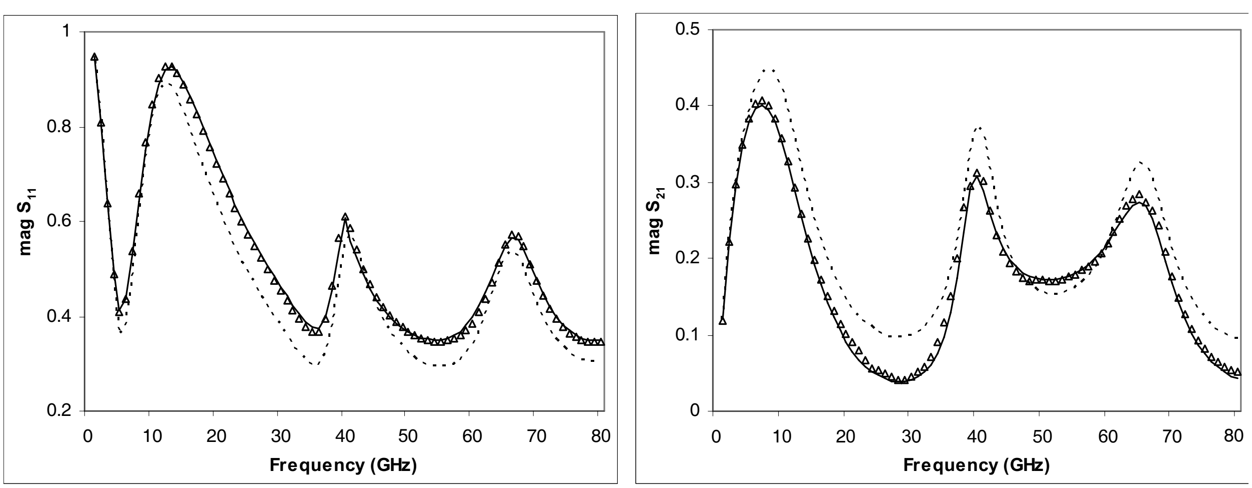


Figure 5: Series RLC circuit: Comparison of S_{11} and S_{21} magnitudes obtained with (—) and without coupling (– – –) with those simulated in Ansoft (Δ).

3.2. Chebyshev Band Pass Filter

To furthermore investigate the coupling effects, a more complex circuit, i. e., a 4th order band pass filter, was designed. As expected, the results showed a close agreement with the original ones (Figure 6) with a significant reduction of the computation time ratio by 1/6000.

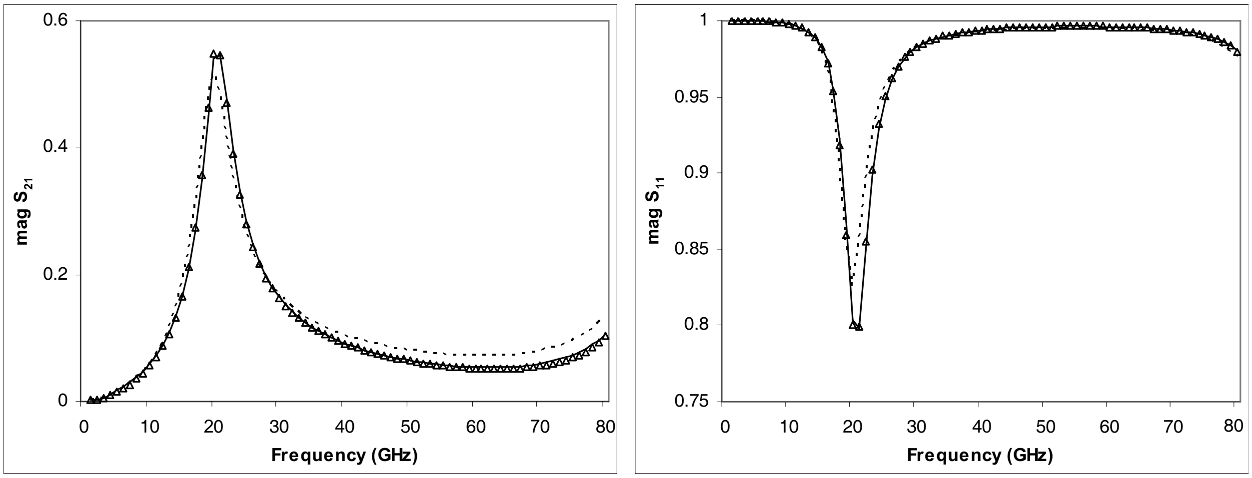


Figure 6: BandPass Filter: Comparison of S_{11} and S_{21} magnitude and angle obtained with (—) and without coupling (– – –) with those simulated in Ansoft (Δ).

4. DISCUSSION

The major advantages of the above method are accuracy, speediness, and simplicity. First, all neural models of passives were generated from accurate EM simulators, making the model response practically as accurate as the EM simulator outputs (the final testing errors of the different neural models never exceeded 2% while the testing error were less than 3% [5]). Second, all T-matrix manipulations were derived using symbolic computation, excluding any approximation or simplification. Third, the codes can be implemented easily in any circuit simulator. Finally, even if the offline neural model generation took significant time, the on-line simulation of the above circuits, i.e., the user time, demonstrated the speediness of our approach in terms of computation time. In fact, the neural model generation was performed in one single step, making the trained models available in form of an internal library inside the circuit simulator for any circuit simulations and analyses.

5. CONCLUSION

An efficient approach for automatic modeling of parasitic coupling has been presented. Applied to various microwave integrated circuits, the proposed technique demonstrated its efficiency in terms of speediness and ease of usage. It helps making the design of microwave integrated circuits faster, more accurate and efficient, contributing to overall reductions in design cycles. The library of neural network models and the N-ports T-parameters internal codes can be subsequently used for the simulation and optimization of more complex microwave circuits.

REFERENCES

1. Dunn, J. M., L. C. Howard, and K. Larson, "An efficient algorithm for the calculation of parasitic coupling between lines in MIC's," *IEEE Trans. Microwave Theory Tech.*, Vol. 41, 1287–1293, 1993.
2. Goldfarb, M. and A. Platzker, "The effects of electromagnetic coupling on MMIC design," *Microwave Millimeter Wave CAE*, Vol. 1, 38–47, 1991.
3. Rizzoli, V., A. Costanzo, D. Masotti, A. Lipparini, and F. Mastri, "Computer-aided optimization of nonlinear microwave circuits with the aid of electromagnetic simulation," *IEEE Trans. Microwave Theory Tech.*, Vol. 52, 362–377, 2004.
4. Ding, X., B. Chattaraj, M. C. E. Yagoub, V. K. Devabhaktuni, and Q. J. Zhang, "EM based statistical design of microwave circuits using neural models," *Int. Symp. on Microwave and Optical Technology*, 421–426, Montreal, QC, Canada, 2001.
5. Yagoub, M. C. E. and P. Sharma, "Characterization of EM effects in RF/microwave integrated circuits," *European Microwave Conf.*, 221–224, Amsterdam, Netherlands, 2004.
6. Baudrand, H., "Electromagnetic study of coupling between active and passive circuits," *Int. Microwave and Optoelectronics Conf.*, 143–152, Natal, Brazil, 1997.
7. Centeno, A., "A comparison of numerical dispersion in FDTD and TLM algorithms," *Asia-Pacific Conf. on Applied Electromagnetics*, 128-131, 2003.

8. *Sonnet-EM*, Sonnet Software Inc., Liverpool, NY.
9. *Ansoft-HFSS*, Ansoft Corp., Pittsburg, PA.
10. Chen, C. and M. Deen, "A general noise and s-parameter deembedding procedure for on-wafer high-frequency noise measurements of MOSFETs," *IEEE Trans. Microwave Theory Tech.*, Vol. 49, 1004–1005, 2001.
11. Gupta, K. C., R. Garg, and R. Chadha, *Computer Aided Design of Microwave Circuits*, Artech House, Dedham, 1981.
12. *Maple*, MapleSoft, Waterloo, ON.
13. McPhee, D. and M. C. E. Yagoub, "A generic procedure for exact high frequency device characterization using post-processing data," *Int. J. of Pure and Applied Mathematics*, Vol. 20, 553–573, 2005.
14. *ADS*, Agilent Technologies, Palo Alto, CA.

On Quantum Corrections to Space Charge Waves in Silicon

A. Garcia-B[1], **V. Grimalsky**[1,3], **E. Gutierrez-D.**[2], **and S. Koshevaya**[3]

[1]INAOE, Electronics Department, Luis Enrique Erro # 1, Tonanzintla, Puebla, Mexico
[2]INTEL-SRCM, Guadalajara, México
[3]Autonomous University of Morelos (UAEM), Cuernavaca 62210, Mexico

Abstract— An analysis that incorporates quantum corrections and the thermal conductivity term into the classical hydrodynamic model of the propagation of space-charge waves in silicon is presented. From numerical simulations we have seen that for frequencies $f < 8\,\mathrm{THz}$ the classical hydrodynamic model, with the thermal conductivity term and with quantum corrections, gives good results where the thermoconductivity seems to be more essential for these frequencies. However for higher frequencies $f > 8\,\mathrm{THz}$ both quantum corrections and thermoconductivity are equally important. These results suggest accurate simulations of ultra-small device require the thermal conductivity term to be included in the model.

1. INTRODUCTION

The amplification of EM waves in the microwave range by using space charge waves has been studied in the last few decades [1, 2]. The microwave technology of monolithic integrated/hybrid circuits is been gradually moving into the millimeter wave range, up and above 100 GHz. The development and manufacturing of microwave or millimeter-wave integrated semiconductor devices depends on the development of computer aided design tools, based on the accuracy and the adequacy of mathematical models by solving quantum hydrodynamic equations rigorously. It was shown in [3] that, when the temperature is not too low, two-dimensional electrons in the channel of the FET behave not as a gas (as conventionally expected) but rather as a fluid. Indeed, as one can show, the electron mean free path for collisions with impurities and phonons is much greater than the mean free path for electron-electron collisions. This means that the theoretical description of the electron flow in the FET channel should be based on the equations of hydrodynamics.

A variety of models have been developed for semiconductor device simulation. However, the classical hydrodynamic model (HD) can be extended to include quantum effects by incorporating the quantum corrections, this model is called the quantum hydrodynamic (QHD) [4]. The QHD model is derived from a moment expansion of the Wigner-Botlzman equation, using a quantum Maxwellian distribution to close the moments [5, 6]. The QHD conservation laws have the same form as the classical hydrodynamic equations, but the energy density and stress tensor include additional quantum terms. These quantum corrections allow particle tunneling through potential barriers and to build up a potential well.

In this investigation, we use the QHD transport equations, expressed in terms of conservation laws for particles, momentum, and energy added by the Poissons equation, to analyze the propagation of volume space charge waves in silicon. The linear modes of propagation of a system are studied by means of the dispersion equation $D(\omega, k) = 0$, which relates the frequency ω to the longitudinal wave number (or propagation constant) $k = 2\pi/\lambda$ and it is obtained by the self-consistent solution of basic equations. In general, we consider the cases where ω is real and $k = k' + ik''$ has real and imaginary parts. The case $k'' > 0$ corresponds to the spatial increment (amplification), whereas the case $k'' < 0$ corresponds to the decrement (damping) [7]. To obtain the dispersion equation for space charge waves, it is necessary to use the linearized equations of the electron dynamics jointly with the Poisson's equation for the electric potential.

The amplification of space charge waves is because of the negative differential conductivity as shown in [2, 8] for the case of GaAs structures. However in silicon the negative differential conductivity does not appear, so an amplification is not possible but only the propagation and damping of space charge waves. In this work, we have studied the influence when the quantum corrections and thermal conductivity term are included into the balance equations and what is more essential (in this model): thermoconductivity or quantum corrections. These results are relevant as they explain the influence of quantum corrections and termoconductivity on the propagation of space charge waves in silicon structures. One can see that for frequencies $f < 8\,\mathrm{THz}$ the classical hydrodynamic model with the thermal conductivity term and with quantum corrections gives good results. However, for frequencies $f > 8\,\mathrm{THz}$ both the thermoconductivity and quantum correction terms give essential inputs, and it is difficult to conclude about which correction term is more important. These results can also be used in the simulation of semiconductor devices with ultra short lengths.

2. THE QUANTUM HYDRODYNAMIC MODEL

The QHD model has the same form as the classical hydrodynamic equations only with quantum corrections in the energy density and stress tensor terms. Thus, we employ the following set of balance equations regarding carrier density, average velocity, and average energy for electrons added by the Poisson's equation [9].

$$\frac{\partial n}{\partial t} + \nabla(nv) = 0 \tag{1}$$

$$\frac{\partial v}{\partial t} + (\nabla v)v = \frac{qE}{m^*} - \frac{1}{nm^*}\nabla(nT) - \gamma_p(w)v \tag{2}$$

$$\frac{\partial w}{\partial t} + \left(\vec{v}\vec{\nabla}\right) w = q\vec{E}\vec{v} - \frac{1}{n}\nabla(nvT - \kappa\nabla T) - \gamma_w(w)\,(w - w_0) \tag{3}$$

where n is the electron density, v is the velocity, m^* is the effective electron mass, q is the electronic charge, w is the energy density. $T = (2/3)(w - m^*v^2/2)$, and the thermal conductivity can be approximated as in Ref. [10] $\kappa = (5/2)\left(nT/m^*\gamma_p(w)\right)$, E is the electric field and γ_p and γ_w are the inverse of momentum and energy relaxation rate, respectively.

We take into account the quantum corrections to stress tensor and energy density like in Ref. [11]

$$P_{ij} = -nT\delta_{ij} + \frac{\hbar^2}{4m^*}\Sigma F_\lambda \frac{\partial^2}{\partial x_i \partial x_j}\log\left(F_\lambda\right); \quad W = \frac{1}{2}mv^2 + \frac{3}{2}T - \frac{\hbar^2}{8m^*n}\Sigma F_\lambda \Delta \log\left(F_\lambda\right) \tag{4}$$

where T is the temperature of electron gas and F_λ is the distribution function (for quantum corrections, it is the shifted Maxwellian one). Actually $F_\lambda = f_\lambda \cdot n$, where n is the concentration and λ are the states of electron. Therefore, after some mathematical treatment the quantum corrections can be written as follows:

$$P_{ij} = -nT\delta_{ij} + \frac{\hbar^2}{4m^*}n\frac{\partial^2}{\partial x_i \partial x_j}\log(n); \quad W = \frac{1}{2}mv^2 + \frac{3}{2}T - \frac{\hbar^2}{8m^*}\Delta\log(n) \tag{5}$$

In the one-dimensional case, the equations with quantum corrections are given as follows; Eq. (1) for the carrier density is the same, however the equations for the average velocity and average energy have some differences:

$$\frac{\partial v}{\partial t} + v\frac{\partial v}{\partial z} = \frac{qE}{m^*} - \frac{1}{nm^*}\frac{\partial}{\partial z}\left[nT - \frac{\hbar}{4m^*}n\frac{\partial}{\partial z^2}\log(n)\right] - v\gamma_p(w) \tag{6}$$

$$\frac{\partial w}{\partial t} + v\frac{\partial w}{\partial z} = qEv - \frac{1}{n}\frac{\partial}{\partial z}\left[v\left(nT - \frac{\hbar^2}{4m^*}n\frac{\partial^2}{\partial z^2}\log(n)\right)\right] - \kappa\frac{\partial T}{\partial z} - \frac{\hbar^2}{8m^*}\left(n\frac{\partial^2 v}{\partial z^2}\right) - (w - w_{00})\,\gamma_w(w) \tag{7}$$

We consider below linear space charge waves $n = n_0 + \tilde{n}$; $\vec{v} = v_0 + \tilde{\vec{v}}$; $w = w_0 + \tilde{w}$; $T = T_0 + \tilde{T}$, where all the small perturbations obey the law $\sim \exp(i(\omega t - kz))$, and use the following parameters: drift velocity is $v_0 = 10^7\,\text{cm s}^{-1}$, the bias electric field is $E_0 = 10^7\,\text{V m}^{-1}$, electron concentration is $n_0 = 10^{23}\,\text{m}^{-3}$, and the momentum and energy relaxation rates are $\gamma_p \approx 4 \times 10^{13}\,\text{s}^{-1}$ and$\gamma_w \approx 3 \times 10^{12}\,\text{s}^{-1}$ [9]. The characteristic spatial scale is $l_n = 10^{-8}\,\text{m}$.

The quantum hydrodynamic equations for perturbations are given as

$$\frac{\partial \tilde{n}}{\partial t} + v_0\frac{\partial \tilde{n}}{\partial z} + n_0\frac{\partial \tilde{v}}{\partial z} = 0 \tag{8}$$

$$\frac{\partial \tilde{v}}{\partial t} + \frac{v_0}{3}\frac{\partial \tilde{v}}{\partial z} + \gamma_p\tilde{v} = -\frac{\gamma_4 T_0}{n_0}\left[1 - \frac{\gamma_c}{T_0}\frac{\partial^2}{\partial z^2}\right]\frac{\partial \tilde{n}}{\partial z} + \gamma_2\tilde{E} - \left[\frac{2}{3}\gamma_4\frac{\partial}{\partial z} + v_0\frac{d\gamma_p}{dw}\right]\tilde{w} \tag{9}$$

$$\frac{\partial \tilde{w}}{\partial t} + \frac{5v_0}{3}\frac{\partial \tilde{w}}{\partial z} - \frac{2}{3}\frac{\kappa}{n_0}\frac{\partial^2 \tilde{w}}{\partial z^2} + \gamma_w\tilde{w} = \left[\gamma_1 E_0 - \left(T_0 - \frac{2}{3}\gamma_1 v_0\right)\frac{\partial}{\partial z} - \frac{2}{3}\frac{\kappa\gamma_3 v_0}{n_0}\frac{\partial^2}{\partial z^2} + \frac{3}{4}\gamma_c\frac{\partial^3}{\partial z^3}\right]\tilde{v} - \frac{v_0 T_0}{n_0}\frac{\partial}{\partial z}\left[1 - \frac{\gamma_c}{T_0}\frac{\partial^2}{\partial z^2} - \frac{\kappa\gamma_c}{2n_0 v_0 T_0}\frac{\partial^3}{\partial z^3}\right]\tilde{n} + \gamma_1 v_0\tilde{E} \tag{10}$$

where γ_1, γ_2, γ_3, γ_4 and γ_c are defined as

$$\gamma_1 = \frac{qE_0 l_n}{w_0}; \quad \gamma_2 = \frac{qE_0 l_n}{m^* v_0^2}; \quad \gamma_3 = \frac{\gamma_1}{\gamma_2} = \frac{m^* v_0^2}{w_0}; \quad \gamma_4 = \frac{\gamma_2}{\gamma_1} = \frac{w_0}{m^* v_0^2} \quad \text{and} \quad \gamma_c = \frac{\hbar^2}{6m^* w_0 l_n^2}$$

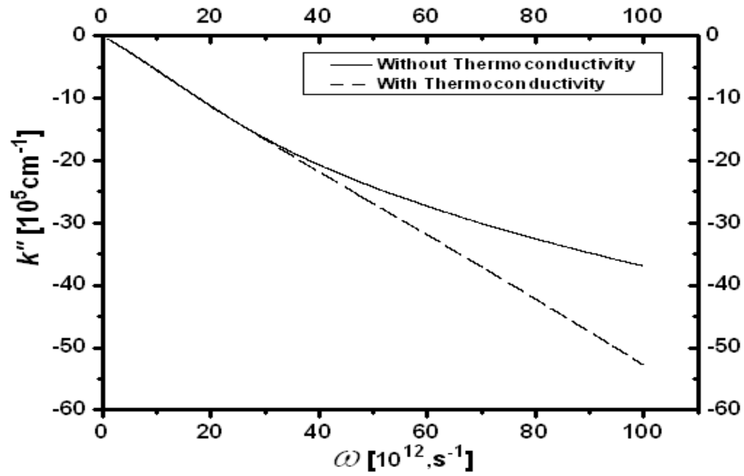


Figure 1: Dependence of spatial decrement of space charge wave on frequency without quantum corrections for two cases; without thermal conductivity term and with thermal conductivity term only.

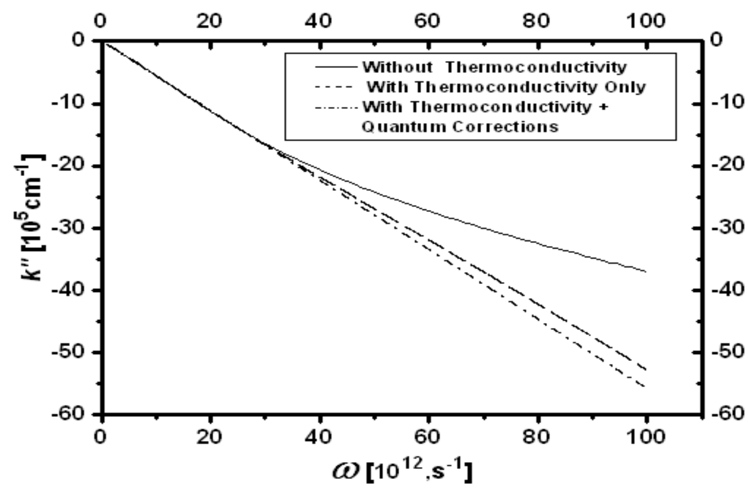


Figure 2: Dependence of spatial decrement of space charge wave on frequency obtained with three different approximations, without thermal conductivity, with thermal conductivity only and with quantum corrections and thermoconductivity jointly.

3. SIMULATIONS AND RESULTS

The dispersion relations $k(\omega)$ have been calculated within the framework of balance equations added by the Poisson's equation. The results of direct simulations of $k''(\omega)$ of linearized equations are shown in the Figures 1 and 2. In Figure 1, the spatial decrement is given for two cases; the classical hydrodynamic equations without the thermal conductivity term and with the thermal conductivity term, one can see the discrepancy when the frequency $\omega > 50 \times 10^{12}\ \mathrm{s}^{-1}$ ($f \sim 8\,\mathrm{THz}$). In the Figure 2 one can see the comparison of the dispersion equation between three different approximations, without thermal conductivity, with thermal conductivity only and with quantum corrections and thermoconductivity jointly. These results are relevant as they explain the influence of quantum corrections and termoconductivity on the propagation of space charge waves in silicon structures. Also the scope of space charge waves applications is not limited to the analysis here described, but can be useful to monolithic phase shifters, delay lines, analog circuits for microwave signals.

4. CONCLUSIONS

The dependencies of the imaginary part $k''(\omega)$ of a complex longitudinal wave number on frequency, which are obtained from balance equations taking into account the thermal conductivity term and quantum corrections, are presented and discussed. For frequencies $f < 8\,\mathrm{THz}$ the thermal conductivity term seems to be more essential than the quantum corrections. For frequencies ($f > 8\,\mathrm{THz}$) both the thermal conductivity term and quantum corrections are important simultaneously.

ACKNOWLEDGMENT

This project has been partially funded by the CONACyT- Mexico (grants 47141, 48955) and by the PhD's scholarship #165390. We also want to acknowledge INTEL-SRCM for providing the time for preparing and reviewing this work. This work is also a result of the partial support of Mexico Federal Government Secretaria de Economía, CECyT Puebla, and COECyTJAL.

REFERENCES

1. Barybin, A. A. and V. M. Prigorovskii, "Waves in thin semiconductor layers with negative differential conductivity," *Russian Physics Journal*, Vol. 24, No. 8, 704–717, August, 1981.
2. Bryksin, V. V., P. Kleinert, and M. P. Petrov, "Theory of space-charge waves in semiconductors with negative differential conductivity," *Physics of the Solid State*, Vol. 45, No. 11, 2044–2052, 2003.
3. Dyakonov, M. and M. Shur, "Shallow water analogy for a ballistic field effect transistor: new mechanism of plasma wave generation by dc current," *Phys. Rev. Lett.*, Vol. 71, 2465–2468, 1993.
4. Gardner, C. L., "The quantum hydrodynamic model for semiconductor devices," *SIAM Journal on Applied Mathematics*, Vol. 54, 409–427, 1994.
5. Gardner, C. L., "Resonant tunneling in the quantum hydrodynamic model," *VLSI Design*, Vol. 3, 201–210, 1995.
6. Wigner, E., "On the quantum correction for thermodynamic equilibrium," *Physical Review*, Vol. 40, 749–759, 1932.
7. Steele, M. C. and B. Vural, *Wave Interactions in Solid State Plasmas*, 236–240, McGraw-Hill, New York, 1969.
8. Garcia-B, A., V. Grimalsky, E. Gutierrez-D, and S. Koshevaya, "Dispersion relation for two-valley quasi-hydrodynamic models in SCW Propagation in n-GaAs thin films," *Proceedings of IEEE-MIEL*, 541–544, Belgrade, Serbia and Montenegro, May 2006.
9. Tomizawa, K., *Numerical Simulation of Submicron Semiconductor Devices*, 186–189, Artech House, London, 1993.
10. Baccarani, G. and M. R Wordeman, "An investigation of steady state velocity overshoot in Silicon," *Solid State Electronics*, Vol. 28, 407–416, 1985.
11. Gardner, C. L. and C. Ringhofer, "The Chapman-Enskog expansion and the quantum hydrodynamic model for semiconductor devices," *VLSI Design*, Vol. 10, 415–435, 2000.

Parametric Models for Electromagnetic Field Systems Related to Passive Integrated Components

Abstract— The paper presents some of the scientific objectives and the achievements of the FP6/IST European research project entitled "*Chameleon RF* — Comprehensive High-Accuracy Modeling of Electromagnetic Effects in complete Nanoscale RF blocks". One of the project's goals is the development of a better Electronic Design Automation — EDA tool dedicated to the signal integrity design verification of passive nano-structures, including interconnects on RFICs. The major contribution of the article is the new, effective methodology for extraction of the parametric models for passive integrated components with field effects, valid for high frequency broad range. The proposed numeric method is systematically based on a dual approach, which provides two complementary approximations of the exact solution. Duality is applied both to the spaces where the discrete solution is found as well as to the open boundary conditions. Adjoint Field Technique — AFT applied to Finite Integral Techniques is used to handle the parameter variability. The new method decreases the necessary computing resources for modeling, comparing with the usual numerical methods. Considering the variability specific to the nowadays nanotechnologies, the fast extraction of parametric models is a must for the present VLSI and RF-IC design environments.

1. INTRODUCTION

With the further downscaling of the IC technology, the operational frequencies of the signals are in the GHz range. As a consequence, the use of lumped-element parameters for the simulation of IC designs is too crude to generate reliable lay-outs. The lumped element models with parameters obtained by static field computation ignore many effects that become pronounced at high frequencies, such as propagation, eddy currents, skin effects, etc. In other words: interconnects and integrated passives (capacitors and spiral coils) should be modeled with Maxwell equations, electromagnetic field playing an essential role in their behavior. The fact that high frequency issues can no longer be ignored in IC design has urged the contributors to the International Technology Roadmap for Semiconductors (ITRS, http://*www.itrs.net)* to declare the high-frequency modeling ($> 5\,\mathrm{GHz}$) as a *grand challenge.* It should be solved in order to continue the pace of progress that was witnessed in the last three decades. Another challenge is related to the high parameter variability, specific to nowadays IC technologies, which requires suitable EDA tools, able to handle them.

To accelerate the design of on-chip structures and interconnects the development of new modeling algorithms is needed. This task was addressed by a European joint research project entitled *Chameleon RF* — "Comprehensive High-Accuracy Modeling of Electromagnetic Effects in complete Nanoscale RF blocks" *(http://www.chameleon-rf.org)* carried out within the Sixth Framework Program, the thematic component Information Society Technologies (IST). The following three principles drive the project:

- In order to capture the high-frequency effects, full wave electromagnetic field is considered.
- Extracted models should be as compact as possible, but computation times should be acceptable.
- Models should describe also the variability of geometric and technological parameters.
- Experimental data provides hardware validation of the developed modeling methodology.

2. ELECTROMAGNETIC FIELD COMPUTATION

The Manhattan geometry, characteristic to IC layout makes Finite Integration Technique — FIT a suitable numerical method for electromagnetic field computation. It is a numerical method able to solve field problems based on spatial discretization "without shape functions". Two staggered orthogonal (Yee type) grids are used as discretization mesh [1]. The degrees of freedom used by FIT are not local field components, but global variables i.e., voltages $\mathbf{u}$, $\mathbf{u_m}$ and fluxes $\boldsymbol{\varphi}$, $\boldsymbol{\psi}$ assigned to grid elements: edges and faces, respectively (Fig. 1). Applying the global form of electromagnetic field equations on mesh elements, a system of DAE is obtained, called Maxwell Grid Equations (MGE):

$$\begin{cases} curl\mathbf{E} = -\dfrac{\partial \mathbf{B}}{\partial t} \\ \Rightarrow div\mathbf{B} = 0 \end{cases} \Rightarrow \begin{cases} \oint \mathbf{E}dr = -\iint \dfrac{\partial B}{\partial t}dA \\ \Rightarrow \oiint \mathbf{B}dA = 0 \end{cases} \Rightarrow \begin{cases} \mathbf{Cu} = -\dfrac{d\varphi}{dt} & (1) \\ \Rightarrow \mathbf{D}'\varphi = 0 & (2) \end{cases}$$

$$\begin{cases} curl\mathbf{H} = \mathbf{J} + \dfrac{\partial \mathbf{D}}{\partial t} \\ \Rightarrow div\mathbf{D} = \rho \end{cases} \Rightarrow \begin{cases} \oint \mathbf{H}dr = \iint (\mathbf{J} + \dfrac{\partial \mathbf{D}}{\partial t})dA \\ \Rightarrow \oiint \mathbf{D}dA = \iiint \rho dv \end{cases} \Rightarrow \begin{cases} \mathbf{C}'\mathbf{u}_m = \mathbf{i} + \dfrac{d\psi}{dt} & (3) \\ \Rightarrow \mathbf{D}\psi = \mathbf{q} & (4) \end{cases}$$

$$\Rightarrow div\mathbf{J} = -\frac{\partial \rho}{\partial t} \Rightarrow \iint \mathbf{J}dA = -\iiint \frac{\partial \rho}{\partial t}dv \Rightarrow \mathbf{Di} = -\frac{d\mathbf{q}}{dt} \tag{5}$$

FIT combines MGE with the Hodge's operators which describe material

$$\mathbf{B} = \mu\mathbf{H},\ \mathbf{D} = \varepsilon\,\mathbf{E},\ \mathbf{J} = \sigma\mathbf{E},\ \Rightarrow \boldsymbol{\varphi} = \mathbf{M}_{\boldsymbol{\mu}}\mathbf{u}_{\mathbf{m}}, \boldsymbol{\psi} = \mathbf{M}_{\boldsymbol{\varepsilon}}\mathbf{u},\ \mathbf{i} = \mathbf{M}_{\boldsymbol{\mu}}\mathbf{u}. \tag{6}$$

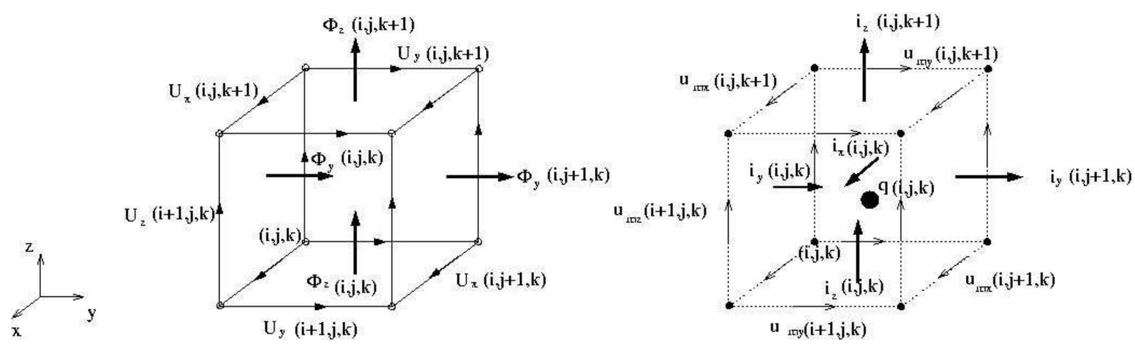


Figure 1: Dofs for FIT numerical method.

The main characteristics of FIT method are:

- There is no discretization errors in MGE fundamental equations;
- MGE are metric-free, sparse, mimetic and conservative system of Differential-Algebraic-Equations (DAE), without spurious modes.

An efficient numerical method called dual Finite Integration Technique (dFIT) was developed as an improved version of FIT [2]. The numerical field problem is solved two times, using both dual staggered grids as graphs of the electric network. The exact lumped port parameters: **R**-resistance, **L**-inductance and **C**-capacitance and the extracted parameters from numeric static solutions (using *p-primary* and *s-secondary* grid) satisfy:

$$\mathbf{R_p} \le \mathbf{R} \le \mathbf{R_s}, \quad \mathbf{L_p} \le \mathbf{L} \le \mathbf{L_s}, \quad \mathbf{C_s} \le \mathbf{C} \le \mathbf{C_p}, \tag{7}$$

(where $\mathbf{A}\le\mathbf{B}$ means $\mathbf{x}^{\mathrm{T}}(\mathbf{A}\text{-}\mathbf{B})\mathbf{x}\le 0$ for any $\mathbf{x}$). The average between s and p accelerates the numeric convergence, preliminary tests showing that typically there are 20 times less DOFs needed to obtain the same solution accuracy as in FIT. dFIT provides not only better accuracy but also it reduces the computational efforts and it provides an effective indicator for the numerical error (the distance between dual solutions).

3. BOUNDARY CONDITIONS

A key aspect of the numerical simulation is the proper boundary conditions for the field problem. The basic formulation of the electromagnetic circuit element assures the field-circuit compatibility [3]. It corresponds to a zero Neumann boundary condition for both electric and magnetic potentials. Formally, we may consider that outside the computational domain the electric and magnetic material constants are zero ($\mu \to 0$, $\varepsilon \to 0$). With these boundary conditions, the system response may be defined in terms of multiport network parameters, e.g., hybrid circuit (transfer) function $\mathbf{H}(\boldsymbol{\omega})$ — in particular impedance matrix $\mathbf{Z}(\boldsymbol{\omega})$ or admittance matrix $\mathbf{Y}(\boldsymbol{\omega})$, as well as S-parameters.

At low frequencies, the system behaviour is defined by **R, L** and **C** matrices, which can be extracted by a static field analysis. According to the material constant variation theorem [4] it follows that $\mathbf{C}_{\mathrm{NBC}} < \mathbf{C}$, $\mathbf{L}_{\mathrm{NBC}} < \mathbf{L}$, where **C** and **L** are the exact parameters that correspond to the unbounded domain, and $\mathbf{C}_{\mathrm{NBC}}$ and $\mathbf{L}_{\mathrm{NBC}}$ are the parameters that correspond to the zero Neumann condition. Therefore, the Neumann boundary conditions provide lower bounds for the extracted parameters. In order to obtain upper bounds for the extracted parameters, the material constants outside the computational domain should be increased. Let consider the degenerate case: $\mu \to \infty$, $\varepsilon \to \infty$. This choice corresponds to zero Dirichlet (dual) boundary conditions. The parameters extracted so satisfy $\mathbf{C}_{\mathrm{DBC}} > \mathbf{C}$, $\mathbf{L}_{\mathrm{DBC}} > \mathbf{L}$. By combining these dual approaches the "Strategic Dual Image" (SDI) method, proposed by Saito [5] is obtained, where the numerical solution considered is the average of the solutions obtained with zero Neumann and Dirichlet boundary conditions

$$\mathbf{C}_a = (\mathbf{C}_{\mathrm{NBC}} + \mathbf{C}_{\mathrm{DBC}})/2,\ \mathbf{L}_a = (\mathbf{L}_{\mathrm{NBC}} + \mathbf{L}_{\mathrm{DBC}})/2, \tag{8}$$

The inequalities satisfied by the exact solution:

$$\mathbf{C}_{\mathrm{NBC}} \leq \mathbf{C} \leq \mathbf{C}_{\mathrm{DBC}},\ \mathbf{L}_{\mathrm{NBC}} \leq \mathbf{L} \leq \mathbf{L}_{\mathrm{DBC}}, \tag{9}$$

reveal the following bounds of the numerical error:

$$\|\mathbf{C} - \mathbf{C_a}\| < \|\mathbf{C_{NBC}} - \mathbf{C_{DBC}}\|/\mathbf{2},\ \|\mathbf{L} - \mathbf{L_a}\| < \|\mathbf{L_{NBC}} - \mathbf{L_{DBC}}\|/\mathbf{2}. \tag{10}$$

Numerical experiments show a reduction of the degrees of freedom of about 10, with respect to the classical Neumann condition, for the same numerical accuracy, if dual boundary conditions (9) are used. The modelling effort is further reduced if it is used an Equivalent Layer for Open Boundary (ELOB) with optimal material constants [4].

4. MODEL EXTRACTION — FRQUENCY ANALYSIS

In order to find solutions of real complex problems, an original methodological approach called *ALROM (All Level Reduced Order Modeling)* has been developed [6–8]. The ALLROM strategy consists of four stages: *macro-modeling, a-priori ROM*, *on the fly ROM* and *a-posteriory ROM*, aiming to reduce the number of DOFs as much as possible in each stage. The basic steps of the proposed algorithm are:

(1) Grid Calibration by dual-Finite Integrals Techniques (dFIT): mesh is refined up to an optimal one, which provides a numerical solution with acceptable accuracy (the scissors of dual solutions is close enough);

(2) Calibration of Virtual Boundary: the domain is ballooned up to an optimal size, when the difference between the solutions related to the dual boundary conditions is low enough;

(3) Frequency Analysis by Adaptive Frequency Samples. For a sequence of frequency samples ω is computed the system responce: $\mathbf{Y}(\boldsymbol{\omega}) = (\mathbf{Y}_p(\boldsymbol{\omega}) + \mathbf{Y}_s(\boldsymbol{\omega}))/2$ and its sensitivity, where

- $\mathrm{Y_p}$ (ω) is admitance computed by FIT on the primary grid with ELOB parameters: $\varepsilon_r = M >> 1$, $\mu_r = 1$;
- $\mathrm{Y_s}(\omega)$ is admittance computed by FIT on the secondary grid with ELOB parameters: $\varepsilon_r = 1$, $\mu_r = M >> 1$.

The only theoretical reason known by authors for this dual averaging is valid only for low frequencies.

5. PARAMETRIC MODELING BY ADJOINT FIELD TECHNIQUE

The problem of efficient sensitivity estimation and optimization with full-wave EM analysis remains a challenge. The most effective approach to handle the small parameter variations is to compute the sensitivities by Adjoint Field Technique — AFT, which is as well fast and accurate [9]. It can be used for efficient gradient-based optimization, in statistical, tolerance and yield analysis. We propose in present paper another application of this method, namely the model extraction. The multidimensional response surface over the design space of a given passive integrated component is built by Hermitte interpolation, considered not only the response in parameter samples but also their derivative, evaluated by AFT.

Adjoint-based sensitivity analysis for circuits has been well studied and extensively covered in the microwave literature [10]. In comparison, sensitivities with full-wave analysis techniques have attracted little attention, and there have been few applications into feasible and versatile algorithms. The great advantage of FIT is the easiness to build the equivalent electrical circuit of full-wave numerical discretization of Maxwell equations. It comprises two coupled circuits: an electrical one (with voltage sources controlled in flux derivatives, which represent the induced voltages) and a magnetic one (with "voltage" sources controlled by currents of electrical circuit). Therefore, the adjoint-circuit method can be applied in order to compute sensitivity of numerical field solution. Considering the general case of a multiport passive component with hybrid excitations, described by:

$$\mathbf{y} = \mathbf{H}\mathbf{x} \tag{11}$$

where $\mathbf{x} = [\mathbf{v}_a;\ \mathbf{i}_b]^{\mathrm{T}} = [\mathrm{V1},\ \mathrm{V2},\ \ldots,\ \mathrm{V}_{na},\ \mathrm{I}_{na+1},\ \ldots,\ \mathrm{I}_{na+nb}]^{\mathrm{T}}$ and $\mathbf{y} = [\mathbf{i}_a;\ \mathbf{v}_b]^{\mathrm{T}} = [\mathrm{I1},\ \mathrm{I2},\ \ldots,\ \mathrm{V}_{na},\ \mathrm{V}_{na+1},\ \ldots,\ \mathrm{V}_{na+nb}]^{\mathrm{T}}$ are the vectors of input and output signals, respectively and the hybrid matrix is:

$$\mathbf{H} = \begin{bmatrix} \mathbf{Y} & \mathbf{N} \\ \mathbf{M} & \mathbf{Z} \end{bmatrix}. \tag{12}$$

As in [10], the component sensitivity may be a result of perturbation in all design parameters p_n, n= 1, ..., N:

$$\Delta G = \sum_{n=1}^{N} \frac{\partial G}{\partial p_n} \Delta p_n = \hat{\mathbf{x}}^T \cdot \Delta\mathbf{H} \cdot \mathbf{x} \tag{13}$$

$\hat{\mathbf{x}} = \left[-\hat{\mathbf{v}}_a; \hat{\mathbf{i}}_b\right]^{\mathrm{T}}$ is the input signal of the adjoint circuit with output $\hat{\mathbf{y}} = \left[-\hat{\mathbf{i}}_a; \hat{\mathbf{v}}_b\right]^T$ given by $\hat{\mathbf{y}} = \mathbf{H}^{\mathrm{T}} \cdot \hat{\mathbf{x}}$.

The hybrid matrix of the adjoint circuit is the transpose matrix of the original circuit. If the circuit consists of M elements, each defined by its constitutive hybrid matrix $\mathbf{H}_m$, according to the consequence of Tellegen's theorem, the currents of voltage excitation-independent sources and the voltage of current sources satisfy:

$$\Delta\mathbf{i}_v^T \cdot \hat{\mathbf{v}}_v - \Delta\mathbf{v}_i^T \hat{\mathbf{i}}_i = \sum_{m=1}^{M} \Delta G_m = \sum_{m=1}^{M} \hat{\mathbf{x}}_m^T \cdot \Delta\mathbf{H}_m \cdot \mathbf{x}_m. \tag{14}$$

Consequently, the sensitivity of any port response f (I or V) with respect to any design parameter p_n is

$$\frac{\partial f}{\partial p_n} = -\sum_{m=1}^{M} \hat{\mathbf{x}}_m^T \cdot \frac{\partial \mathbf{H}_m}{\partial p_n} \cdot \mathbf{x}_m \tag{15}$$

The only terms that need to be evaluated for different n are matrix $\mathbf{H}_m$ derivatives. For any pair of corresponding branches associated to the edge/face m (in Fig. 1), the constitutive matrix of the FIT equivalent circuit is [7]:

$$\mathbf{H}_m = \begin{bmatrix} \dfrac{1}{G_m + j\omega G_m} & j\omega \\ 1 & R_m \end{bmatrix}, \quad \text{with} \quad G_m = \frac{\sigma_m A_m}{l_m}, \quad C_m = \frac{\varepsilon_m A_m}{l_m}, \quad R_m = \frac{l_m}{\mu_m A_m}. \tag{16}$$

If p_n is geometric parameter, e.g., a domain size, for the cells placed on the border of that domain:

$$\frac{\partial \mathbf{H}_m}{\partial p_n} = \begin{bmatrix} \dfrac{\pm 1}{A_m(\sigma_m + j\omega\varepsilon_m)} & j\omega \\ 1 & \dfrac{\pm 1}{A_m \mu_m} \end{bmatrix}, \quad \text{and it is zero in rest.} \tag{17}$$

REFERENCES

1. Clemens, M. and T. Weiland, "Discrete electromagnetism with the finite integration technique," *Progress in Electromagnetics Research, PIER*, Vol. 32, 65–87, 2001.
2. Ioan, D., M. Radulescu, and G. Ciuprina, "Fast extraction of static electric parameters with accuracy control," *Scientific Computing in Electrical Engineering*, (W.H.A. Schilders et al. Eds), Vol. 8, 248–256, Springer, 2004.
3. Ioan, D. and I. Munteanu, "Missing link rediscovered: the electromagnetic circuit element concept," *JSAEM Studies in Applied Electromagnetics and Mechanics*, Vol. 8, 302–320, 1999.
4. Ioan, D., G. Ciuprina, and M. Radulescu, "Absorbing boundary conditions for compact modeling of on-chip passive structures," *COMPEL: The International Journal for Computation and Mathematics in Electrical and Electronic Engineering*, Vol. 25, No. 3, 652–659, 2006.
5. Chen, Q. and A. Konrad, "A review of finite element open boundary techniques for static and quasi-static electromagnetic field problems," *IEEE Trans. on Magnetics*, Vol. 33, No. 1, 663–675, 1997.
6. Ioan, D., G. Ciuprina, M. Radulescu, and E. Seebacher, "Compact modeling and fast simulation of on-chip interconnect lines," *IEEE Transactions on Magnetics*, Vol. 42, No. 4, 547–550, 2006.
7. Ioan, D., G. Ciuprina, and M. Radulescu, "Algebraic sparsefied partial equivalent electric circuit — ASPEEC," *Scientific Computing in Electrical Engineering*, (Anile, A. M.; Alì, G.; Mascali, G., Eds.), Vol. 9, 45–50, Springer, 2006.
8. Meuris, P., G. Ciuprina, and E. Seebacher, "High-frequency simulations and compact models compared with measurements for passive on-chip component," *International Journal of Numerical Modelling*, Vol. 18, No. 3, 189–201, John Wiley, 2005.
9. Ioan, D., I. Munteanu, and G. Ciuprina, "Adjoint field technique applied in optimal design of a nonlinear inductor," *IEEE Transactions on Magnetics*, Vol. 34, No. 5, 2849–2852, 1998.
10. Nikolova, N. K., J. W. Bandler, and M. H. Bakr, "Adjoint techniques for sensitivity analysis in high-frequency structure," *IEEE Transactions on MTT*, Vol. 52, No. 1, 403–413, 2004.

Space Mapping and Neuro-space Mapping for Microwave Design

J. W. Bandler[1] and Q. J. Zhang[2]

[1]McMaster University, Hamilton, Canada
[2]Carleton University, Ottawa, Canada

Abstract— Space Mapping (SM) is one of the most recognized new developments in microwave CAD in recent years. It overcomes the limitations of speed and accuracy in conventional modeling and optimization and addresses the challenges of ever increasing design complexity, tighter component tolerances and shorter design cycles. It helps achieve EM/physics oriented design with much high efficiency than straightforward use of conventional EM/physics solvers. A highlight of recent advance is neuro-space mapping combining the advantages of neural network learning with the power of space mapping optimization, for both passive and active device modeling and optimization. This paper is an overview of advances in these areas.

1. INTRODUCTION

Over the recent years, we have witnessed significant advances in modeling and design methodologies to meet the challenges of next generation RF/microwave design. Increasing design complexity, coupled with tighter component tolerances and shorter design cycles, demand design methodologies that are faster, more accurate and automated than possible with conventional CAD methodology. Modeling and optimization with physical/geometrical information, including electromagnetic (EM)/physics effects, become necessary. This trend poses significant challenge to design automation process such as yield optimization requiring highly repetitive evaluation of component models. Due to the computational expense of EM/physics models, simply substituting conventional equivalent electrical models by EM/physics models by conventional optimization methods will not work, because of extremely long or prohibitive computation. Such gap between low-level EM/physics simulations versus high-level design optimization must be addressed in order to meet the needs of next generation CAD. CAD procedures such as statistical analysis and yield optimization taking into account process variations and manufacturing tolerances in the components demands that the component models are accurate and fast so that design solutions can be achieved feasibly and reliably. To achieve first-pass success in the next generation of high-frequency electronic design, we need EM/physics based component solutions on a much larger scale in circuit and systems design optimization.

A design frontier that remains to be conquered is the application of optimization procedures when direct application of traditional optimization approaches is not practical. The recent exploitation of iteratively refined surrogates in conjunction with "fine" or accurate models and the implementation of Space Mapping technology address this issue. Space Mapping is a recent breakthrough in engineering optimization, allowing expensive EM optimization to be performed more efficiently with the help of ideal, fast, approximate or low-fidelity "coarse" or surrogate models [1, 2].

Space Mapping has been applied with great success to otherwise expensive direct EM optimizations of microwave components and circuits with substantial computation speedup [3–14]. Recent advances have been made on mathematical motivation to place Space Mapping into the context of classical optimization. The aim of Space Mapping is to achieve a satisfactory solution with a minimal number of computationally expensive "fine" model evaluations. Space Mapping procedures iteratively update and optimize surrogates based on fast physically-based "coarse" models. The methodology has also been adopted in diverse engineering design applications such as electronic components, magnetic systems, and civil and mechanical engineering structures. Space mapping facilitates efficient optimization while avoiding direct optimization of the detailed fine model. A highlight of recent advance is neuro-space mapping combining the advantages of Artificial Neural Network (ANN) learning with the power of space mapping optimization, for both passive and active device modeling [15, 16] and optimization [17].

2. ADVANCES IN SPACE MAPPING

Significant progress is underway towards creating an automated user-friendly Space Mapping Framework (SMF, 2006). Recent approaches to space mapping-based optimization have included the original algorithm, the Broyden-based aggressive space mapping algorithm, various trust region approaches, neural space mapping and implicit space mapping [6–14].

A powerful optimization algorithm that combines Space Mapping (SM) with a novel Output Space Mapping (OSM) has been recently presented. In a handful of fine model evaluations, it delivers for the first time the accuracy expected from classical direct optimization using sequential linear programming. The method employs an SM-based interpolating surrogate (SMIS) framework that aims at locally matching the surrogate with the fine model. A highly optimized six-section H-plane waveguide filter design emerges after only four HFSS EM simulations, excluding necessary Jacobian estimations, using our algorithm with sparse frequency sweeps.

A comprehensive microwave design framework was developed for implementing the original, aggressive, implicit, and response residual SM approaches through widely available software. An instructive "multiple cheese-cutting" example demonstrates the SM approach to engineering design and some possible pitfalls. An Agilent-ADS framework implements the SM steps interactively. A three-section transformer example illustrates the approach, step by step. A six-section H-plane waveguide filter design emerges after four iterations, using the implicit SM and the response residual (RRSM) optimization entirely within the design framework. An RRSM surrogate is developed to match the fine (HFSS) model. We use sparse frequency sweeps and do not require Jacobians of the fine model.

Theoretical justification of SM methods were recently presented. A formal definition of optimization algorithms using surrogate models based on SM technology is given. Convergence conditions for the chosen subclass of algorithms are discussed and explained using a synthetic example, the so-called generalized cheese-cutting problem. An illustrative, circuit-theory based example is also utilized to further demonstrate the practicality of the theory. A space mapping design framework, space-mapping-based interpolation for engineering optimization, and TLM-based modeling and design exploiting space mapping have been developed. A recent review of the SM technique is presented in [12]. The most recent breakthrough in the area includes a software package for user-friendly space mapping optimization (SMF) [13], and extension of SM to modeling of microwave devices [14].

3. ADVANCES IN NEURO-SPACE MAPPING

Significant progress has been made in combining artificial neural network technology and space mapping into an advanced CAD algorithm [15–17]. Recently a method called neuro-space mapping (Neuro-SM) is developed for automated modeling of nonlinear devices. It is a systematic computational method to address the situation where an existing device model cannot fit new device data well. By modifying the current and voltage relationships in the model, Neuro-SM produces a new model exceeding the accuracy limit of the existing model. In a further step, a novel analytical formulation of Neuro-SM is developed to achieve the same accuracy as the basic formulation of Neuro-SM (known as circuit-based Neuro-SM) with much higher computational efficiency. Through our derivations, the mapping between the existing (coarse) model and the overall Neuro-SM model is analytically achieved for DC, small-signal, and large-signal simulation and sensitivity analysis. The analytical formulation is a significant advance over the circuit-based Neuro-SM, due to the elimination of extra circuit equations needed in the circuit-based formulation. A 2-phase training algorithm utilizing gradient optimization is also developed for fast training of the analytical Neuro-SM models. Application examples on modeling HBT, MESFET, and HEMT devices and the use of Neuro-SM models in harmonic balance simulations demonstrate that the analytical Neuro-SM is an efficient approach for modeling various types of microwave devices. It is useful for systematic and automated update of nonlinear device model library for existing circuit simulators.

REFERENCES

1. Bandler, J. W., Q. S. Cheng, S. Dakroury, A. S. Mohamed, M. H. Bakr, K. Madsen, and J. Søndergaard, "Space mapping: the state of the art," *IEEE Trans. Microwave Theory Tech.*, Vol. 52, 337–361, 2004.
2. Bandler, J. W., R. M. Biernacki, S. H. Chen, P. A. Grobelny, and R. H. Hemmers, "Space mapping technique for electromagnetic optimization," *IEEE Trans. Microwave Theory Tech.*, Vol. 42, 2536–2544, 1994.
3. Bandler, J. W., R. M. Biernacki, S. H. Chen, R. H. Hemmers, and K. Madsen, "Electromagnetic optimization exploiting aggressive space mapping," *IEEE Trans. Microwave Theory Tech.*, Vol. 43, 2874–2882, 1995.

4. Bakr, M. H., J. W. Bandler, R. M. Biernacki, S. H. Chen, and K. Madsen, "A trust region aggressive space mapping algorithm for EM optimization," *IEEE Trans. Microwave Theory Tech.*, Vol. 46, 2412–2425, 1998.
5. Bandler, J. W., N. Georgieva, M. A. Ismail, J. E. Rayas-Sánchez, and Q. J. Zhang, "A generalized space mapping tableau approach to device modeling," *IEEE Trans. Microwave Theory Tech.*, Vol. 49, 67–79, 2001.
6. Bandler, J. W., M. A. Ismail, and J. E. Rayas-Sánchez, "Expanded space-mapping EM-based design framework exploiting preassigned parameters," *IEEE Trans. Circuits SystemsI*, Vol. 49, 1833–1838, 2002.
7. Bandler, J. W., Q. S. Cheng, N. K. Nikolova, and M. A. Ismail, "Implicit space mapping EM-based modeling and design using preassigned parameters," *IEEE Trans. Microwave Theory Tech.*, Vol. 52, 378–385, 2004.
8. Ismail, M. A., D. Smith, A. Panariello, Y. Wang, and M. Yu, "EM-based design of large-scale dielectric-resonator filters and multiplexers by space mapping," *IEEE Trans. Microwave Theory Tech.*, Vol. 52, 386–392, 2004.
9. Bandler, J. W., Q. S. Cheng, D. H. Hailu, and N. K. Nikolova, "A space mapping design framework," *IEEE Trans. Microwave Theory Tech.*, Vol. 52, 2601–2610, 2004.
10. Koziel, S., J. W. Bandler, and K. Madsen, "Space-mapping-based interpolation for engineering optimization," *IEEE Trans. Microwave Theory Tech.*, Vol. 54, 2006.
11. Bandler, J. W., A. S. Mohamed, and M. H. Bakr, "TLM-based modeling and design exploiting space mapping," *IEEE Trans. Microwave Theory Tech.*, Vol. 53, 2801–2811, 2005.
12. Bandler, J. W., S. Koziel, and K. Madsen, "Space mapping for engineering optimization," *SIAG/Optimization Views-and-News Special Issue on Derivative-free and Surrogate Optimization*, Vol. 17, No. 1, 19–26, March 2006.
13. Koziel, S., J. W. Bandler, and K. Madsen, "Space mapping optimization algorithms for engineering design," *IEEE MTT-S Int. Microwave Symp. Dig.*, San Francisco, CA, June 2006.
14. Koziel, S. and J. W. Bandler, "Space-mapping-based modeling utilizing parameter extraction with variable weight coefficients and a data base," *IEEE MTT-S Int. Microwave Symp. Dig.*, San Francisco, CA, June 2006.
15. Bandler, J. W., M. A. Ismail, J. E. Rayas-Sánchez, and Q. J. Zhang, "Neuromodeling of microwave circuits exploiting space mapping technology," *IEEE Trans. Microwave Theory Tech.*, Vol. 47, 2417–2427, 1999.
16. Zhang, L., J. J. Xu, M. Yagoub, R. Ding, and Q. J. Zhang, "Efficient analytical formulation and sensitivity analysis of neuro-space mapping for nonlinear microwave device modeling," *IEEE Trans. Microwave Theory Tech.*, Vol. 53, 2752–2767, 2005.
17. Bakr, M. H., J. W. Bandler, M. A. Ismail, J. E. Rayas-Saánchez, and Q. J. Zhang, "Neural space mapping optimization for EM-based design," *IEEE Trans. Microwave Theory Tech.*, Vol. 48, 2307–2315, 2000.

Applications of Artificial Neural Network Techniques in Microwave Filter Modeling, Optimization and Design

H. Kabir[1], **Y. Wang**[2], **M. Yu**[2], **and Q. J. Zhang**[1]
[1]Department of Electronics, Carleton University, Ottawa, Ontario, Canada
[2]COM DEV Ltd., Cambridge, Ontario, Canada

Abstract— This paper reviews state-of-the-art microwave filter modeling, optimization and design methods using artificial neural network (ANN) technique. Innovative methodologies of using ANN in microwave filter analysis and synthesis are discussed. Various ANN structures including wavelet and radial basis function have been utilized for this purpose. ANN also finds application in filter yield prediction and optimization. The results from different work demonstrate that ANN technique can reduce the cost of computation significantly and thus can produce fast and accurate result compared to the conventional electromagnetic (EM) methods.

Microwave filters are widely used in satellite and ground based communication systems. The full wave EM solvers have been utilized to design these kinds of filters for a long time. Usually several simulations are required to meet the filter specifications which takes considerable amount of time. In order to achieve first pass success with only minor tuning and adjustment in the manufacturing process, precise electromagnetic modeling is an essential condition. The design procedure usually involves iterating the design parameters until the final filter response is realized. The whole process needs to be repeated even with a slight change in any of the design specifications. The modeling time increases as the filter order increases. With the increasing complexity of wireless and satellite communication hardware, there is a need for faster method to design this kind of filters. Artificial neural network (ANN) or simply neural network (NN) has been proven to be a fast and effective means of modeling complex electromagnetic devices. It has been recognized as a powerful tool for predicting device behavior for which no mathematical model is available or the device has not been analyzed properly yet. ANN can be trained to capture arbitrary input-output relationship to any degree of accuracy. Once a model is developed it can be used over and over again. The trained model delivers the output parameters very quickly. This avoids any EM simulation where a simple change in the physical dimension requires a complete simulation. For these attractive qualities, ANN has been applied to different areas of engineering and biomedical. While the present state of the art neural network modeling technique for EM modeling and optimization is available in detail in [1–3], it is beyond the scope of this paper. This paper reviews the ANN techniques dealing with microwave filter.

Waveguide cavity filters are very popular in microwave applications. Several results have been reported using neural network techniques to model cavity filters including E-plane metal-insert filter [4], rectangular waveguide H-plane iris bandpass filter [5–8], dual mode pseudo elliptic filter [9], cylindrical posts in waveguide filter [10] and etc. The simplest form of modeling is the direct approach where the geometrical parameters are related to its frequency response. Response of a filter is sampled at different frequency points to generate the training data. Result shows that ANN can provide accurate design parameters and after learning phase the computational cost is lower than the one associated with full wave model analysis [4]. In a similar work the performance of filter obtained from the ANN was much better than obtained from parametric curve and faster than finite element method (FEM) analysis [5].

Simpler structure or lower order filter is feasible to realize the whole model in a single NN model. For higher order filter several assumptions and simplifications are required to lower the number of NN inputs. Filter can be modeled by segmentation finite element (SFE) method and using ANN [6]. Filter structure was segmented into small regions connected by arbitrary cross section and then the smaller sections were analyzed separately. The generalized scattering matrix (GSM) was computed by FEM and the response of the complete circuit was obtained by connecting the smaller sections in proper order. In general the optimization of microwave circuits is time consuming. To attain a circuit response by analytical method is too slow. Therefore, ANN based analytical models were used. The method was applied to a three-cavity filter. The response of the filter rigorously found from SFE was compared with the same response obtained from the GSMs of the irises computed from ANN and excellent agreement was observed. In similar approach smooth piecewise linear

(SPWL) neural network model can be utilized for design and optimization of microwave filter [7]. SPWL has the advantage of smooth transitions between linear regions through the use of logarithm of hyperbolic cosine function. This feature suits well for the inductive iris modeling. A rectangular waveguide inductive iris band pass filter was modeled using SPWL neural network model. Several multi section Chebyshev band pass filters in different bands have been tested and each showed very good agreement with full 3D electromagnetic solution. Again using the NN speeds up the design process significantly.

Waveguide dual-mode pseudo-elliptic filters are often used in satellite applications due to its high Q, compact size and sharp selectivity [11]. Recently NN modeling technique has been applied to design wave-guide dual-mode pseudo-elliptic filter [9]. The coupling mechanism for dual mode filters is complex in nature and the numbers of variables are quite high. This makes the data generation and NN training an overwhelmingly time-consuming job. Therefore, filter structures were decomposed into different modules each representing different coupling mechanism. This ensures faster data generation, NN training and better accuracy. This model may be applied to filter with any number of poles as long as the filter structure remains the same. Due to the coupling between orthogonal models, GSM of the discontinuity junctions in the filter is necessary to characterize most of the modules. Equivalent circuit parameters such as coupling values and insertion phase lengths were extracted from EM data first. Neural network models were then developed for the circuit parameters instead of EM parameters. The method was applied to a four pole filter with 2 transmission zeros. The filter was decomposed into three modules: input-output coupling iris, internal coupling iris and tuning screw. NN models were developed for each module and irises and tuning screw dimensions were calculated using the trained NN models. The dimensions found from the NN models are within 1% of the ideal ones.

The other popular type of microwave filters is built in planar configuration such as microstrip and strip line. Numerous works have been published modeling microwave filters using ANN including low pass microstrip step filter [12], coupled microstrip band pass filter [13–22], microstrip band rejection filter [23], coplanar waveguide low pass filter [24] etc. The trained neural networks become fast filter model so that a designer can get the parameters quickly by avoiding long EM simulations. Wide bandwidth band pass filters were designed using microstrip line coupling at the end [14]. Coupling gaps are critical for designing these kinds of filters and the optimization of gaps require significant amount of time. To speed up the optimization of coupling gaps ANN models were developed and these models were used to design a filter. For a given filter specifications, physical parameters were obtained using ANN models. With these physical dimensions the filter was analyzed using a circuit simulator. A significant improvement in terms of speed has been realized using ANN models. The method can be generalized for low-pass, high pass, band pass or band rejection filters using planar configuration. A little modification is needed if the structure of the filter is changed from microstrip to strip line, but the general process remains the same. ANN models can be developed to model the entire filter if the number of variables is kept low. For larger dimensions some parameters are kept constant to keep the model simple.

Multi-layer asymmetric coupled microstrip line has been modeled using ANN [16]. The ANN replaces the time-consuming optimization routines to determine the physical geometry of multi-conductor multi-layer coupled line sections. ANN models for both synthesis and analysis were developed. The methodology was applied to a two layer coupled line filter and compared with segmentation and boundary element method (SBEM). Circuit elements were obtained much faster by ANN models than the optimization method. Circuit parameters can also be used as modeling parameters for this kind of filter. For all these cases ANN models are capable of predicting the dimensions or circuit parameters accurately compared to that obtained from the analytical formulas.

Microstrip filter on PBG structure were also designed using neural network models [17]. A new NN function called sample function neural network (SFNN) was employed for the modeling purpose. The PBG structures are periodic structures that are characterized by the prohibition of electromagnetic wave propagation at some microwave frequencies. A 2 dimensional square lattice consisting of circular holes were considered as the modeling problem. Radius of the circle of the periodic holes and frequency was input and s-parameters were considered as output of the neural network. Regular MLP was unable to converge to right solutions. RBF and wavelet functions improved the result but it was not accurate enough. Due to these reasons a new activation function called the sample activation function were used. The result shows that the SFNN can produce complex input-output relationship and could model the PBG filters on microstrip circuits accurately.

Neural network has been combined with some other optimization process in order to achieve filter design parameters quickly. A design technique combining finite-difference-time domain (FDTD) and neural network was proposed [19]. Two-stage time reduction was realized by utilizing an autoregressive moving-average (ARMA) signal estimation technique to reduce the computation time of each FDTD run and then the number of FDTD simulations was decreased using a neural network as a device model. The neural network maps geometrical parameters to ARMA coefficients. The trained network was incorporated with an optimization procedure for a microstrip filter design and significant time saving was achieved. Different algorithms can be developed combining NN and optimization method for faster and accurate filter solution. A Neuro-genetic algorithm was developed for microwave filter [20]. NN models were combined with genetic algorithms to synthesize millimeter wave devices. The method has been used to synthesize low pass and band pass filters in microstrip configuration. While the method worked well for low pass filters it showed limited accuracy for band pass filter. In order to overcome the problem some modification is required in the layout and design space.

Wavelet neural network (WNN) [22] and radial basis function (RBF) [12] can be advantageous for some special applications. Wavelet radical and the entire network construction have a reliable theory, which can avoid the fanaticism of network structure like back propagation (BP) NN. Also it can radically avoid the non-linear optimization issue such as local most optimized during the network training and have strong function study and extend ability. For these qualities WNN was chosen in [22]. Microstrip band pass filter was optimized where the geometrical parameters were changed to obtain the desired output response. The result was compared with that obtained using ADS optimizer. Fast and accurate results were obtained. In a similar work, radial basis function neural networks (RBF-NN) were used to model microstrip filter. Segmentation of the structure was employed for a 13 sections microwave step filter. Using the RBF-NN shows much faster and better accurate result than full wave analysis.

NN also finds applications in the design of microwave filters consists of dielectric resonator [25]. A rigorous and accurate EM analysis of the device was performed with FEM and combined with a fast analytical model. The analytical model was derived using segmented EM analysis applying to neural network. The method was then applied to dielectric resonator (DR) filters and good agreement between theoretical and experimental result has been achieved within a few iterations.

NN has been employed to obtain starting point for optimizer used for yield prediction algorithm [26]. The yield was computed as a ratio of the number of cases passing the specification to the total number of simulations performed. For efficient calculation of yield, the choice of starting point is critical. It requires the knowledge of final solution, which is not available. Neural network was used to predict this solution and then the solution was used as the starting point of the optimization. Different structures realizing the same response was used to calculate the yield. Result suggests that by using neural network models, computational effort can be reduced significantly.

In conclusion this paper has reviewed the role of ANN in microwave filter modeling, optimization and design. The ANN method provides fast and accurate results and reduces the computational costs associated with a time consuming EM solver in the design of microwave filters. These methods can be used in combination with standard filter design methods to design complex microwave filters. It helps to improve the speed and accuracy of filter design for communication circuit and systems.

REFERENCES

1. Zhang, Q. J., K. C. Gupta, and V. K. Devabhaktuni, "Artificial neural networks for RF and microwave design: from theory to practice," *IEEE Trans. Microwave Theory Tech.*, Vol. 51, 1339–1350, April 2003.
2. Rayas-Sanchez, J. E., "EM-Based optimization of microwave circuits using artificial neural networks: the state-of-the-art," *IEEE Trans. Microwave Theory Tech.*, Vol. 52, No. 1, 420–435, January 2004.
3. Zhang, Q. J. and K. C. Gupta, *Neural Networks for RF and Microwave Design*, Artech House, Boston, 2000.
4. Burrascano, P., M. Dionigi, C. Fancelli, and M. Mongiardo, "A neural network model for CAD and optimization of microwave filters," *IEEE MTT-S Int. Microwave Symp. Digest*, Vol. 1, 13–16, June 1998.
5. Fedi, G., A. Gaggelli, S. Manetti, and G. Pelosi, "Direct-coupled cavity filters design using a hybrid feed forward neural network- finite elements procedure," *Int. Journal of RF and Microwave CAE*, Vol. 9, 287–296, May 1999.

6. Cid, J. M. and J. Zapata, "CAD of rectangular waveguide H-plane circuits by segmentation, finite elements and artificial neural networks," *IEE Electronic Letters*, Vol. 37, 98–99, January 2001.
7. Mediavilla, A., A. Tazon, J. A. Pereda, M. Lazaro, I. Santamaria, and C. Pantaleon, "Neuronal architecture for waveguide inductive iris bandpass filter optimization," *Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks*, Vol. 4, 395–399, July 2000.
8. Fedi, G., A. Gaggelli, S. Manetti, and G. Pelosi, "A finite-element neural-network approach to microwave filter design," *Microwave and Optical technology letters*, Vol. 19, 36–38, September 1998.
9. Wang, Y., M. Yu, H. Kabir, and Q. J. Zhang, "Effective design of waveguide dual mode filter using neural networks," *IEEE MTT-S Int. Microwave Symposyum*, San Francisco, USA, June 2006.
10. Fedi, G., S. Manetti, G. Pelosi, and S. Selleri, "Design of cylindrical posts in rectangular waveguide by neural network approach," *IEEE International Symposium of Antenna and Propagation*, Vol. 2, 1054–1057, July 2000.
11. Kudsia, C., R. Cameron, and W. C. Tang, "Innovations in microwave filters and multiplexing networks for communications satellite systems," *IEEE Trans. Microwave Theory Tech.*, Vol. 40, No. 6, 1133–1149, June 1992.
12. Nunez, F. and A. K. Skrivervik, "Filter approximation by RBF-NN and segmentation method," *IEEE MTT-S Int. Microwave Symp. Digest*, Vol. 3, 1561–1564, June 2004.
13. Ciminski, A. S., "Artificial neural networks modeling for computer aided design of microwave filter," *International Conference on Microwaves, Radar and Wireless Communications*, Vol. 1, 96–99, May 2002.
14. Chao, C. and K. C. Gupta, "EM-ANN modeling of overlapping open-ends in multiplayer lines for design of band pass filters," *IEEE AP-S Int Symp. Digest*, 2592–2595, Orlando, USA, July 1999.
15. Li, X., J. Gao, J. Yook, and X. Chen, "Bandpass filter design by artificial neural network modeling," *Proceedings of APMC*, 2005.
16. Watson, P. M., C. Cho, and K. C. Gupta, "Electromagnetic-Artificial neural network model for synthesis of physical dimensions for multilayer asymmetric coupled transmission structures," *Int. Journal of RF and Microwave CAE*, Vol. 9, 175–186, May 1999.
17. Fernandes, E. N. R. Q., P. H. F. Silva, M. A. B. Melo, and A. G. d'Assuncao, "A new neural network model for accurate analysis of microstrip filters on PBG structure," *Proceedings of European Microwave Conference*, Italy, September 2002.
18. Gunez, F. and N. Turker, "Artificial neural networks in their simplest forms for analysis and synthesis of RF/Microwave planar transmission lines," *Int. Journal of RF and Microwave CAE*, Vol. 15, No. 6, 587–600, November 2005.
19. Banciu, M. G., E. Ambikairajah, and R. Ramer, "Microstrip filter design using fdtd and neural networks," *Microwave and Optical Technology Letters*, Vol. 34, No. 3, 219–224, August 2002.
20. Pratap, R. J., J. H. Lee, S. Pinel, G. S. May, J. Laskar, and E. M. Tentzeris, "Millimeter wave RF front end design using neuro-genetic algorithms," *Proceedings of Electronic Component and Technology Conference*, Vol. 2, 1802-1806, May–June 2005.
21. Peik, S. F., G. Coutts, and R. R. Mansour, "Application of neural networks in microwave circuit modeling," *Proceedings of IEEE Canadian Conference on Electrical and Computer Engineering*, 928–931, May 1998.
22. Li, M., X. Li, X. Liao, and J. Yu, "Modeling and optimization of microwave circuits and devices using wavelet neural networks," *IEEE International Conference on Communications, Circuits and Systems*, 1471–1475, June 2004.
23. Ciminski, A. S., "Artificial neural networks modeling for computer-aided design for planar band-rejection filter," *International Conference on Microwave, Radar and Wireless Communications*, Vol. 2, 551–554, 2004.
24. Gati, A., M. F. Wong, G. Alqui, and V. F. Hanna, "Neural networks modeling and parameterization applied to coplanar waveguide components," *Int. Journal of RF and Microwave CAE*, Vol. 10, 296–307, September 2000.

25. Bila, S., D. Baillargeat, M. Aubourg, S. Verdeyme, and P. Guillon, "A full electromagnetic CAD tool for microwave devices using a finite element method and neural networks," *Int. Journal of Numerical Modeling*, 167–180, 2000.
26. Harish, A. R., "Neural network based yield prediction of microwave filters," *Proceedings of IEEE APACE*, 30–33, Shah Alam, Malaysia, 2003.

Neural-based Transient Behavioral Modeling of IC Buffers for High-speed Interconnect Design

Yi Cao[1], **Qi-Jun Zhang**[1], **and Ihsan Erdin**[2]

[1]Department of Electronics, Carleton University, Ottawa, ON, K1S 5B6, Canada
[2]Nortel Networks, Ottawa, ON, Canada

Abstract— Artificial neural networks (ANN) have gained attention as fast and flexible vehicles to microwave modeling and design. This paper reviews a recent advance of neural network modeling, i.e., state-space dynamic neural network (SSDNN) for transient behavioral modeling of high-speed nonlinear circuits. The SSDNN model can be directly trained from the input and output waveforms without relying on the circuit internal details. A training algorithm exploiting adjoint sensitivities is summarized for training the model in an efficient manner. An example of the SSDNN technique for IC buffer modeling and its use with transmission line elements in high-speed interconnect design are included.

With the continuous increase of signal speed and frequency, signal integrity (SI) in VLSI packages becomes more and more prominent. Fast and accurate representations of the nonlinear analog behaviors of driver/receiver buffers are the key to the success of SI-based design of high-speed interconnects with nonlinear terminations [1, 2]. As such, developing efficient buffer models for transient applications has become an important topic [3–5]. To ensure model reliability in circuit simulations, the model stability remains one of the most critical aspects of nonlinear transient modeling. In the neural network community, global asymptotical stability and global exponential stability have been studied for some special classes of dynamic networks, e.g., Hopfield neural networks [6], recurrent neural networks [7], and discrete-time state-space neural networks [8]. Recently stability for ANN-based analog microwave modeling has also been addressed [9]. This paper summarizes a state-space dynamic neural network (SSDNN) technique for modeling nonlinear transient behaviors of IC drivers and receivers [5]. We describe the detailed structure of SSDNN and how to train the model based on the transient waveforms from the original circuits. An example is provided to demonstrate the application of the SSDNN model in coupled transmission line environments.

Let $\boldsymbol{u} \in \Re^M$ be transient input signals of a nonlinear circuit, e.g., input voltages and currents, and $\boldsymbol{y} \in \Re^K$ be transient 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. Based on combining state-space concept and continuous recurrent neural network method [10], the SSDNN nonlinear model is formulated as [5]

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

where $\boldsymbol{x} = [x_1, \ldots, x_N]^T \in \Re^N$ is the vector of the state variables and N denotes the dimension of the state-space, or order of the model. Here $\boldsymbol{g}_{ANN} = [g_{ANN\_1}, \ldots, g_{ANN\_N}]^T$ represents a feed-forward Multilayer Perceptron (MLP) neural network [11], 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 state-space into output-space.

Training is an essential step to ensure the SSDNN model can accurately represent the original nonlinear circuit. The training data for SSDNN is generated from the detailed nonlinear circuit in the form of transient waveforms. 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 $\mathbf{y}^i(t)$ represent SSDNN prediction of 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

$$E_d = \frac{1}{2} \sum_{i=1}^{L} \int_0^{T_1} \|\mathbf{y}^i(t) - \mathbf{y}_d^i(t)\|^2 dt \tag{2}$$

is minimized. Since training is essentially an optimization process, fast and accurate sensitivity information is important in order to utilize gradient-based training methods. Based on the concept of adjoint dynamic neural network [12], training error derivatives can be derived as

$$\frac{\partial E_d}{\partial \boldsymbol{w}} = -\int_0^{T_1} \tau \hat{\boldsymbol{x}}^T \frac{\partial \boldsymbol{g}_{ANN}}{\partial \boldsymbol{w}} dt \tag{3}$$

where $\hat{\boldsymbol{x}} \in \Re^N$ is the vector of the adjoint state variables computed from the adjoint SSDNN system

$$\dot{\hat{\boldsymbol{x}}} = \hat{\boldsymbol{x}} - \tau \frac{\partial \boldsymbol{g}_{ANN}^T}{\partial \boldsymbol{x}} \hat{\boldsymbol{x}} + \boldsymbol{C}^T(\boldsymbol{y} - \boldsymbol{y_d}) \tag{4}$$

with initial condition $\hat{\boldsymbol{x}}(T_1) = 0$.

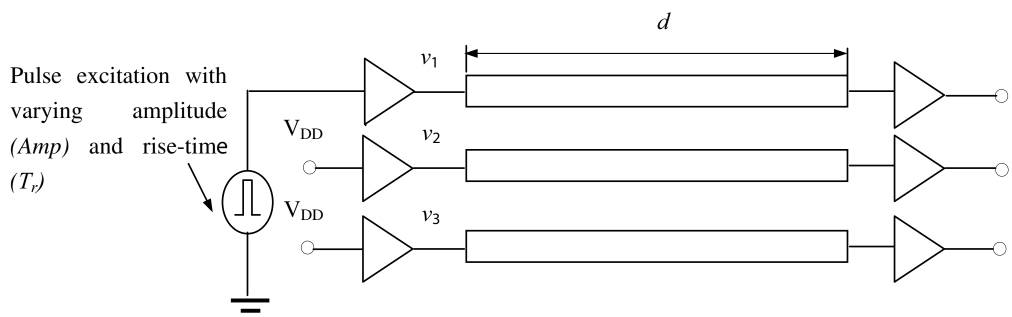


Figure 1: A coupled 3-conductor interconnect circuit terminated with nonlinear buffers. This circuit is used to generate test waveforms for testing trained DNN models of nonlinear drivers. Varying parameters in the pulse signal generator allows us to obtain different sets of testing waveforms.

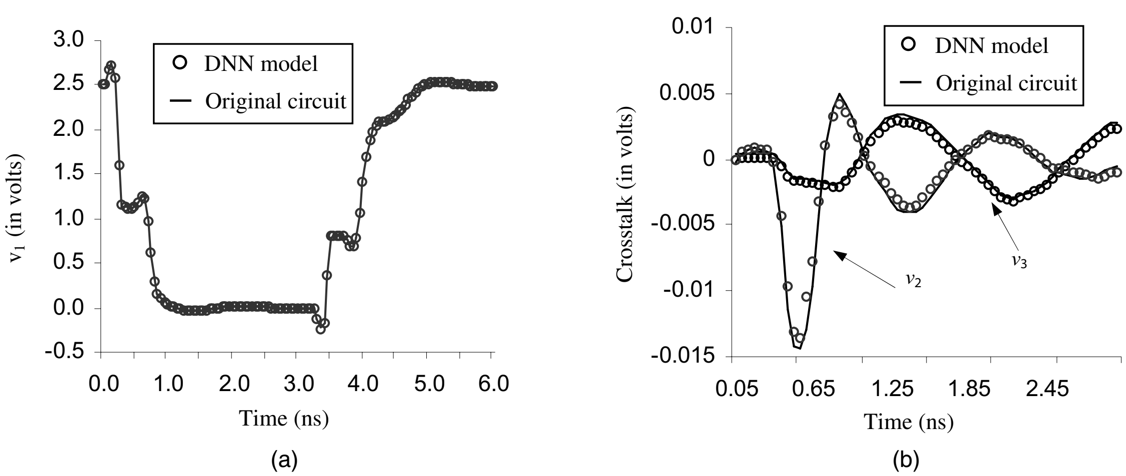


Figure 2: (a) Signal propagation waveform v_1 and (b) crosstalk waveforms v_2 and v_3 of the VLSI interconnect example of Fig. 1 where $Amp = 2.4$v, $d = 2.5$ cm, and $T_r = 0.2$ ns. Excellent agreement is achieved between the responses of DNN models with those obtained from original HSPICE simulation.

Here we illustrate the SSDNN modeling of a circuit-based high-speed interconnect driver [12]. A specialized version of SSDNN, dynamic neural network (DNN), is used in this example. Two sets of training waveforms, namely, signal data and crosstalk data are generated based on level-49 BSIM3V3 *HSPICE* transistor models. After training is finished, we use the trained DNN for VLSI interconnect simulations with coupled transmission lines. Specifically, a trained DNN model of order 2 with 50 hidden neurons is used 3 times as the nonlinear terminations in the coupled transmission line circuit of Fig. 1 for signal delay and crosstalk analysis of VLSI interconnect networks. As can be seen in Fig. 2, excellent agreement is achieved between *HSPICE* simulations of the original interconnect circuit (including coupled transmission lines and the original *HSPICE* CMOS inverter drivers) and our DNN-based interconnect simulations. This example verifies the accuracy of the SSDNN technique in signal integrity analysis of high-speed interconnect networks.

REFERENCES

1. Achar, R. and M. S. Nakhla, "Simulation of high-speed interconnects," *Proc. IEEE*, Vol. 89, No. 5, 693–728, 2001.
2. Lum, S., M. Nakhla, and Q. J. Zhang, "Sensitivity analysis of lossy coupled transmission lines with nonlinear terminations," *IEEE Trans. Microwave Theory Tech.*, Vol. 42, No. 4, 607–615, 1994.
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. Mutnury, B., M. Swaminathan, and J. Libous, "Macro-modeling of non-linear I/O drivers using spline functions and finite time-difference approximation," *Proc. Electrical Performance of Electronic Packaging*, 273–276, Princeton, NJ, Oct. 2003.
5. 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.
6. Guan, Z. H., G. R. Chen, and Y. Qin, "On equilibria, stability, and instability of Hopfield neural networks," *IEEE Trans. Neural Networks*, Vol. 11, No. 3, 534–540, 2000.
7. Chu, T. G., C. S. Zhang, and Z. D. Zhang, "Necessary and sufficient condition for absolute stability of normal neural networks," *Neural Networks*, Vol. 16, No. 8, 1223–1227, 2003.
8. Zamarreño, J. M. and P. Vega, "State space neural network: properties and application," *Neural Networks*, Vol. 11, No. 6, 1099–1112, 1998.
9. 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.
10. Leistritz, L., M. Galicki, H. Witte, and E. Kochs, "Training trajectories by continuous recurrent multilayer networks," *IEEE Trans. Neural Networks*, Vol. 15, No. 3, 283–291, 2002.
11. Haykin, S., *Neural Networks, A Comprehensive Foundation*, Macmillan, New York, NY, 1994.
12. 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.

A Space-Mapping Based CAD Methodology for Modeling Temperature Characteristics of Combline Resonators

A. Kashi[1], **P. Kumar**[2], **M. Caron**[2], **and V. K. Devabhaktuni**[1]

[1]Department of ECE, Concordia University
S-EV5.139, 1455 de Maisonneuve W., Montreal, H3G 1M8, Canada
[2]Mitec Telecom Inc., 9000 Trans-Canada Highway, Montreal, H9R 5Z8, Canada

Abstract— In this paper, we propose a new space-mapping based CAD modeling methodology for modeling temperature characteristics of combline resonators. With the aid of two commercial simulation tools, namely Ansoft HFSS and Agilent ADS, this methodology generates accurate temperature models of combline resonators. The method is generic i. e., it is capable of generating corresponding models for a variety of resonator structures including the mushroom and straight, using only one frequency sweep in HFSS. Fine EM simulations from HFSS are used to create an ultra-fast model in ADS, which is extremely useful to RF/microwave designers.

1. INTRODUCTION

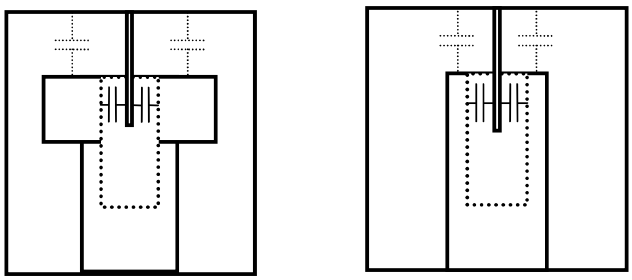
Frequency drift due to temperature changes is a key concern in RF/microwave filter design. This aspect is particularly critical in the case of combline resonators used in a wide variety of RF/microwave systems. Several conventional temperature compensation techniques have been employed. For instance, special materials with low thermal coefficients of expansion, e. g., invar, have been used. However, such materials could result in other problems such as low thermal conductivity and high cost. Alternatively, it was shown that for a combline resonator with given dimensions of housing and resonator rod, perfect temperature compensation can be achieved at two resonant frequencies *via* appropriate selection of materials for the housing and the rod [1–4]. Another recent temperature compensation method [5] advocated use of a shape memory alloy (SMA) actuator.

In order to integrate temperature compensation solutions such as the above into RF/microwave filter design, an accurate and fast temperature drift model of a combline resonator is essential. A simple relationship between the resonant frequency of a combline resonator and the temperature exists [5]. However, this traditional model is not accurate enough and fails to cover different rod structures (see Fig. 1). As such, there is a demand for a reasonably accurate model, which takes into account all the physical dimensions of a combline resonator. A direct approach to address this problem is to use detailed theoretical models such as the full-wave EM solvers. In practice, full-wave EM simulations tend to be CPU-expensive and hence unsuitable for design/optimization tasks. This scenario leads to research on novel EM based CAD models, which is the subject of this paper.

In this paper, a CAD methodology based on space-mapping is proposed for the development of accurate/fast temperature models of combline resonators. The methodology is implemented using appropriate coarse model (Agilent ADS) and fine model (Ansoft HFSS) as in any space-mapping framework. All the steps involved in the proposed methodology are illustrated through a practical example.

2. REVIEW OF SPACE-MAPPING FRAMEWORK

Space-mapping tool [6–9] effectively connects fast coarse models to align with accurate/CPU-intensive fine models in the design parameter space. For instance, the response-residual space-mapping (RRSM) surrogate is matched with the fine model response through a parameter-extraction process [10]. As shown in [10], SM-optimization/modeling algorithms involve four essential steps as shown in Fig. 2. Initially, suitable coarse and fine models are selected. In comparison with the fine model, the coarse model should possess acceptable accuracy and relatively high computation-speed. In the second step, a suitable mapping function is selected. Commonly used mathematical functions include simple linear, polynomial, or rational functions and complex Artificial Neural Network (ANN) functions. In the third step, the fine model is simulated, e. g., simulation of an EM structure in HFSS workspace. The final step is parameter extraction, which is an iterative process leading to computation of mapping parameters. At the end of the fourth step, the mapping is updated and the surrogate model is derived. Subsequently, the response of the surrogate model is



(a) (b)

Figure 1: (a) Mushroom rod and (b) straight rod resonator structures. Capacitors shown by dotted and solid lines contribute to equivalent gap capacitor (C_{gap}) and equivalent screw capacitor (C_{screw}) respectively.

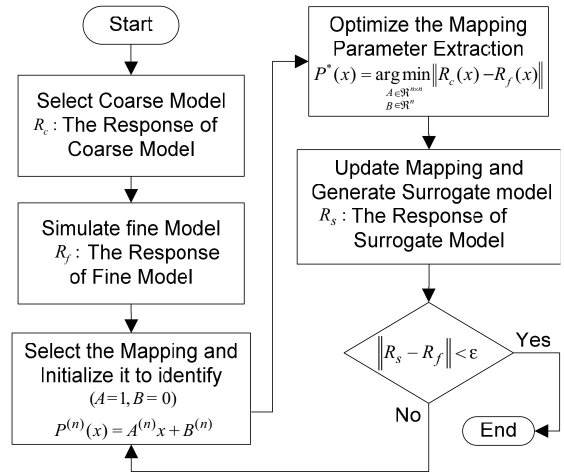


Figure 2: Flowchart of the space-mapping algorithm [10] used in this work for temperature modeling of combline resonators.

verified to see whether or not it matches the fine model's response over the entire frequency range of interest. If the two models do not match, the algorithm repeats steps two through four.

2.1. Coarse Model

Agilent ADS has a vast library of circuit models that can be used as "coarse" models. As shown in Fig. 3, the idealized coarse model for a combline resonator is a coaxial transmission line, which is loaded by two types of capacitors. The effective capacitance between the rod and the cavity, shown in Fig. 1, can be calculated using parallel plate capacitor equation

$$C_{\text{gap}} = \frac{\varepsilon_0 A}{d} \tag{1}$$

where ε_0, A and d represent permittivity of the free space, area of the parallel plates, and separation between the plates respectively. Another type of capacitance, which is caused by the penetration of the tuning screw into the rod, can be calculated using the cylindrical capacitance equation

$$C_{\text{screw}} = \frac{2\pi k \varepsilon_0}{\ln(b/a)} L, \tag{2}$$

where ε_0, k, L, b and a denote permittivity of the free space, relative permittivity, length of screw penetration, diameter of the rod hole, and diameter of the screw respectively. Based on the above calculations, the input impedance of the coarse model is given by

$$Z_{\text{in}} = -jZ_0 \tan\frac{\omega}{c} l + \frac{1}{j\left(C_{\text{gap}} + C_{\text{screw}}\right)\omega} \quad Z_0 = \sqrt{\frac{\mu}{\varepsilon}} \frac{\ln b/a}{2\pi}, \tag{3}$$

where Z_0, l, and c represent characteristic impedance, length of the rod, and speed of the light respectively.

In practice, the geometry of a combline resonator changes with temperature. As such, all the lengths such as cavity height and rod diameter change with respect to three factors: temperature drift, initial length and the material's coefficient of thermal expansion. In a case, where the rod is made up of two materials and has a different housing, the equivalent thermal expansion coefficient γ can be calculated as

$$\gamma = \frac{a}{(a+b)}\alpha + \frac{b}{(a+b)}\beta, \tag{4}$$

where α and β are the coefficients of thermal expansion of the two materials. In (4), a and b are the initial lengths of the two materials. In an attempt to take into account the effects of temperature,

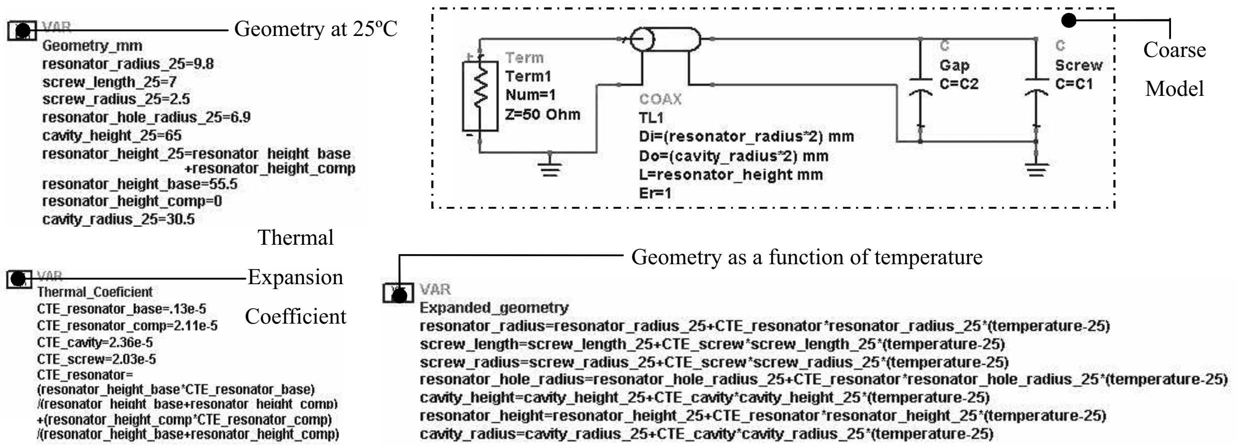


Figure 3: Representation of the coarse model as a function of temperature and γ in ADS.

we replaced the absolute geometrical values by temperature-dependent symbolic expressions as shown in Fig. 3.

2.2. Fine Model

In this work, EM simulations of combline resonator based on Ansoft HFSS are considered as the fine model. EM data from HFSS consisting of S-parameters versus frequency is in touchstone format. An S-parameter file, namely SnP in ADS, can import such data. In SnP, n stands for the number of ports. For example, Fig. 4 depicts a one-port S-parameter file component (S1P). It is to be noted that S-parameters can then be converted into Z-parameters by library functions of ADS.

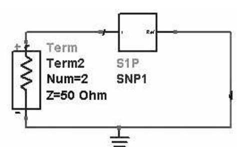


Figure 4: S1P representation in ADS.

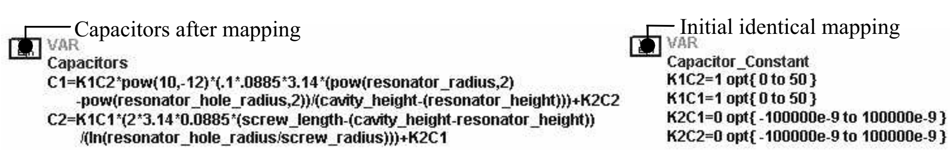


Figure 5: Implementation of the proposed linear mapping in ADS taking into account the fringing effects.

2.3. Linear Mapping

The capacitance values calculated using Equations (1) and (2) do not account for certain phenomena, e. g., fringing that depends on the geometry of the rod. As such, it is necessary to adjust these approximate values to accurate ones. The fringing effects are accounted for using the linear mapping

$$C_{\text{gap+fringing}} = ratio_{\text{gap}} \times C_{\text{gap}} + \mathit{offset}_{\text{gap}} \quad C_{\text{screw+fringing}} = ratio_{\text{screw}} \times C_{\text{screw}} + \mathit{offset}_{\text{screw}}, \quad (5)$$

where *ratios* and *offsets* are mapping parameters. Implementation details are shown in Fig. 5.

2.4. Parameter Extraction

Responses of coarse and fine models are matched through a parameter-extraction process. More specifically, $ratio_{\text{gap}}$, $ratio_{\text{screw}}$, $\mathit{offset}_{\text{gap}}$ and $\mathit{offset}_{\text{screw}}$ (5) are iteratively adjusted based on full-wave EM simulations from the fine model i.e., Ansoft HFSS. To implement this step, an optimization framework shown in Fig. 6 has been setup in ADS. The goal for such optimization process is to accurately match both magnitude and phase of Z_{in} computed using coarse and fine models. *Hybrid optimization* option in ADS is employed.

3. NUMERICAL RESULTS

For the purpose of illustration, a combline resonator with mushroom structure is considered. EM simulations have been performed using HFSS and resulting S-parameter data is loaded into ADS *via* an S1P file. Fig. 8 shows the matching of Z_{in} from the coarse model (after parameter extraction and coarse model update) and the fine model. The combline resonator is simulated for two different

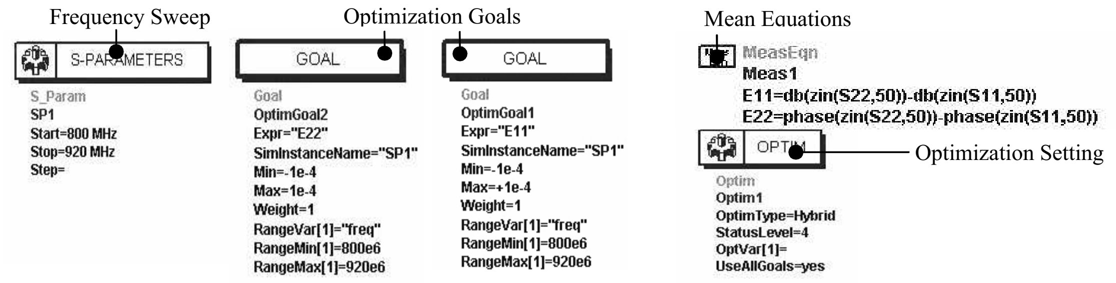
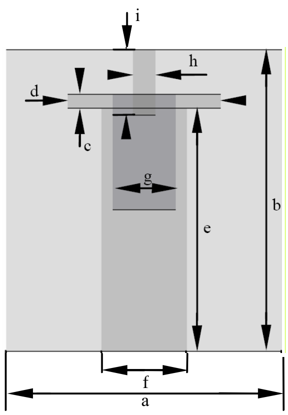
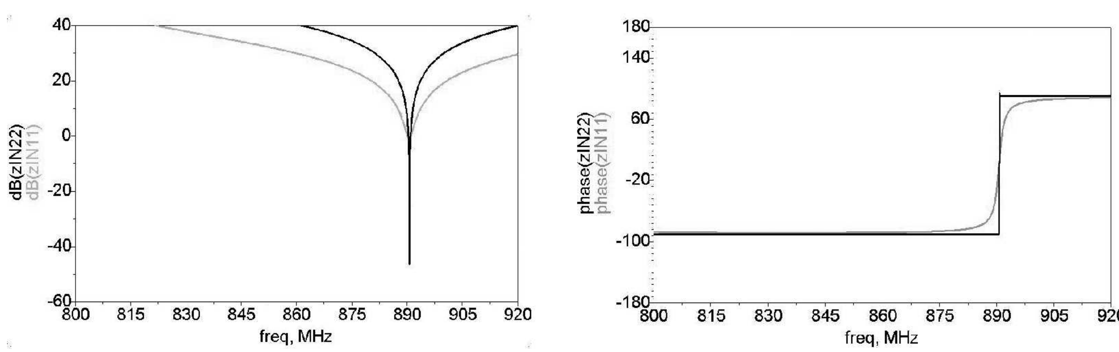


Figure 6: Implementation of the optimization procedure is ADS leading to parameter extraction.



Geometry	mm
Cavity Dia. (*a*)	61
Cavity Height (*b*)	65
Resonator Mushroom Height (*c*)	3
Resonator Mushroom Dia. (*d*)	34
Resonator Base Height (*e*)	52.5
Resonator Base Dia. (*f*)	19.06
Resonator Hole Dia. (*g*)	13.8
Screw Dia. (*h*)	5
Screw Length (*i*)	15

Figure 7: Geometry of the combline resonator modeled in this work.



(a) (b)

Figure 8: (a) Magnitude of the input impedance computed using the coarse (zIN11) and the fine (zIN22) models, (b) phase of the input impedance computed using the coarse (zIN11) and the fine (zIN22) models.

Table 1: Comparison of numerical results from simulations and measurements.

	CTE Cavity	Aluminum 6061 (23.6 ppm/ C)
	CTE Screw	**Brass (20.3 ppm/ C)**
Case #1	CTE Resonator	SS (9.9 ppm/ C)
	Temp. Drift (Simulated)	-6.714 ppm/ C
	Temp. Drift (Measured)	-4.93 ppm/ C
Case #2	CTE Resonator	Invar (1.3 ppm/ C)
	Temp. Drift (Simulated)	+1.856 ppm/ C
	Temp. Drift (Measured)	+1.82 ppm/ C

rod materials: stainless steel and invar. The results show a ± 2ppm/°C error, which is satisfactory in terms of model accuracy. Table 1 shows a comparison of the results from both simulations and measurements.

4. CONCLUSIONS

A new CAD modeling algorithm based on space-mapping has been proposed. Using the systematic steps of this methodology, temperature behavior of a variety of combline resonator structures can be modeled. The resulting models are accurate and fast and hence can be of practical significance in RF/microwave filter design/optimization.

ACKNOWLEDGMENT

The authors wish to gratefully acknowledge the support from the Canada Research Chairs (CRC) program, the Natural Sciences and Engineering Council of Canada (NSERC) and FQRNT/NATEQ. Mr. Iwan Jemczyk (Vice-President, Engineering) of Mitec Telecom is thanked for creating opportunities for collaboration.

REFERENCES

1. Wang, C. and K. A. Zaki, "Temperature compensation of combline resonators and filters," *IEEE MTT-S Int. Microwave Symp. Dig.*, Anaheim, CA, 1041–1044, June 1999.
2. Hui-Wen Y. and A. E. Atia, "Temperature characteristics of combline resonators and filters," *IEEE MTT-S Int. Microwave Symp. Dig.*, Phoenix, AZ, 1475–1478, May 2001.
3. Alford, N. M., J. Breeze, S. J. Penn, and M. Poole, "Tunable, temperature-compensated high Q and high thermal conductivity dielectrics for Ku and Ka band communications," *IEE Microwave Filters and Multi-plexers Colloquium*, Ref. 2000/117, 6/1–6/4, November 2000.
4. Chen, S. W., K. A. Zaki, and R. F. West, "Tunable, temperature compensated dielectric resonators and filters," *IEEE Trans. Microwave Theory Tech.*, Vol. 38, 1046–1052, 1990.
5. Keats, B. F., R. B. Gorbet, and R. R. Mansour, "Design and testing of SMA temperature-compensated cavity resonator," *IEEE Trans. Microwave Theory Tech.*, Vol. 51, 2284–2289, 2003.
6. Bandler, J. W., R. M. Biernacki, S. H. Chen, P. A. Grobelny, and R. H. Hemmers, "Space mapping technique for electromagnetic optimization," *IEEE Trans. Microwave Theory Tech.*, Vol. 42, 2536–2544, 1994.
7. Bandler, J. W., R. M. Biernacki, and S. H. Chen, "Fully automated space mapping optimization of 3D structures," *IEEE MTT-S Int. Microwave Symp. Dig.*, San Francisco, CA, 753–756, June 1996.
8. Bakr, M. H., J. W. Bandler, R. M. Biernacki, S. H. Chen, and K. Madsen, "A trust region aggressive space mapping algorithm for EM optimization," *IEEE Trans. Microwave Theory Tech.*, Vol. 46, 2412–2425, 1998.
9. Bandler, J. W., R. M. Biernacki, S. H. Chen, R. H. Hemmers, and K. Madsen, "Electromagnetic optimization exploiting aggressive space mapping," *IEEE Trans. Microwave Theory Tech.*, Vol. 43, 2874–2882, 1995.
10. Bandler, J. W., "A space-mapping design framework," *IEEE Trans. Microwave Theory Tech.*, Vol. 52, 2601–2610, 2004.

A Chebyshev Tapered TEM Horn Antenna

S. Bassam and J. Rashed-Mohassel

Center of Excellence on Applied Electromagnetic Systems
ECE Department, University of Tehran
Tehran, Iran

Abstract— An exponential TEM horn antenna with Chebeyshev impedance taper is designed and constructed. The results show improved directivity of the antenna and a wider range of frequencies with a VSWR less than two, over the entire band.

1. INTRODUCTION

TEM horn antennas have been used as wideband antennas for various applications. Typical applications for these antennas include EMC experiments, Ground Penetrating Radar (GPR) and feeds for reflectors [1, 2]. These antennas have the advantages of wideband, no dispersion, unidirectional and easy construction. However it has the disadvantage of a large size. Several methods have been introduced to improve the performance of the antenna and reduce its size. Some researchers have tried to reduce the size of the antenna by resistive loading, but this reduces the antenna efficiency. One other solution is to modify the shapes of the plates.

In this paper we have used a Chebyshev impedance taper for the antenna resulting in a better VSWR and improved gain.

2. TEM HORN PRELIMINARIES

The basic TEM horn antenna consists of two linear or exponential tapered metal plates [3]. The spacing between the plates can be linear or exponential. Linear tapered antennas can be built easily in contrast to an exponential taper; however exponential taper has the advantage of smooth impedance variations. The antenna acts as a transformer to match the transmission line and the free space. The antenna can be assumed as a multistage impedance transformer. To increase the frequency band of VSWR< 2, the number of steps should be increased such that the impedance varies continuously and smoothly. An alternative is using tapered transmission line.

Among the various impedance tapers, Chebyshev has the preferred optimum characteristics. This taper has equal-amplitude minor lobes and is an optimum design in the sense that it gives the smallest minor-lobe amplitude for a fixed taper length and vice versa [4]. In this model a Chebyshev impedance taper is used to match the characteristic impedance at the feed point to the impedance of free space at the antenna aperture. The impedance at each point of the antenna can be calculated by [4]:

$$\ln \bar{Z} = \left(\frac{p}{2\pi} + \frac{1}{2} - \frac{p}{2\pi \cosh \pi u_0}\right) \ln \bar{Z}_L + \frac{\ln \bar{Z}_L}{\pi \cosh \pi u_0} \sum_{n=1}^{\infty} \frac{\cos \pi \sqrt{n^2 - u_0^2} - \cos n\pi}{n} \sin np \tag{1}$$

Where, $\bar{Z}$ is the normalized impedance at each point, $\bar{Z}_L$ is the normalized load impedance which in this case is the intrinsic impedance of free space ($377\,\Omega$), Z_0 is the characteristic impedance ($100\,\Omega$) of the feed line, $\pi u_0 = \beta_0 L$, $\pi u = \beta L$, $p = 2\pi \frac{x - L/2}{L}$ where $-\pi < p < \pi$ and L is the taper length. If we assume the TEM horn as a parallel waveguide, we obtain the equation for $w(z)$ as:

$$w(x) = \eta d(x)/Z(x) \tag{2}$$

Where, $d(x)$ is separation of the plates, $w(x)$ is the width of the plates and η is the intrinsic impedance of the free space. Let $d(x)$ be an exponential of the form

$$d(x) = \alpha e^{(\beta x)} \tag{3}$$

Where α and β are constants to be identified. Design parameters can be calculated by equations (1), (2), and (3).

3. MODELING AND SIMULATION OF THE ANTENNA

A TEM horn antenna is simulated using FEKO which is based on method of moments. Chebyshev impedance taper is used to match the characteristic impedance at the feed point to the impedance of the free space at the antenna aperture. The separation of the plates is exponential as shown in Figure 1. The area around the feed is locally distorted such that a gap source can be placed between the two planar elements.

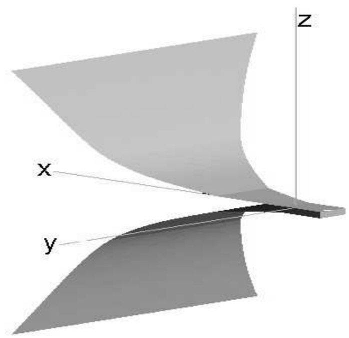


Figure 1: Tapered TEM horn antenna.

The dimensions of the structure are shown in Table 1. The antenna length is 60 cm and the aperture of the antenna is 50 cm×50 cm.

Table 1: Characteristics of the simulated antenna

X [cm]	Z[Ω]	d [cm]	W [cm]
−10	100	1	3.7
0	100	1	3.7
5	103.05	1.36	4.9
10	110.05	2.06	6.84
15	123.49	3.1	9.18
20	145.68	4.68	11.82
25	178.25	7.06	14.8
30	220.62	10.64	18.38
35	313.72	16.04	23.12
40	348.41	24.2	29.98
45	368.79	36.5	40.52
50	377	50	50

With the same dimensions we have simulated an exponential impedance taper antenna. Figure 2 shows the simulated VSWR over the entire frequency band of 400–5000 MHz for the Chebyshev

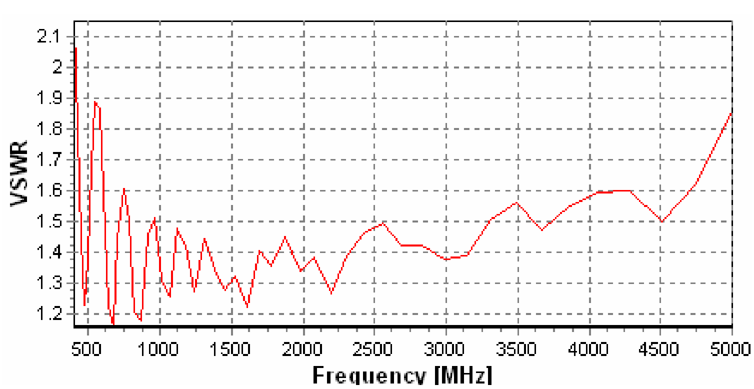


Figure 2: VSWR of the Chebyshev impedance taper antenna vs. frequency.

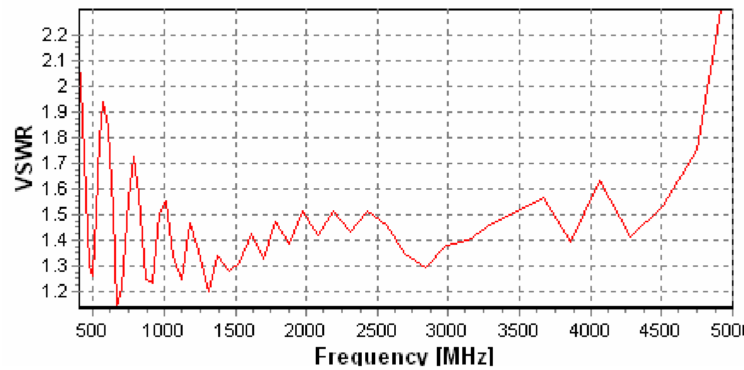


Figure 3: VSWR of the exponential impedance taper antenna vs. frequency.

antenna and Figure 3 shows the VSWR for the exponential antenna. The comparison of these two figures show that the Chebyshev taper has better matching over a wider range of frequencies.

4. EXPERIMENTATION AND MEASUREMENT

An experimental antenna is constructed using the results of the simulation, as shown in Figure 4. The material for the conducting plates is copper with 1-mm thickness, and for simplicity a ground plane and the image theory is used. For supporting and fixing the plate, a Styrofoam with $\varepsilon_r = 1.05$ is used. The antenna length is 60 cm with the center frequency of 2.7 GHz and with dimensions of the aperture as 50 cm×50 cm. The antenna is fed through a coaxial cable connected directly to the antenna plate. The parallel plate section of the feeding point has the dimensions of 1 cm×1.89 cm with 12 cm in length.

Figure 4: The experimental antenna.

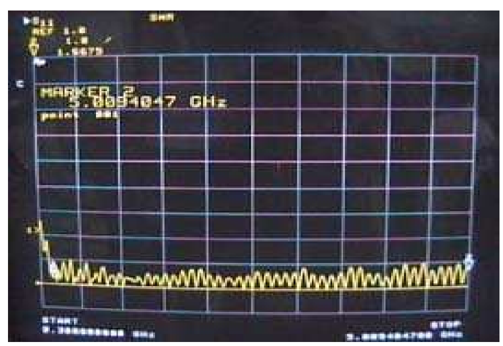


Figure 5: Measured VSWR of the antenna.

The VSWR of the antenna is measured by a 8410A network analyzer and the measured VSWR of the TEM horn over the frequency band of 447.4 MHz to 5 GHz is observed to be less than 2.0.

5. CONCLUSION

An exponential TEM horn antenna with Chebyshev impedance taper is simulated and constructed. The results show, better matching at the feed point and over wide range of frequencies. In this design, impedance changes smoothly from the feed line to the antenna aperture. In comparison with exponential impedance taper, it has wider frequency band with improved VSWR less than 2.

REFERENCES

1. Licul, S., "Ultra wideband characterization and measurement," PhD Thesis, Faculty of Virginia Polytechnic Institute & State University, September 2004.
2. Turk, A. S., "Ultra wideband TEM horn design for ground penetrating impulse radar system," *Microwave and Optical Technology Letters*, Vol. 41, No. 5, 333–336, June 2004.
3. Smith, G. S. and T. Lee, "A design study of the basic TEM horn antenna," *IEEE AP. Magazine*, Vol. 46, No. 1, 86–92, February 2004.
4. Collin, R., *Foundations For Microwave Engineering*, McGraw-Hill, Singapore, 1992.

5. Lee, S. and H. Choi, "Design of an exponentially tapered TEM horn antenna for the wide broadband communication," *Microwave and Optical Tech. Lett.*, Vol. 40, No. 6, March 2004.
6. Kolokotronis, D. A., Y. Huang, and J. T. Zhang, "Design of TEM horn antenna for impulse radar," High Frequency Postgraduate Student Colloquium, 120–126, September 1999.
7. Chang, K. and S. Yang, "Design of a wideband TEM horn antenna," *IEEE Trans. AP-S.*, Vol. 1, 229–232, June 2003.

Low Power Dissipation SEU-hardened CMOS Latch

Yuhong Li, **Suge Yue**, **Yuanfu Zhao**, **and Guozhen Liang**
Beijing Microelectronics Technology Institute, 100076, China

Abstract— This paper reports three design improvements for CMOS latches hardened against single event upset (SEU) based on three memory cells appeared in recent years. The improvement drastically reduces static power dissipation, reduces the number of transistors required in the VLSI, especially when they are used in the Gate Array. The original cells and the new improved latches are compared. It is shown that the new latch-NDICE latch has the best compositive capability and the best SEU immunity.

1. INTRODUCTION

To ensure the reliability of electronic systems in space, electronic circuitry must be resilient when exposed to the radiant space. The effects of the single event (SEE) have become more critical as clocks speeds have increased and feature sizes have decreased in modern IC processes. The main SEE includes SEU and single event latch (SET). SEU phenomenon may result in the corruption of data in memory cells and in the propagation of erroneous signals in combinational logic circuits. As a result, it is increasingly clear that hardening approaches for both combinational logic and memory elements are needed.

In recent years, there have been many designed SEU-Hardened CMOS memory cells which were designed by Rockett, Whitaker, Liu respectively and the HIT [1–5], DICE [6], DICE with Guard-gates [7]. Through [5] we know that the HIT is better than the pioneering three SEU-hardened memory cells. But the original HIT, DICE and DICE with Guard-gates have their advantages and disadvantages respectively.

The combinational logic and memory elements are mainly from latch, so this paper reports three SEU-Hardened CMOS latches considering of their static power dissipation and the number of the transistors synthetically. The three latches in the paper are mainly improved from the HIT, DICE and DICE with Guard-gates respectively. The results make them have lower static power dissipation and fewer transistors.

In Section 2, the design method, functionality and performance of the proposed SEU-hardened latches are described. Section 3 compares the improved latches with the original cells [5–7].

2. LATCH DESIGN

The three original cells used to design the SEU-hardened latches are shown in Figures 1, 2 and 3 respectively.

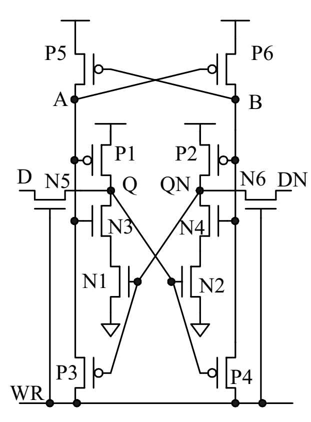


Figure 1: Original HIT cell.

2.1. The HIT Cell

The HIT cell is composed of 12 transistors organized as two storage structures interconnected by feedback paths. The read/write operation needs a single phase clock CK and differential input D and DN. The output Q and its complement QN are both available.

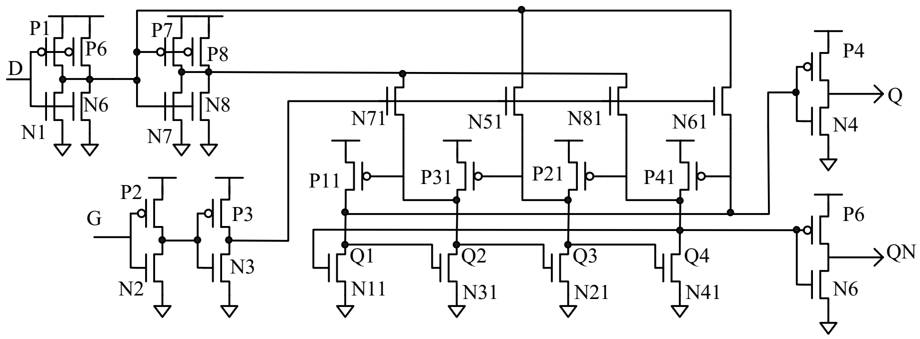


Figure 2: Original latch using DICE.

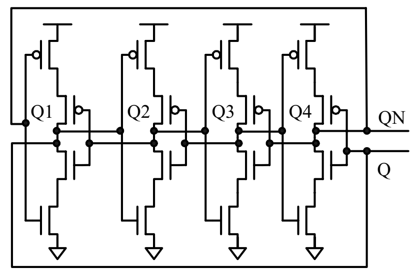


Figure 3: Original DICE with guard-gates.

It has three sensitive notes: Q, QN, A/B. If a particle strikes the drain of transistor to make only one sensitive node to upset, it can recover quickly. But if both nodes upset, the contents of the memory cell will be corrupted. To cope with single upset, specific transistors ratios have been used (Table 1).

But our aim is to use the cell in gate array. So the transistors must be in the same size. In order to make it suitable for our demand, we connect one PMOS transistor in parallel with P5 and P6 respectively.

Table 1: HIT cell device sizes.

Transistors	Wn(μm)	Ln(μm)	Transistors	Wp(μm)	Lp(μm)
N1	6.0	1.2	P1	8.0	1.2
N2	6.0	1.2	P2	8.0	1.2
N3	6.0	1.2	P3	4.0	1.2
N4	6.0	1.2	P4	4.0	1.2
N5	6.0	1.2	P5	6.0	1.2
N6	6.0	1.2	P6	6.0	1.2

All the transistors of the new HIT latch are the same size, none is special. Though it has two more PMOS transistors, there is no extra pre-existing word line driver needed, no matter what it is connected in the front. Also the parallel-connection make the part of P5 and P6 stronger than the P3 and P4. If a particle strikes the drain of transistor which is controlling the sensitive note: Q, QN, A/B, it can recover its initial state all. But it will lead to corruption of the data stored if both were stricken. There is no extra static power dissipation. In addition, the total dose effect is considered.

2.2. The DICE

The original DICE (Figure 2) puts no particular constraints on transistor sizes and it does not evidence the high sensitivity to total dose of the ration designs. Though it has four sensitive nodes: Q1, Q2, Q3 and Q4, the recovery process is very fast (much less than 1ns). It will lead to corruption of the data stored if both were stricken.

But it requires two pre-existing word line drivers and hence a small increase in the dynamic power dissipation. Additional design changes are required to adapt the word line routing, the write buffer drive capability and column pass gate width to the increased load requirements.

To overcome these limitations the new improved DICE latch is developed. In order to reduce the write transistor, take out N71 and N81. Connect PMOS transistor P51 and P61 in parallel with N51 and N61 respectively, it will reduce the static power dissipation and the loss of threshold. Connect P12, N12 with P11 and N11 respectively in series, connect P22, N22 with P21 and N21 respectively. It can prevent the write state from the strife of the stored state. So it will reduce the pre-existing word line driver. There is no extra pre-existing word line drive needed, no matter what it connects in the front. So the P6, N6, P7, N7, P8, and N8 are all taken out. There is only P1 and N1 needed. For the dual interlocked structure, it presents excellent SEU immunity.

2.3. The Guard-gates DICE

The radiation performance for the guard-gates DICE (Figure 3) is better than DICE, because the only way to upset these cells with ion hits is by depositing charges on at least three of the storage nodes — an extremely low probability event. The proposed design isolates the gate outputs during a strike to preserve the data. As a result, it provides better protection against multiple node hits than regular DICE design. The write time will be improved because the feedback loop within the latch is disconnected. Additionally, in the presence of a single-event hit, DICE latches are subjected to short circuit current because p- and n-channel devices within an *inverter* will turn on simultaneously creating a short circuit between supply voltage and ground. This facilitates the upset by changing the output node voltage away from the supply rails. The guard-gate based designs avoid such a short circuit by disconnecting the output from the power supply and ground during the presence of an SET.

But the total number of transistors in the four guard-gates DICE is twice as that for a conventional DICE latch. In order to reduce the number of the transistor and the static power dissipation., use the same principle as the New DICE latch, we get the new improved Guard-gates DICE latch.

3. COMPARISON AND SPICE SIMULATION

Table 2 and Table 3 show some performance comparisons between the three original memory cells and the three new improved latches. HIT, DICE and GDICE stand for the original HIT, DICE and Guard-gates DICE cells respectively. NHIT, NDICE and NGDICE stand for the new HIT latch, DICE latch and Guard-gates DICE latch.

Table 2: Comparisons between the three original memory cells.

Original cell	Power	N.PMOS	N.NMOS
HIT	2.2689×10^{-5}W	12	12
DICE	6.6599×10^{-5}W	12	16
GDICE	2.1262×10^{-5}W	14	18

From Table 2, we can't know the best cell. That is to say, the original cells have their own disadvantage and advantage. Through our improvement, all the power become smaller. Also we can see the best one is the NDICE which has the fewest transistors and least Power dissipation. The delay time is smaller.

Table 3: Comparisons between the three new latches.

New latch	Tplh (ns)	Tphl (ns)	Power (W)	N.PMOS	N.NMOS
NHIT	0.1	0.6	2.2689×10^{-5}	14	12
NDICE	0.2	0.3	1.0419×10^{-5}	13	13
NGDICE	0.3	0.4	1.1773×10^{-5}	17	17

For the sake of the SEU immunity of the latches, we design a single event transient model (Figure 4) to simulate the SEU. VCC is reduced from 1.8 V to 0 for 0.18 μm.

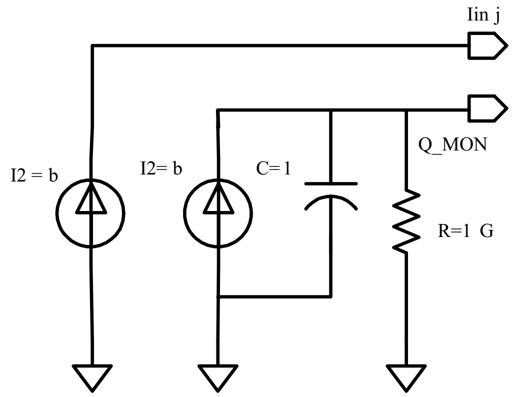


Figure 4: Single event transient model.

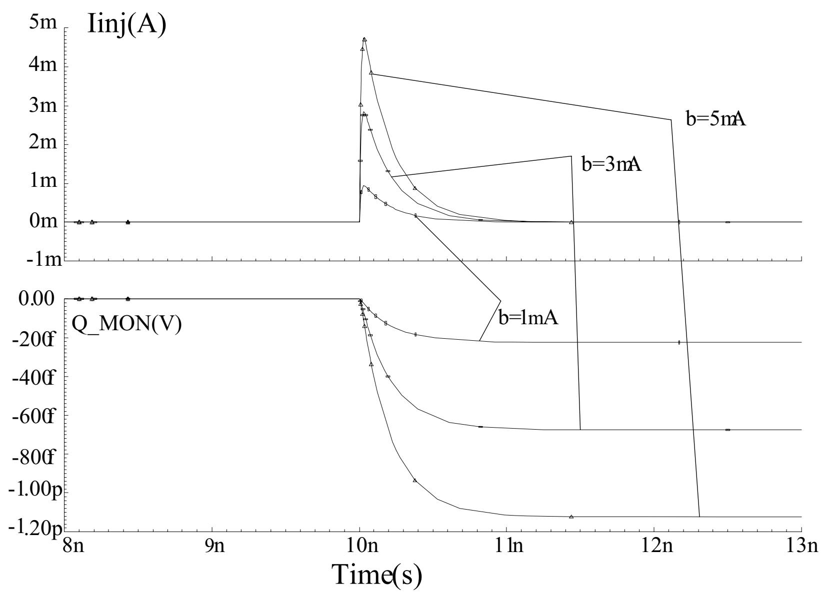


Figure 5: the comparision of b to $Q_{\rm crit}$.

In the SPICE simulator, the heavy ion effect is modeled by injecting a hit current pulse at the sensitive note. This current pulse is of a rough triangular shape with rise time equal to 30 ps, rise and fall damping factor equal to 10 ps and 200 ps. The injection node is clamped to either GND or VCC by a variable current during charge collection. The method assumes the hit pulse is much faster than the circuit response and ignores the detailed timing information carried in the pulse shape. It is the-worse-case-design simulation to generate a relatively conservation critical charge for SEU. In the model, the I2 is variable. From Figure 5 we can get that $b = 1\,\text{mA}$ cause crit $Q_{\rm cirt} = 0.2\,\text{PC}$.

Make use of the conclusion in [8], we know for the most sensitive node, the crit $Q_{\rm crit}$ to cause a propagating pulse is 0.02 PC and the threshold LET is 2 MeV-cm^2/mg. So we get that: $b = 1\,\text{mA}$ $\Rightarrow Q_{\rm crit}$=0.2 PC$\Rightarrow$ LET=20 MeV-cm^2/mg.

Connect the model to one of the sensitive node of the three new latches, get Table 4 (when the value of b become **the biggest current**, the sensitive note will not recover, and the stored state will be wrong).

From Table 4, we can see the SEU-immunity of the NHIT latch is weaker than that of the NDICE latch, and then, we generally define it as crashing SEU-immunity when its LET$_{\rm th}$ is above 100 MeV-cm^2/mg according to the former experience. At last we know that the NDICE latch has the best compositive capability and the best SEU immunity.

Table 4: The SEU-immunity of the three new latches.

New latch	The biggest current (b)	LET
NHIT latch	3.375 mA	67.5 MeV-cm^2/mg
NDICE latch	>5 mA	>100 MeV-cm^2/mg
NGDICE latch	>5 mA	>100 MeV-cm^2/mg

REFERENCES

1. Weaver, H. T., "An SEU tolerant memory cell derived from fundamental studies of SEU mechanisms in SRAM," *IEEE Trans. on Nucl. Sci*, Vol. NS-34, No. 6, 1281–1286, 1987.
2. Canaris, J., "An SEU immune logic family," *Proceedings of the Third NASA Symposium on VLSI Design, Moscow*, 2.3.1–2.3.12, October 1991.
3. Wiseman, D., J. Canaris, and S. Whitaker, "Design and testing of SEU/SEL immune memory and logic circuits in a commercial CMOS process," NSREC workshop, July 1993.
4. Liu, M. N. and S. Whitaker, "Low power SEU immune CMOS memory circuits," *IEEE Trans. Nucl. Sci.*, Vol. 39, No. 6, 1679–1684, 1992.
5. Bessot, D. and R. Velazco, "Design of seu-hardened CMOS memory cells: the hit cell," *Proceeding 1993 RADECS Conference*, 563–570.
6. Calin, T., M. Nicolaidis, and R. Velazco, *Upset Hardened Memory Design for Submicron CMOS Technology, IEEE Trans. Nucl. Sci.*, Vol. 43, No. 6, 2874–2878, 1996.
7. Bhuva, B. L., J. D. Black, and L. W. Massengill, "RHBD Techniques for mitigating effects of single-event hits using guard-gates," *IEEE Trans. Nucl. Sci.*, Vol. 52. No. 6, 2531–2535, 2005.
8. Wang, J. J., R. B. Katz, J. S. Sun, B. E. Cronquist, J. K. Mccollum, T. M. Speers, and W. C. Plants, "SRAM based re-programmable FPGA for space applications," *IEEE Trans. Nucl. Sci.*, Vol. 46, No. 6, 1728–1735, 1999.

Mutual Coupling of Rectangular DRA in a Four Element Circular Array

S. Jarchi and J. Rashed-Mohassel
Center of Excellence on Applied Electromagnetic Systems, ECE Department, University of Tehran, Iran

M. H. Neshati
University of Sistan and Baluchestan, Iran

Abstract— Mutual coupling of DRAs in a four element circular array is investigated. For several values of the array radius mutual coupling and return loss are computed and illustrated. The FDTD method with a UPML boundary condition is used for simulation.

1. INTRODUCTION

Specific features of DRAs has made them suitable for a variety of applications specially MMW applications. DRAs have small size and low cost. They can be easily coupled to almost all types of transmission lines [1]. They can be integrated easily with MMIC circuits. In MMW applications conductor loss of metallic antennas become severe and the antenna efficiency decreases significantly, conversely the only loss for a DRA is that due to the imperfect material of the DRA which can be very small in practice [2]. Therefore DRAs have high radiation efficiency. In comparison to microstrip antennas, DRAs have wider impedance bandwidths. For a typical DRA with dielectric constant of 10 the impedance bandwidth of 10% can be achieved. Avoidance of surface waves is another attractive advantage of DRAs over microstrip antennas.

Single DRAs of different shapes has been studied, including rectangular, cylindrical, hemispherical, triangular, conical, etc. Among these different shapes cylindrical and rectangular are the most common and the rectangular has the advantage of having one more degree of freedom for design purposes.

There are a variety of feed configurations, which electromagnetic fields can be coupled to DRAs. Most common feed arrangements are microstrip aperture coupling, direct microstrip coupling, probe coupling and conformal strip coupling. Among these feed configurations, aperture coupling is more suitable for MMW applications. In aperture coupling configuration, since the DRA is placed on the ground plane of the microstrip feed, parasitic radiation from the microstrip line is avoided. Isolation of the feed network from the radiating element is another advantage of the aperture coupling method.

In many cases with a single element DRA, desired specifications can not be achieved. For example a high gain, directional pattern can not be synthesized with a single DRA of any shape. In these applications, a DRA array with appropriate element arrangement and feed configurations, can be used to provide desired specifications. In DRA arrays proximity of elements produces mutual coupling. Usually this mutual coupling is considered as an undesired phenomenon because it can alter the array characteristics. However with an exact knowledge of mutual coupling between different elements of an array, this undesired phenomenon may be optimally used to provide specific desired characteristics.

In this paper mutual coupling of different elements of a four element circular array of DRAs and its influence on return loss is computed and illustrated. For simulation, the FDTD method with a UPML boundary condition is applied.

Section 2 provides a brief introduction to the FDTD method including source modeling and the frequency domain parameter definitions. In Section 3 the single DRA dimensions and the simulated response with the FDTD is presented. Section 4 is main contribution of this paper. In this section mutual coupling between different elements of a four element circular array is investigated and simulated with the FDTD method. Section 5 concludes the work.

2. THE FDTD METHOD

FDTD is one of the most common numerical techniques in electromagnetics. With simple formulation it can be easily applied to complex structures. In order to have valid results from the FDTD simulations a boundary condition should be applied. In this paper for investigating circular array, the UPML is applied. Since UPML has the potential of modeling inhomogenuities which extend to infinity and in the DRA with aperture coupling it is assumed that the ground plane and microstrip line extends to infinity.

FDTD is a time domain technique and the source should be applied in time domain. A Gaussian resistive voltage source with the following parameters is used for simulations [3]:

$$v = e^{-\frac{(t-t_0)^2}{T^2}} \tag{1}$$

$$T = 15\,ps, \qquad t_0 = 4T \tag{2}$$

where T and t_0 are chosen to have desired frequencies in the output and the source resistance is selected $50\,\Omega$ which is approximately the resistance of the microstrip feed at resonant frequency. Once FDTD analysis is done, frequency domain parameters can be obtained with a FFT procedure.

Return loss and mutual couplings are computed with the following definitions [4]:

$$S_{ij} = \frac{b_i}{a_j} \tag{3}$$

$$a_i = \frac{1}{2}\left(\frac{V_i}{\sqrt{Z_0}} + \sqrt{Z_0} I_i\right) \tag{4}$$

$$b_i = \frac{1}{2}\left(\frac{V_i}{\sqrt{Z_0}} - \sqrt{Z_0} I_i\right) \tag{5}$$

With these definitions we don't need to run the FDTD simulation program twice as in [5].

3. SINGLE DRA

The single DRA configuration which is fed by aperture coupling method is shown in Figure 1. The DRA dimensions are: $\varepsilon_{rs} = 10.2$, $\varepsilon_r = 10.8$, $2a = 15\,\text{mm}$, $b = 7.5\,\text{mm}$, $d = 3\,\text{mm}$, $L = 6.1\,\text{mm}$, $L_s = 2.2\,\text{mm}$, $w = 1.2\,\text{mm}$, $w_f = 0.64\,\text{mm}$ and $h = 0.64\,\text{mm}$. Return loss, computed with the FDTD method is plotted in Figure 2. Resonant frequency of the single DRA is approximately 7.2 GHz.

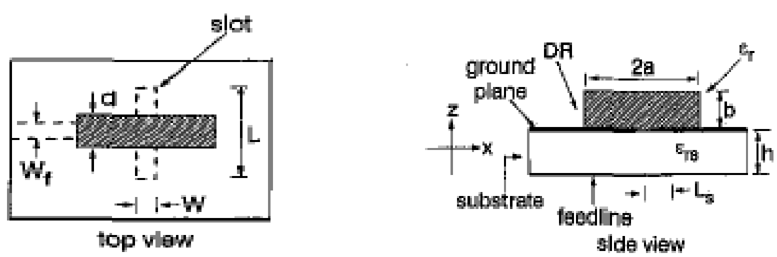


Figure 1: A single DRA [2]: (a) top view, (b) side view.

4. CIRCULAR ARRAY

The four element circular array configuration is shown in Figure 3. The array consists of four rectangular DRAs which are placed symmetrically on the circumference of a circle with radius D. A port is assigned to each DRA. Investigating this array is more complicated than linear arrays, since in the proposed configuration in addition to mutual coupling between E-plane elements, S_{31}, there is mutual coupling between circular staggered DRAs, i.e., S_{21} and S_{41}. Circular staggered configuration is thoroughly discussed in [6].

For investigating the proposed circular array, the array radius, D, is increased and mutual coupling and return loss are computed.

Mutual coupling between the E-plane DRAs, S_{31}, is computed and plotted in Figure 4. It is observed from this figure that for D equal to 0.25λ, S_{31} is high while for the other two values of D, mutual coupling is weak and below $-20\,\text{dB}$. It is expected from this figure that for D equal to 0.25λ, mutual coupling influence on return loss should be observable.

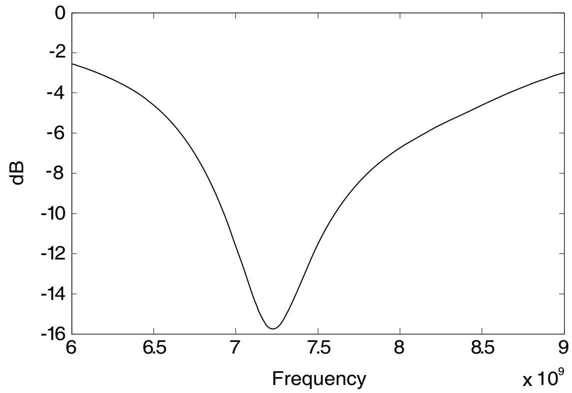


Figure 2: Single DRA return loss.

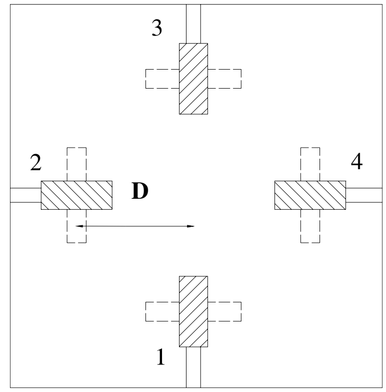


Figure 3: Circular array configuration.

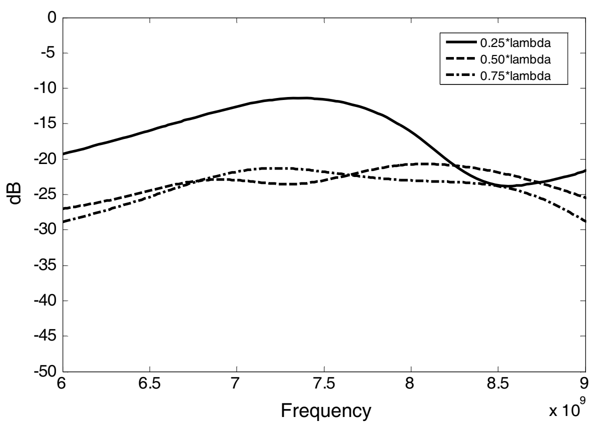


Figure 4: Mutual coupling between E-plane DRAs for several values of the array radius.

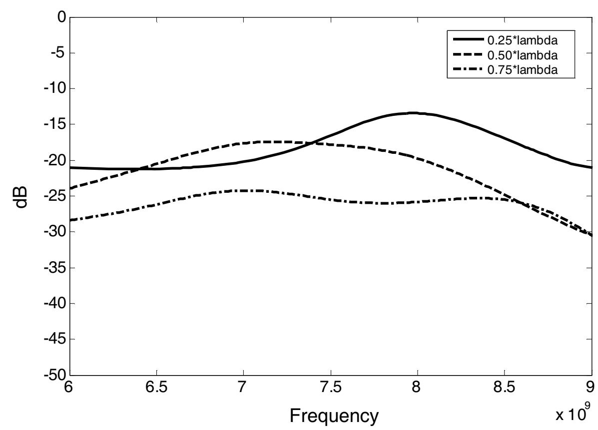


Figure 5: Mutual coupling between staggered DRAs for several values of the array radius.

Mutual coupling between the staggered DRAs, i.e., S_{21}, is computed and illustrated in Figure 5. This figure describes that for D equal to 0.25λ mutual coupling can be as high as $-12\,\mathrm{dB}$ while for the other two values of D, S_{21} is lower. With this explanation it is expected that return loss curve for D equal to 0.25λ should be considerably different from return loss for the other two values of D. The same phenomenon was observed in [6] where a circular staggered DRA was investigated.

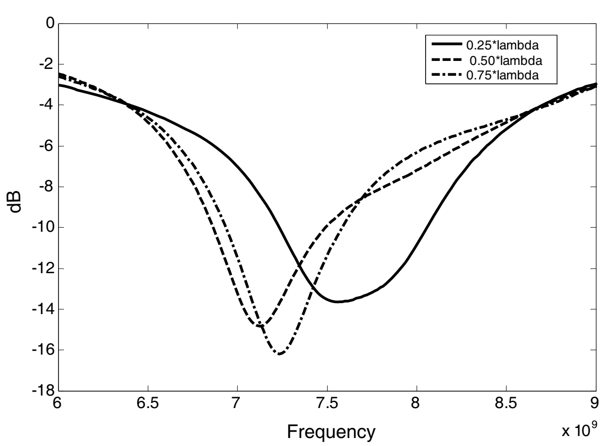


Figure 6: Return loss for several values of the array radius.

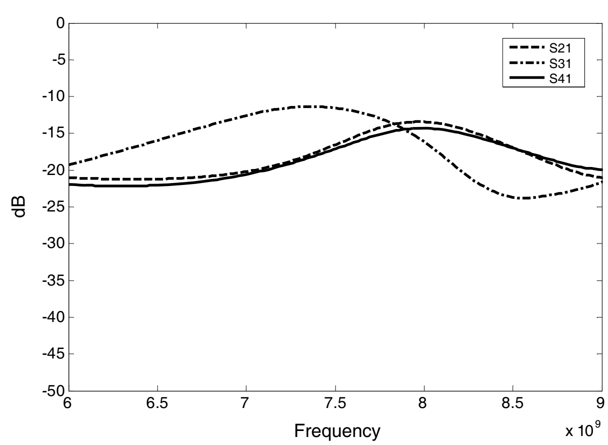


Figure 7: Mutual couplings for D equal to 0.25λ.

For different values of D, return loss is computed and plotted in Figure 6. It is clear from this figure that for D equal to 0.25λ return loss curve is considerably different from the curves corresponding to the other greater values of D. This is because of the high mutual couplings S_{21}, S_{31} and S_{41} for D equal to 0.25λ, as it was explained in previous paragraphs. It is also observed that for this value of D, mutual couplings have resulted in improved bandwidth. For the other values of D equal to 0.5λ and 0.75λ, return loss curves are approximately the same and the same as a single DRA. This phenomenon indicates that mutual coupling for these values of D are not strong enough to affect return loss considerably.

In order to have a comparison among different mutual couplings of the array, in Figure 7, S_{21}, S_{31} and S_{41} are plotted for D equal to 0.25λ. This figure illustrates that, S_{31} is considerable and for frequencies above 7.4 GHz, S_{21} and S_{41} are also significant.

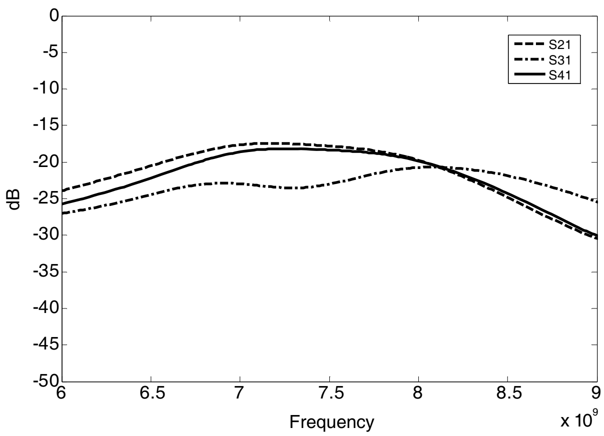


Figure 8: Mutual couplings for D equal to 0.5λ.

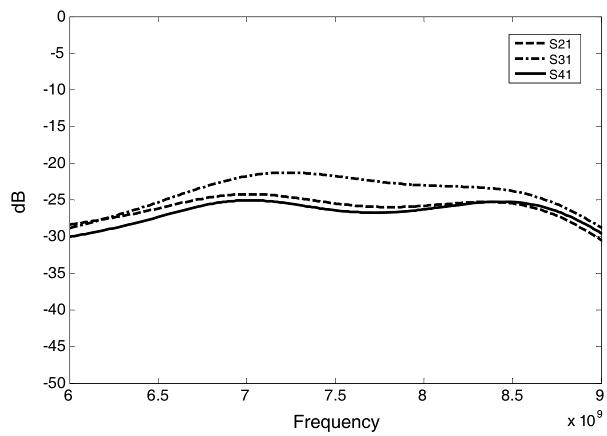


Figure 9: Mutual couplings for D equal to 0.75λ.

Mutual couplings for $\mathrm{D} = 0.5\lambda$ are plotted in Figure 8. It is observed that all mutual couplings are below -18 dB.

Finally mutual couplings for D equal to 0.75λ are shown in Figure 9. This figure demonstrates weak coupling between the arrays' elements for D equal to 0.75λ.

5. CONCLUSION

A four element circular array of rectangular DRAs was investigated. The FDTD method with a UPML boundary condition was used for simulations. For different values of the array radius return loss and mutual couplings were computed and illustrated. It was observed that for D equal to 0.25λ, mutual couplings are high and affect return loss considerably. In this case mutual couplings result in improved bandwidth. For greater values of D mutual couplings are not strong enough to affect the return loss considerably.

ACKNOWLEDGMENT

The support of Center of Excellence on Applied Electromagnetic Systems in preparing this work is gratefully acknowledged.

REFERENCES

1. Ittipiboon, A., R. K. Mongia, Y. M. M. Antar, P. Bhartia, and M. Cuhaci, "Aperture fed rectangular and triangular dielectric resonators for use as magnetic dipole antennas," *Electronics Letters*, Vol. 29, No. 23, 2001–2002, 11 November 1993.
2. Luk, K. M. and K. W. Leung, *Dielectric Resonator Antennas*, Research Studies Press LTD.
3. Luebbers, R. J. and H. S. Langdon, "A Simple feed model that reduces time steps needed for FDTD antenna and microstrip calculations," *IEEE Trans. on Antennas and Propagation*, Vol. 44, No. 7, 1000–1005, July 1996.
4. Taflov, A. and S. C. Hagness, *Computational Electrodynamics: The Finite-difference Time-domain Method*, Artech House, Boston, London, 2005.
5. Gentili, G. B., M. Morini, and S. Selleri, "Relevance of coupling effects on DRA array design," *IEEE Trans. on Antennas and Propagation*, Vol. 51, No. 3, 399–404, March 2003.

6. Jarchi, S., J. Rashed-Mohassel, and M. H. Neshati, "Mutual coupling of DRA's in staggered configuration," *Proceedings of ANTEM/URSI 2006*, 245–248, Montreal, Canada, July 2006.

Wideband Microstrip Antenna Array Using U-slot

Wenwen Chai[1,2], **Xiaojuan Zhang**[1], **and Shuguang Liu**[1,2]

[1]Institute of Electronics, Chinese Academy of Sciences, Beijing 100080, China
[2]Graduate School of Chinese Academy of Sciences, Beijing 100039, China

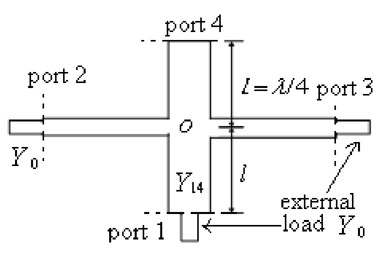
Abstract— A new element antenna array topology of U-slot aperture coupled microstrip patches is introduced. The array is fed by a novel wideband self-adapting matching network and the energy is coupled to the U-slot patches through the feed slot on the ground. The operating frequency range of antenna array is from 10.4 to 16.7 GHz, corresponding to an impedance bandwidth of 46.5%. The gain is above 13 dBi from 11–15 GHz.

1. INTRODUCTION

It is well known that a standard approach to design arrays consists of the design of the array element and a feeding network [1]. The main weakness of an ordinary microstrip element is its narrow bandwidth. There are several ways to overcome this problem. A well known way is based on the introduction of an additional stacked [2, 3] or coupled patch [4, 5]. This makes the configuration more complex. As a consequence the array becomes more costly in production especially in the case of stacked patches. Also the dimension of the array increases. An alternative way consists in the use of specially shaped patches with slits [6, 7] and the rectangular patches with slot [8]. We have considered the U-slotted patch because they have a wider bandwidth and they are simpler in production. For the feeding network, the key components are power dividers. At the present, $\lambda/4$ impedance transformers are used most often to make a power divider. A $\lambda/4$ impedance transformers is well performed at the center frequency, but slight deviation from the center frequency would result in the deterioration of matching. Therefore, conventional feeding network is always narrow-band structure. In this paper, a broadband power divider is analyzed, and based on it, a wideband self-adapting microstrip power-splitting network is made. Fed by this network, the U-slot patch antenna array is fabricated.

2. DESCRIPTION OF FEEDING NETWORK

The model of the broadband power divider is shown in Fig. 1(a). Ports 1, 2 and 3 are all terminated in Y_0, while port 4 is shorted at the end. Characteristic admittance of transmission line between ports 1 and 4 is Y_{14}. Its equivalent transmission line model is shown in Fig. 1(b).



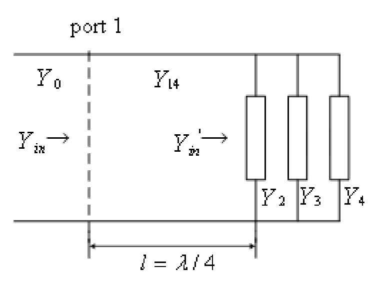


Figure 1: (a) The model of the broadband power divider (b) Equivalent transmission line model.

The equivalent admittances of ports 2, 3, and 4 as seen at the center O are Y_2, Y_3 and Y_4, respectively. The total equivalent admittances as seen at the center O and at port 1 are Y_{in}' and Y_{in}, respectively. The power splitting ratio at output ports 2 and 3 is $1:k$. The total equivalent admittances as seen at the center O can be written as:

$$Y_{in}' = Y_2 + Y_3 + Y_4. \tag{1}$$

Due to short circuit, the equivalent admittance of port 4 as seen at the center O can be expressed as:

$$Y_4 = -jY_{14}\mathrm{ctg}\beta l. \tag{2}$$

Assuming $Y_2 = Y_0$, in terms of the power splitting ratio, we can obtain: $Y_3 = kY_0$. By the use of these parameters and (2), we have:

$$Y'_{in} = (1+k)Y_0 - jY_{14}\mathrm{ctg}\beta l \tag{3}$$

The normalized admittance can be written as:

$$y'_{in} = (1+k) - jy_{14}\mathrm{ctg}\beta l \tag{4}$$

The total equivalent admittance as seen at port 1 can be calculated as:

$$y_{in} = y_{14}\frac{(1+k) + j(\mathrm{tg}\beta l - \mathrm{ctg}\beta l)y_{14}}{2y_{14} + j(1+k)\mathrm{tg}\beta l} \tag{5}$$

The matching at port1 can be achieved at the center frequency only when y_{in} equals to 1. Combining $y_{in} = 1$ and (5), we obtain:

$$y_{14} = \sqrt{1+k} \tag{6}$$

Using (6), y_{in} is rewritten as:

$$y_{in} = 1 + \frac{k - 3 - j[(k-1)\sqrt{1+k}\mathrm{tg}\beta l + 2\sqrt{1+k}\mathrm{ctg}\beta l]}{4 + (1+k)\mathrm{tg}^2\beta l} \tag{7}$$

From (7), it is concluded that when the operating frequencies deviate from the center frequency, βl tends to be $\pi/2$, and therefore ctgβl and tgβl tend to be zero and infinite, respectively. Especially when k equals to 1, the denominator of (7) becomes the high-order term which can guarantee that y_{in} is close to 1 for perfect matching at port 1 in a wide frequency band. Therefore, according to the theoretical analysis, the wideband self-adapting feeding network can be achieved by such series and parallel dividers.

3. DESIGN OF THE ANTENNA ARRAY

Figure 2(a) shows the array element design using a layer of 1.6 mm foam ($\varepsilon_2 = 1.06$) and three substrates of Duroid 5880 ($h = 0.8$ mm, $\varepsilon_1 = \varepsilon_F = 2.2$). The feed strip line is printed on the top of the third piece of Duroid 5880. A large ground reflector is added to greatly alleviate the back-lobe level. The detailed dimensions are shown in Fig. 2(b).

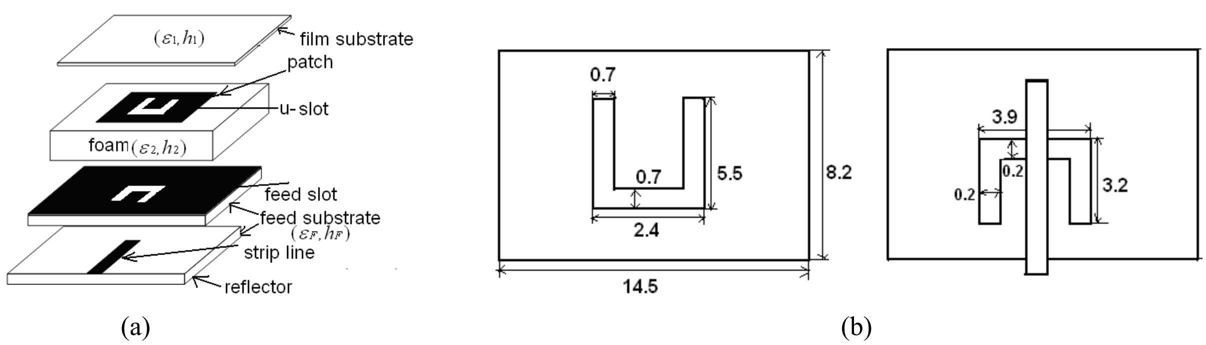


Figure 2: Geometry of U-slot patch antenna. (a) Three-dimensional view (b) Top view.

The geometry and detailed dimensions of the feeding network of planar array are shown in Fig. 3. For the single power divider, $k = 1$, $Z_0 = 50\,\Omega$ are chosen. Z_{14} is 35.4 Ω with a width of 2.12 mm. Characteristic impedances of transmission line between ports 2 and 3 are chosen to be 50 Ω with a width of 1.32 mm. Short circuit at port 4 is realized by soldering a wire from the port to the ground through the substrate. The well matched broadband network are made up of such three power dividers which are connected in a way of hybrid junction and its return loss characteristic curve is shown in Fig. 4 which proves the effectiveness of the theoretical model.

4. RESULTS AND DISCUSSIONS

The standing wave ratio of the antenna array is shown in Fig. 5. The pass-band bandwidth (VSWR <= 2) is about 46.5%. Fig. 6 shows co-polarization gain of the antenna array, which is above 13 dBi from 11–15 GHz. Fig. 7 shows the radiation patterns of the antenna array at 13 GHz.

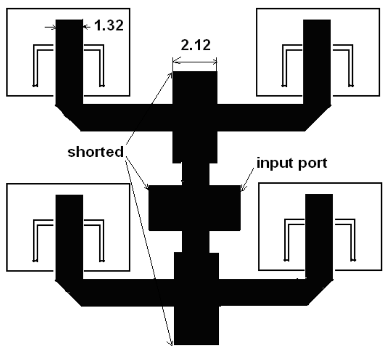


Figure 3: Geometry of feeding network.

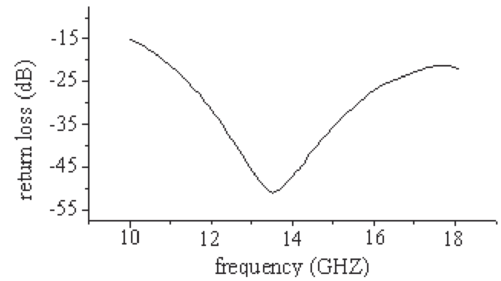


Figure 4: Return loss characteristic of feeding network.

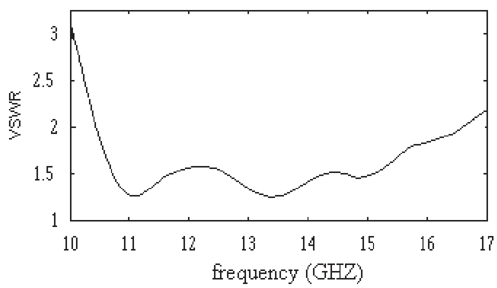


Figure 5: Array bandwidth characteristics.

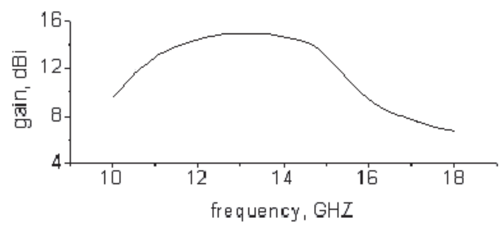


Figure 6: Array gain characteristics.

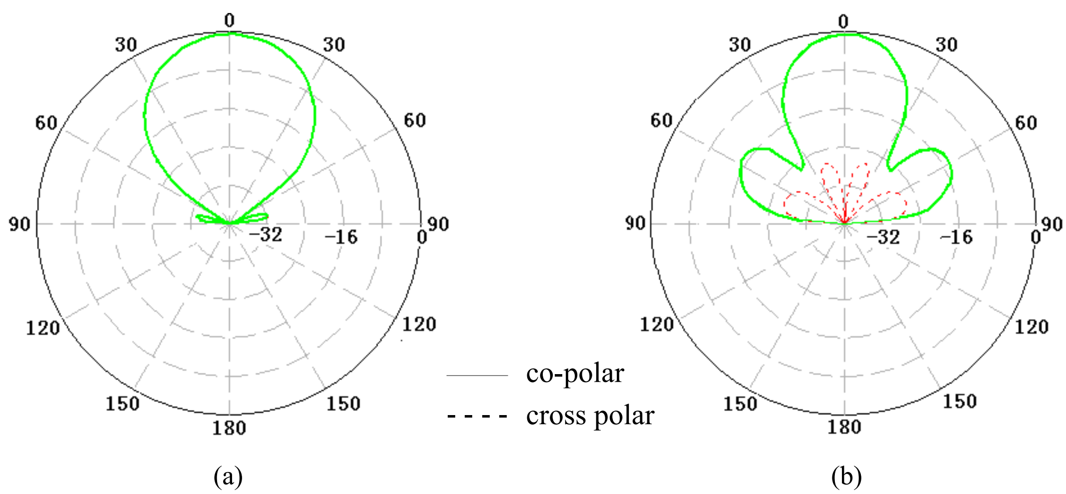


Figure 7: Array radiation patterns at 13 GHz. (a) E plane (b) H plane.

5. CONCLUSIONS

We have successfully designed a wideband self-adapting splitting network fed antenna array composed of four elements of aperture coupled U-slot patch. The antenna has a bandwidth of 46.5%, a maximum gain of 15 dBi and a cross-polarization of less than −25 dB at the center frequency.

REFERENCES

1. Garg, R., P. Bhartia, and I. Bahl, *Microstrip Antenna Design Handbook*, 875, Artech House, 2001.
2. Lee, R. Q., K. F. Lee, and J. Bobinchak, "Characteristics of a two-layer electromagnetically coupled rectangular patch antenna," *Electronic Letters*, Vol. 23, No. 20, 1070–1072, 1987.
3. Croq, F. and D. M. Pozar, "Millimeter-Wave design of wide-band aperture-coupled stacked microstrip antennas," *IEEE Transactions on Antennas and Propagation*, Vol. 39, No. 12, 1770–1776, 1991.

4. Kumar, G. and K. C. Gupta, “Directly coupled multiple resonator wide-band microstrip antenna,” *IEEE Transactions on Antennas and Propagation*, Vol. 33, No. 6, 588–593, 1985.
5. Clenet, M. and L. Shafai, “Wideband single layer microstrip antenna for array applications,” *Electronic Letters*, Vol. 35, No. 16, 1292–1293, 1999.
6. Yang, F., X. X. Zhang, and Y. Samii, “Wide-band E-shaped patch antennas for wireless communications,” *IEEE Transactions on Antennas and Propagation*, Vol. 49, No. 7, 1094–1100, 2001.
7. Wong, K. L. and W. H. Hsu, “A broad-band rectangular patch antennas with a pair of wide slits,” *IEEE Transactions on Antennas and Propagation*, Vol. 49, No. 9, 1345–1347, 2001.
8. Huynh, T. and K. F. Lee, “Single-layer single patch wideband microstrip antennas,” *Electronics Letters*, Vol. 31, No. 16, 1310–1312, 1995.

Numerical and Experimental Analysis of Bow-tie Antennas with Changing of Shield's Height

Kai Zhou[1,2], **Wenji Zhang**[1,2], **and Xiaojuan Zhang**[2]

[1]Graduate University of Chinese Academy of Science, Beijing 100049, China
[2]The Key Laboratory of High Power Microwave and Electro Magnetic Radiation Institute of Electronics Academia Sinica, Beijing 10080, China

Abstract— In this paper the properties of the transmitted waveforms and VSWR of the bow-tie antenna are studied when the shielded height is changed. The results of the simulations and the measurements show agreement well, which can offer the reference to design of the bow-tie antenna.

1. INTRODUCTION

Bow-tie antenna is a transfiguration of the dipole antenna. Bandwidth of the dipole antenna is broadened when the diameter of the wire is broadened. So the double cone antenna, and furthermore, bow-tie antenna as a transfiguration of the double cone antenna, have been introduced. The antenna is designed to work with a widen bandwidth in actual applications. From the transmission line theory when the shield and the terminal resistors with proper value are added to the antenna, some advantages, such as working at the traveling wave state, increasing the radiation efficiency, having better transmitting waveforms and the shield effect, coupling more EM energy into the desired direction, appears. In this paper, the effects of shield's height on antenna's property are studied.

2. STRUCTURE

The planform of the bow-tie antenna is shown in Fig. 1. The angle between two neighbored arms is 30°. In order to achieve the traveling wave state, the loaded resistors are added between the ends of the antenna and the shield. These resistors are used to attenuate the currents coming from the feed points. Both the antenna and the shield are fixed on the floor.

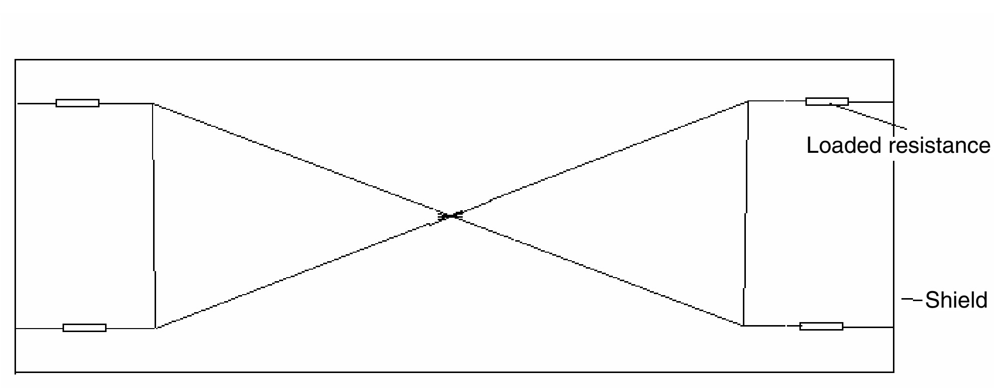


Figure 1: Planform of bow-tie antenna.

When an ideal GPR antenna is excited, it radiates only during the duration of the excitation pulse. But in fact, the antenna continues to radiate even after the excitation pulse has died down. This is due to the internal reflections of the charges that occur at the ends of the bow-tie wings and the excitation points. As a result, antenna ringing, which is undesirable as it corrupts the transmitted pulse waveform and reduces antenna's performance, is formed. The antenna can be resistively loaded by connecting load resistors at the bowtie wing ends to let some of the charges dissipate into the antenna shield and not reflect, thereby suppressing the ringing [1]. The value of the resistance in this paper is 220 Ω.

3. SIMULATION AND TEST

The height of shield has a certain influence on the performance of antenna. In this paper the pattern of antenna with shield height of 8 cm, 15 cm and 18 cm is studied respectively and the result about VSWR of the antenna with shield height of 15 cm is provided in Fig. 3.

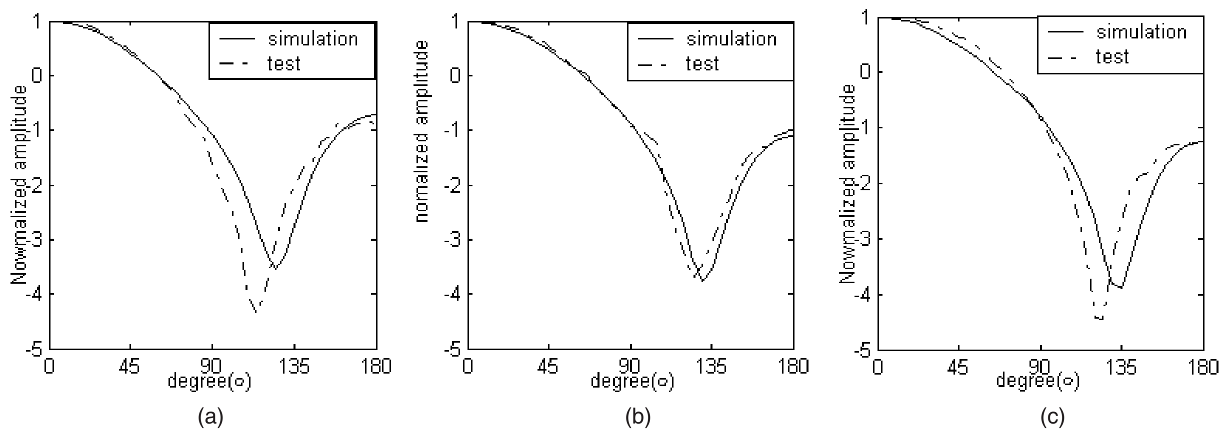


Figure 2: (a) shield height is 8 cm, (b) shield height is 15 cm, (c) shield height is 18 cm.

The central frequency of the antenna is 200 MHz. Fig. 2(a), (b) and (c) show the results of simulation and test to each height of shield. For the symmetry of the antenna, the pattern from 0 to 180 degree only is shown. In Fig. 2. It is found from the figures that the directivity of antenna decreases as the increasing of shield's height and that the beamwidth of the antenna is broadened and sidelobe level is reduced when the height of the shield increases. From Fig. 3 one can find that the VSWR is less than 2 from 100 MHz to 300 MHz, which validate the conclusion that the loaded resistance can attenuate the currents coming from the feed points.

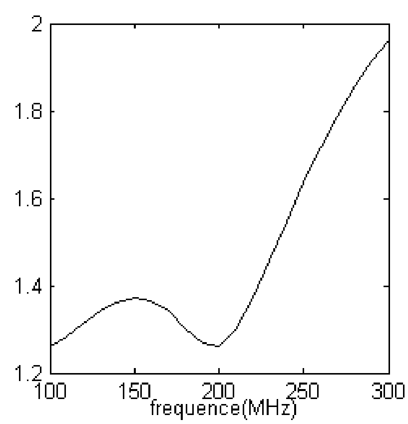


Figure 3: VSWR (shield height is 15 cm).

4. SUMMARY

In this paper, the radiation pattern of bow-tie antenna by changing its shield's height is studied. The simulation results and the tested results are compared, which show basic agreement.

ACKNOWLEDGMENT

The authors thank Ph.D Haitao Hu. Without his help, the authors would not finish this paper successfully. The authors also thank Ph.D Lianlin Li for his reviewing. Supported by the National Natural Science Foundation of China (No. 60371003).

REFERENCES

1. Jayakumar, S., "Antenna optimization for ground penetrating radar using the finite difference time domain technique," Thesis for Master of Science in Electrical Engineering, South Dakota School of Mines and Technology, Rapid city, South Sakota, 1999.
2. Kuchikulla, A., "Design and development of a wideband coherent radar depth sounder," Thesis for Master of Science, Electronics and Communication Engineering, Osmania University, 2001.
3. Kiminami, K., A. Hirata, and T. Shiozawa, "Double-sided printed bow-tie antenna for UWB communications," *IEEE Antennas and Wireless Propagation Letters*, Vol. 3, 2004.

Terahertz Radiation from the Grating Waveguide Loaded Plasma

Wenxin Liu[1], **Ziqiang Yang**[1], **Zheng Liang**[1], **Limei Qi**[1], **Yu Liu**[1]
D. Li[2] **and K. Imasaki**[2]

[1]Institute of High Energy Electronic, University of Electronic Science and Technology of China
Chengdu, Sichuan 610054, China
[2]Institute for Laser Technology, 2-6 yamada-oka, Suita, Osaka 565-0871, Japan

Abstract— A model of generating Terahertz (THz) radiation from a grating waveguide loaded plasma is presented in this paper. The unbounded plasma and electron beam are guided by an infinite magnetic field paralleling the surface of grating. The linear dispersion relation with the effect of plasma is derived by making use of fluid theory, and the interactions of beam-wave with plasma are analyzed by numerical calculations. We find that the operation frequency of beam-wave is enhanced, the enhancement of growth rate is considerable due to the effect of plasma, and there is an optimum plasma density where the growth rate achieves the maximum value.

1. INTRODUCTION

Terahertz sources, a currently active research area, are of importance for a variety of applications in far-infrared spectroscopy, medical and industrial imagings, biomedical research and materials science [1]. At present, much interest has been raised regarding grating-waveguide devices, which is a promising alternative for the development of a tunable, compact, powerful THz radiation source, since Urata et al. [2] and Bakhtyari et al. [3] observed the superradiant Smith-Purcell (SP) emission in the THz regime by an electron beam passing through a single rectangular grating in their experiments. Korbly et al. [4] employed a high-energy bunched electron beam to drive a triangular grating, and observed a coherent radiation at the THz frequency.

To improve the performance of such kind of device, it is necessary to find an efficient mechanism for the beam-wave interaction. Loading plasma is an important physical mechanism that has been successfully used in some high-power microwave systems [5], such as backward wave oscillators, Cerenkov maser, etc, which can be enhanced the interaction efficiency and radiation power. The improved BWO performance may be attribution to the excitation of beam-plasma wave instability and the two-stream instability and an in increase in the drive current of the system. There are also interests to inject plasma into a dielectric Cerenkov maser to study the effect of plasma on the beam-waveguide instability. In this study, we apply loading plasma to a grating-waveguide device, as shown in Fig. 1. A sheet beam is guided by an infinite magnetic field and the device is immersed in the unbounded boundary plasma.

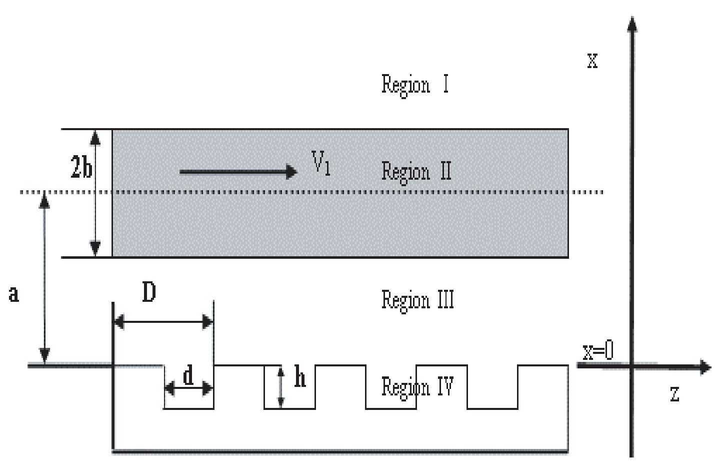


Figure 1: Physical model.

2. THEORY

2.1. Physical Model

The system of beam-wave interaction is shown in Fig. 1. A sheet beam with thickness 2b is located above the rectangular grating, which drifts in the z-direction at constant velocity v_1. It is guided by an infinite magnetic field and immersed in the unbounded boundary of plasma. The quantities D, h, d, and a denote the grating period, groove depth, slot width and distance between the beam center and the surface of grating, respectively. For convenience, the operation area is divided into four regions: $x > a + b$ is region I, $a - b \leq x \leq a + b$ is region II, $0 \leq x < a - b$ is region III and $-h \leq x < 0$ is region IV. To obtain the dispersion equation with plasma, we assume that (i) the external magnetic field is so strong that the perturbed electron motions are restricted in the z-direction; (ii) the effect of beams self static fields are negligible; (iii) the perturbations are uniform along the y-direction, and (iv) the plasma is uniform, cold and the device is immersed in a strong axial magnetic field. In the following, we only concentrate on the TM waves, for which the component of the magnetic field in the z-direction vanishes. Maxwell's equations and relativistic hydrodynamic equations with plasma are employed to describe this system as follows:

$$\nabla \times E + \mu_0 \frac{\partial H}{\partial t} = 0 \tag{1}$$

$$\nabla \times \mu_0 H - \frac{1}{c}\frac{\partial}{\partial t}(\vec{\varepsilon} \cdot E) = \frac{4\pi}{c}\vec{J} \tag{2}$$

$$(\frac{\partial}{\partial t} + v_0 \frac{\partial}{\partial z})\delta v = -\frac{e}{\gamma^3 m_0}E_z \tag{3}$$

$$(\frac{\partial}{\partial t} + v \frac{\partial}{\partial z})\delta n = -n\frac{\partial}{\partial z}\delta v \tag{4}$$

where $\gamma = 1 + eV/m_0c^2$ is the relativistic mass factor, c is the speed of light, V is the voltage of beam, $v_1 = c \cdot [(\gamma^2 - 1)/\gamma^2]^{1/2}$ is the velocity beam, $-e, m_0$, n, δn, δv are the electron charge, the electron rest mass, the unperturbed density, the perturbed density and perturbed velocity, respectively. $\vec{\varepsilon}$ is the dielectric tensor of plasma. The assumption of an infinite axial magnetic field implies that only the nonvanishing terms in the dielectric tensor are along the major diagonal. In the absence of the beam and in the case of a homogeneous and collisionless plasma, the dielectric tensor has the simpler form [6]:

$$\vec{\varepsilon} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \varepsilon_{zz} \end{bmatrix} \quad , \tag{5}$$

where $\varepsilon_{zz} = 1 - (\omega_p/\omega)^2$, $\omega_p = (e^2 n_p/m_0\varepsilon_0)^{1/2}$ is the plasma angular frequency. Using Floquet's theorem, all the fields and beam parameters can be written as a sum of space harmonic:

$$f(x, z) = \sum_{n=-\infty}^{\infty} f_n(x) \exp[j(k_n z - \omega t)], \tag{6}$$

where $k_n = k_0 + 2\pi n/D$, k_0 is the axial wave number, ω is the angular frequency, the n is an integer and labels the nth space harmonics.

2.2. Dispersion Equation

For the slot with d much smaller than the wavelength in free space, the fields in the slot or in region IV of Fig. 1 can be expressed approximately in terms of TEM standing wave, using the boundary conditions of electric field and magnetic filed at the beam surface $x = a \pm b$ and the grating surface $x = 0$, then through the detailed algebraic operation, we can obtain the dispersion relation as follows:

$$\frac{D}{d \cdot \omega/c}\cot(\frac{\omega}{c}\sqrt{\varepsilon_{1,n}}h) + \sum_{n=-\infty}^{\infty} \frac{1}{j\beta_n}\frac{1 + R_{3,n}}{1 - R_{3,n}}\left(\frac{\sin(k_n d/2)}{k_n d/2}\right)^2 = 0, \tag{7}$$

where $R_{3,n} = \frac{(\varepsilon_{2,n}-\varepsilon_{1,n})\exp[2j\beta_n\sqrt{\varepsilon_{1,n}}(a-b)]}{(\varepsilon_{2,n}+\varepsilon_{1,n})+2j\sqrt{\varepsilon_{1,n}\varepsilon_{2,n}}\cot(2\sqrt{\varepsilon_{2,n}}\beta_n b)}$ is the reflection coefficient for the model field at boundary interface $x = a - b$, $\beta_n = (\omega^2/c^2 - k_n^2)^{1/2}$ is the transverse wave number. $\varepsilon_{1,n} = \varepsilon_{3,n} =$

$1-(\omega_p/\omega)^2$ and $\varepsilon_{2,n} = 1 - \frac{\omega_p^2}{\omega^2} - \frac{\omega_b^2}{\gamma^3(\omega - k_n v_1)^2}$ are the dielectric functions at regions I, III and II, respectively.

In the absence of beam and plasma, the dispersion Eq. (7) becomes

$$\frac{D}{d\cdot\omega/c}\cot(\frac{\omega}{c}h) + \sum_{n=-\infty}^{\infty}\frac{1}{j\beta_n}\left(\frac{\sin(k_n d/2)}{k_n d/2}\right)^2 = 0 \tag{8}$$

It should be note that the Eq. (8) has been studied in Refs. [7].

3. NUMERICAL SOLUTIONS

In order to have a numerical appreciation of the effect of the plasma loading, we analyze the dispersion curves in the absence of beam, which is obtained by solving the Eq. (7), as shown in the Fig. 2. The beam line with a voltage of 150 kV is plotted for reference. Obviously, the frequency of the operation point (solid slot), which is the beam-wave intersections, moves up-shifts (from 112 GHz to 119 GHz) as the plasma density increasing. Similar effects are too observed in the absence of plasma grating waveguide either by decreasing the groove depth and width keeping the grating period constant. The introduction of the plasma into the device allows another option and additional way to tune its frequency and output. Moreover, it will be much easier to change the plasma characteristics than to alter the grating parameters.

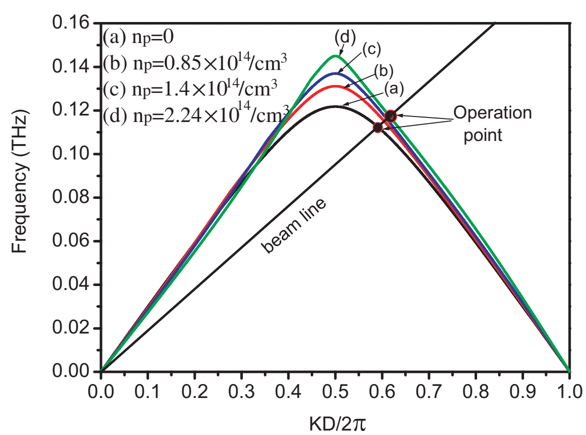


Figure 2: Dispersion Curves. Parameters for grating: Period $D = 1$ mm, groove width $d = 0.4$ mm and groove depth $h = 0.3$ mm. The voltage for beam is 150 kV.

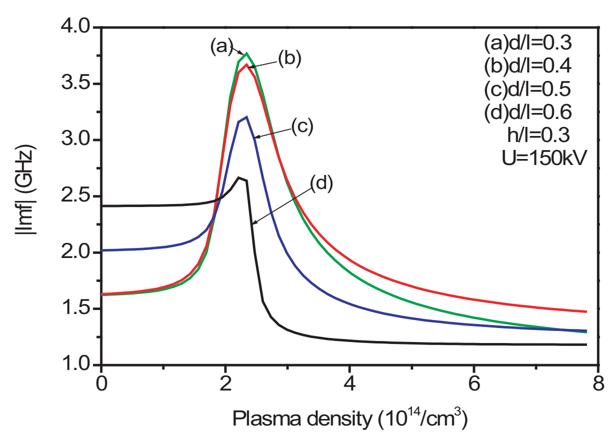


Figure 3: Peak growth rate versus plasma density at the fixed parameter: period $D = 1$ mm, groove depth $h = 0.3$ mm and voltage 150 kV for different groove width (a) $d/l = 0.3$, (b) $d/l = 0.4$, (c) $d/l = 0.5$, and (d) $d/l = 0.6$.

The variations of peak growth rate versus plasma density at the fixed parameters of grating period $D = 1$ mm and groove depth $h = 0.3$ mm. Obviously, the peak growth rate increase with the increasing of plasma density and it achieves a maximum value at an optimum plasma density $n_p = 2.29 \times 10^{14}/\,\text{cm}^3$. It will decrease when the plasma density exceeds the optimum density. It is because that there is a stronger instability of beam-wave interaction at the optimum plasma density, resulting in the enhancements of growth rate; otherwise, the growth rate will be decreased. As observed, the optimum plasma density is similar to keeping a constant with the increasing of grooves width, however, the peak growth rate is decreased from 3.7 to 2.7 GHz as the increasing of groove width. It is thought that the coupling impendence is decreased as the increasing of width.

Figure 3 shows the variations of peak growth rate versus plasma density at the constants of grating period $D = 1$ mm and groove width $d = 0.4$ mm. Similarity, there is an optimum plasma density where the peak growth rate achieves the maximum value, which is increased with the enhancements of groove depth h from 0.3 up to 0.5 mm.

Figure 5 is illustrated the variations of peak growth rate versus plasma density at the fixed of grating parameters: grating period $D = 1$ mm, groove width $d = 0.4$ mm and groove depth $h = 0.3$ mm. Likewise, the peak growth rate is enhanced with the increasing of plasma density. For increasing voltage from 110 up to 180 kV, the peak growth rate becomes less and occurs at

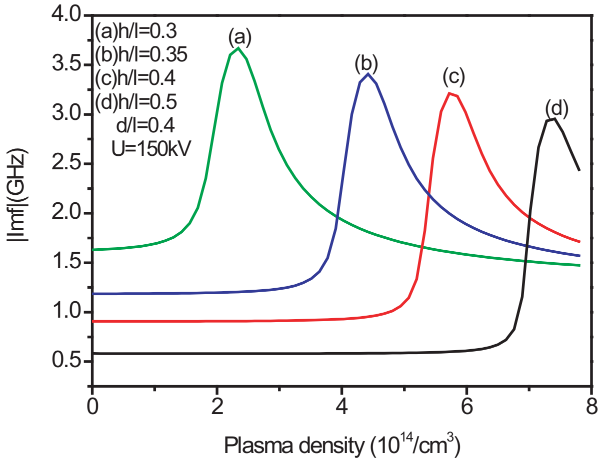


Figure 4: Peak growth rate versus plasma density at the fixed parameter: period $D = 1\,\mathrm{mm}$, groove width $d = 0.3\,\mathrm{mm}$ and voltage $150\,\mathrm{kV}$ for different groove width (a) $h/l = 0.3$, (b) $h/l = 0.35$, (c) $h/l = 0.4$, and (d) $h/l = 0.5$.

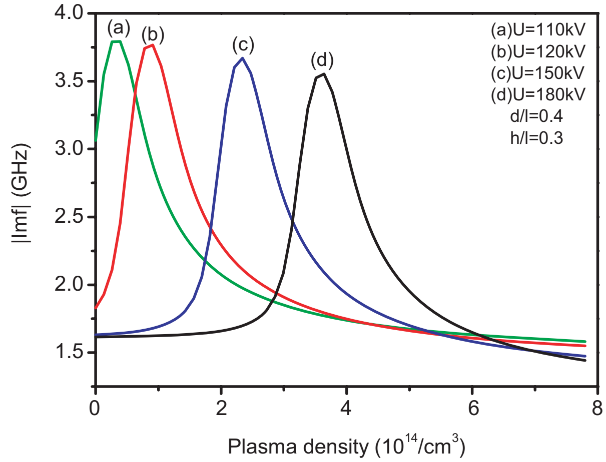


Figure 5: Peak growth rate versus plasma density at the fixed parameter: period $D = 1\,\mathrm{mm}$, groove depth $h = 0.3\,\mathrm{mm}$ and slot width $d = 0.4\,\mathrm{mm}$ for different voltage (a) $U = 110\,\mathrm{kV}$, (b) $U = 120\,\mathrm{kV}$, (c) $U = 150\,\mathrm{kV}$, and (d) $U = 180\,\mathrm{kV}$.

$n_p = 3.61 \times 10^{14}/\,\mathrm{cm}^3$. The increase of peak growth rate with the increment of plasma density may be attributed to the increase of longitudinal electric field, which improves the efficiencies of beam-wave interaction and leading to the enhancements of peak growth rate.

4. CONCLUSIONS

In this study, we present a possible method of THz radiation from grating waveguide loaded plasma. The dispersion equation with plasma is derived and solved by the numerical solutions. The enhancements of operation frequency are observed due to the presence of plasma. The increasing of peak growth rate is observed with the increasing of plasma density, and it achieves the maximum value at an optimum plasma density. The optimum plasma density is similar to keep a constant with increasing of groove width, however, it is increased when the groove depth is enhanced and the operation voltage is increased.

ACKNOWLEDGMENT

This work is supported by National Nature Science of Foundation of P. R. China (60571020).

REFERENCES

1. Siegel, P. H., "Terahertz technology," *IEEE Trans. Microwave Theory Tech.*, Vol. 50, No. 3, 910–928, 2002.
2. Urata, J., M. Goldstein, M. F. Kimmitt, A. Naumov, C. Platt, and J. E. Walsh, "Superradiant Smith-Purcell emission," *Phys. Rev. Lett.*, Vol. 80, No. 3, 516–519, 1998.
3. Bakhtyari, A., J. E. Walsh, and J. H. Brownell, "Amplified-spontaneous-emission power oscillation in a beam-wave interaction," *Phys. Rev. E*, Vol. 65, No. 6, 066503-1–066503-4, 2002.
4. Korbly, S. E., A. S. Kesar, J. R. Sirigiri, and R. J. Temkin, "Observation of frequency-locked coherent terahertz Simth-Purcell radiation," *Phys. Rev. Lett.*, Vol. 94, No. 11, 054803-1–054803-4, 2005.
5. Carmel, Y., K. Minami, R. A. Kehs, et al., "Demonstration of efficiency enhancement in a high-power backward-wave oscillator by plasma injection," *Phys. Rev. Lett.*, Vol. 62, No. 20, 2389–2392, 1989.
6. Zaginaylov, G. I., Yu. V. Gandel, and P. V. Turbin, "Modeling of plasma effect on the diffraction radiation of relativistic beam moving over a grating of finite exent," *Microwave and Optical Tech. Lett.*, Vol. 167, No. 1, 50–54, 1997.
7. Mehrany, K. and B. Rshidaian, "Dispersion and gain investigation of a cerenkov grating amplifier", *IEEE Trans. On Electron Devices*, Vol. 50, No. 61562–1565, 2003.

A Novel Linear Beamforming Algorithm

Shuangning Shi, Yong Shang, and Qinglin Liang
Satellite and Wireless Communication Lab. of Electronic Engineering Department
Peking University, 100871, China

Abstract— Quantum Signal Processing (QSP) beamformer, which is based on Quantum Signal Processing framework, is presented in this paper. This new beamformer can realize the same output gain and signal-to-interference-and-noise ratio (SINR) as conventional beamformers through adjusting impact factors without pre-estimating the covariance matrix of received signals. Finally, the effectiveness and feasibility of the new beamformer are verified by simulation results.

1. INTRODUCTION

Beamforming, which is widely used in the fields of radar [1], communications [2] and sonar [3], is a key technique in array signal processing. It can be classified into two main categories, namely conventional beamforming [2, 4] and blind beamforming [5, 6], depending on the availability of the source location information. Conventional beamforming is widely used since it needs only direction of arrival (DOA) of the interested signal and can be easily implemented. However, estimation of covariance matrix of received signals is indispensable in conventional beamforming, which restricts its application when the number of data samples is limited. A novel linear beamformer, which is defined as QSP beamformer, is presented in this paper by using ideas from Quantum Signal Processing framework [7]. The performance of the new beamformer approaches to that of conventional beamformer, and only the information of DOA is needed in this method. And for the reason that the covariance matrix of received signals does not need to be pre-estimated, the new beamformer can be used in fast fading channels.

The remainder of this paper is organized as follows. In Section 2, basic signal model of array processing and mathematical expressions of the new beamformer are introduced. In Section 3, the simulations of QSP beamformer are carried out and the results are analyzed respectively. Finally, concluding remarks are drawn in Section 4.

2. SIGNAL MODEL AND QSP BEAMFORMER

2.1. Signal Model [8]

Assume that there are J narrow-band stationary signals in far field $\{i_j(t),\, j = 1, 2, \ldots, J\}$, and the DOAs of the signals are $\{\theta_j,\, j = 1, 2, \ldots, J\}$ respectively. At the receiver antenna array is composed of M elements. Suppose that the additive white Gaussian noises at each array element are $\{n_k(t),\, k = 1, 2, \ldots, M\}$ with the same variance σ^2, then the received signal of the kth array element can be denoted as follows:

$$x_k(t) = \sum_{j=1}^{J} a(\theta_j) i_j(t) + n_k(t) \tag{1}$$

And the vector form of the received signals is:

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{n}(t) = \sum_{j=1}^{J} \mathbf{a}(\theta_j) i_j(t) + \mathbf{n}(t) \tag{2}$$

where $\mathbf{a}(\theta_j) = [a_1(\theta_j) \ldots a_M(\theta_j)]$ is the steering vector corresponding to signals with DOAs $\{\theta_j, j = 1, \ldots, J\}$, $\mathbf{x}(t) = [x_1(t) \ldots x_M(t)]^T$, $\mathbf{A} = [(\mathbf{a}(\theta_1) \ldots \mathbf{a}(\theta_J)]^T$, $\mathbf{s}(t) = [i_1(t) \ldots i_J(t)]^T$, and $\mathbf{n}(t) = [n_1(t) \ldots n_M(t)]^T$.

Equation (2) is the matrix form of received signals. According to the principle of adaptive beamforming, the beam pattern can be obtained through adjusting the weight coefficient of each array element. All the weight coefficients form the weight vector $\mathbf{w}$. The output signal of the beamformer is denoted as:

$$y(t) = \mathbf{w}^H \mathbf{x}(t) \tag{3}$$

2.2. QSP Beamformer

As discussed above, the search for optimal weight vector $\mathbf{w}$ is to achieve both a high output gain in the direction of interested signal and a low gain in the direction of interference signals, which means that the interference is suppressed to the maximum extent. Meanwhile, since the signals are always exposed to influence of noises, an ideal beamformer should handle influences and noises effectively at the same time. In order to meet the requirement mentioned above, we set the principle of new beamformer as follows:

Assume that DOA θ_j is given. If we expect that high signal-to-noise ratio (SNR) can be acquired in this DOA, the weight vector $\mathbf{w}(\theta_j)$ of the signal with DOA θ_j should either equal or approach to the steering vector $\mathbf{a}(\theta_j)$ under some constraint. Similarly, the same way can be used to achieve high SNR for other users. Thus we summarize the discussion above as follows: when a group of steering vectors $\{\mathbf{a}(\theta_j),\, 1 \le j \le J\}$ are given, we could look for a set of vectors $\{\mathbf{w}(\theta_j),\, 1 \le j \le J\}$ which equal or approach to $\{\mathbf{a}(\theta_j),\, 1 \le j \le J\}$ with some constraint to achieve high SNR in each DOA of the interested signals.

Meanwhile, the maximum signal-to-interference ratio (SIR) is also expected in beamforming, so any weight vector $\mathbf{w}(\theta_j)$ should be orthogonal to steering vectors $\{\mathbf{a}(\theta_{j'}),\, 1 \le j' \le J,\, j' \ne j\}$ to the greatest extent. And according to the discussion above the weight vectors $\{\mathbf{w}(\theta_{j'}),\, 1 \le j' \le J,\, j' \ne j\}$ should approach to the steering vectors $\{\mathbf{a}(\theta_{j'}),\, 1 \le j' \le J, j' \ne j\}$ to the greatest extent, so $\mathbf{w}(\theta_j)$ should be orthogonal to $\{\mathbf{w}(\theta_{j'}),\, 1 \le j' \le J, j' \ne j\}$ to the greatest extent.

Based on the discussion above and under suitable constraint, the realization of an ideal beamformer can be expressed as follows:

Assume that there are J steering vectors $\{\mathbf{a}(\theta_j),\, 1 \le j \le J\}$ in Hilbert space $\boldsymbol{\mathcal{H}}$, when a group of vectors $\{\mathbf{a}(\theta_j),\, 1 \le j \le J\}$ are given, we can construct a group of vectors $\{\mathbf{w}(\theta_j),\, 1 \le j \le J\}$ whose element is orthogonal to each other and which approach to the given group of vectors $\{\mathbf{a}(\theta_j),\, 1 \le j \le J\}$ under Least-Square constraint as close as possible. Furthermore, we introduce different impact factors $\{q_j,\, 1 \le j \le J\}$ for different signals, which satisfy $q_1 + q_2 + \ldots + q_J = 1$, then the vectors $\{\mathbf{w}(\theta_j),\, 1 \le j \le J\}$ should have the following equation minimum [9]:

$$\varepsilon_{LS} = \sum_{j=1}^{J} q_j \Big\langle \mathbf{a}(\theta_j) - \mathbf{w}(\theta_j), \mathbf{a}(\theta_j) - \mathbf{w}(\theta_j) \Big\rangle \tag{4}$$

Meanwhile subject to:

$$\langle \mathbf{w}(\theta_j),\, \mathbf{w}(\theta_{j'}) \rangle = c^2 \delta_{jj'} \tag{5}$$

where c is a constant value greater than zero, $\delta_{jj'}$ indicates that the vectors should be orthogonal to each other and the constant q_j denotes the impact factor of the j-th signal. After calculation we can get:

$$\mathbf{w}(\theta_j) = c\mathbf{a}(\theta_j)\mathbf{Q}\left((\mathbf{QA}^*\mathbf{AQ})^{1/2} \right)^{\dagger} \quad \{1 < j < J\} \tag{6}$$

where $(\cdot)^{\dagger}$ denotes the Moore-Penrose pseudo-inverse of $(\cdot)$ and $\mathbf{Q}$ is a diagonal matrix $\mathbf{Q} = diag(q_1,\, q_2,\, \ldots,\, q_J)$.

The beamformer defined by Equation (6) is called QSP beamformer in this paper.

According to Equation (6), different performance outputs can be obtained when different impact factors chosen. If the impact factors of all J signals are equal to each other, each signal makes the same contribution to ε_{LS} in Equation (4). Now without loss of generality DOA of the interested signal is θ_j, if its impact factor q_j approximates to 1 and impact factors $\{q_{j'},\, 1 \le j' \le J,\, j' \ne j\}$ from other DOAs equal to each other, the weight vector $\mathbf{w}(\theta_j)$ of the signal from DOA θ_j should approximate to its steering vector $\mathbf{a}(\theta_j)$ more than the weight vectors $\{\mathbf{w}(\theta_{j'}),\, 1 \le j' \le J,\, j' \ne j\}$ of other signals in the solving process of minimizing Equation (4), therefore the interested signal obtains high SNR. Similarly, if the impact factor q_j approximates to 0 and the impact factors $\{q_{j'},\, 1 \le j' \le J,\, j' \ne j\}$ are the same, the vectors $\{\mathbf{w}(\theta_{j'}),\, 1 \le j' \le J,\, j' \ne j\}$ will be much closer to vectors $\{\mathbf{a}(\theta_{j'}),\, 1 \le j' \le J, j' \ne j\}$, that means the weight vector $\mathbf{w}(\theta_j)$ will be approximately orthogonal to vectors $\{\mathbf{a}(\theta_{j'}),\, 1 \le j' \le J,\, j' \ne j\}$ subject to the constraint Equation (5), and high SIR is achieved.

Based on the discussion above, the conclusion can be expressed as follows: the novel beamformer can achieve the optimal output gain and high SINR through configuring different impact factors flexibly under different circumstances.

3. SIMULATION RESULTS AND ANALYSIS

In the computer simulation of this paper, we utilize a four-element Uniform Linear Array (ULA) with the elements separated by half-wavelength interval and assume that the interested signal comes from the direction of 0 degree, and an interference signal comes from the direction of 15 degrees with 20 dB SNR and 0 dB SIR, $c = 1$ in QSP beamformer.

Simulation 1: In this simulation, we focus on the beampatterns of the beamformers. Assume that the estimation of DOA is accurate. 1024 samples are used to estimate the auto covariance matrix of received signals in Minimum Variance Distortionless Response (MVDR) beamformer.

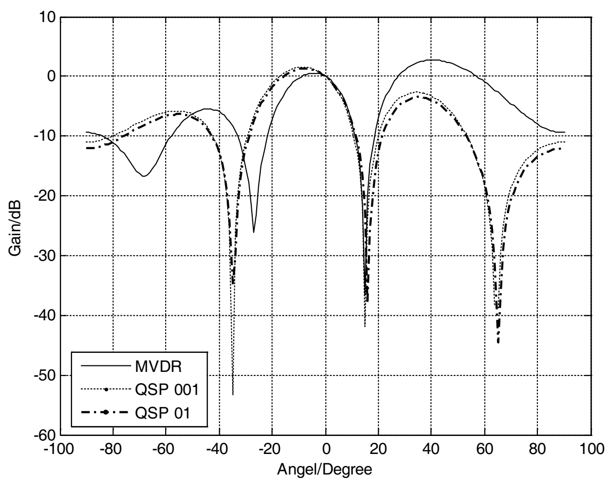


Figure 1: Beampatterns of the beamformers.

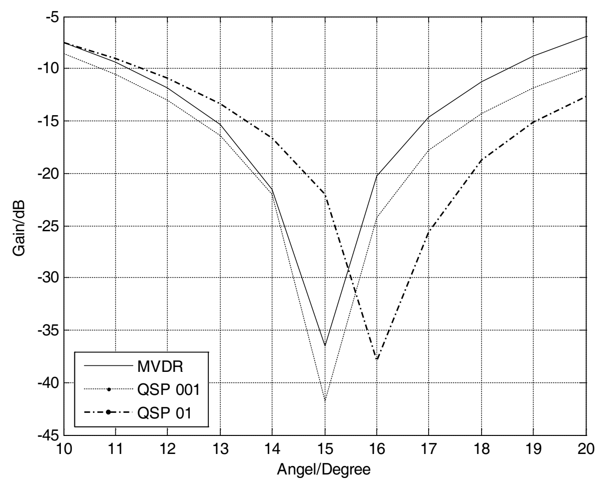


Figure 2: Beampatterns of the beamformers.

Figure 2 is a part of Fig. 1. According to Figs. 1 & 2, we can get: QSP beamformer can not aim at the exact angle of the interrupt signal when q_1 is 0.1 corresponding to curve QSP 01 compared with the curve MVDR or curve QSP 001 which is the output when q_1 is 0.01. That means when the value of q_1 decreases, the performance of QSP will approach to that of MVDR. The result shows the consistency with the analysis above.

With the same environment, the output SINR of the beamformers is presented and compared.

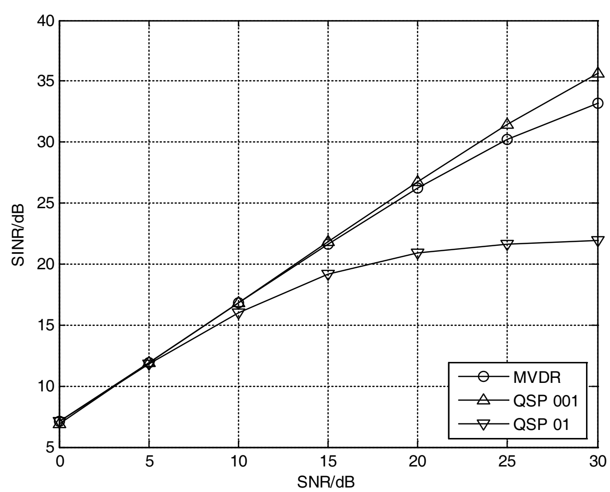


Figure 3: The output SINR of three beamformers.

According to the Fig. 3, as the impact factor of the interested user decreases, the output SINR of QSP beamformer rises gradually. At last its performance surpasses the performance of the MVDR. The result seems hard to accept, but considering the difference between the ideal MVDR outputs with the real MVDR output using 1024 samples, we think the result is acceptable.

4. CONCLUSION

A novel beamformer named QSP beamformer is presented in this paper through introducing the idea of Quantum Signal Processing into the field of beamforming. It can achieve better output performance through adjusting impact factors flexibly under different circumstances when noise's effect is unknown. Compared with the fixed working pattern of conventional beamformers, the new beamformer has more flexibility and stable performance.

ACKNOWLEDGMENT

This work was sponsored by the following project: (1) Project No. 60302006 is supported by National Natural Science Foundation of China; (2) Project No. 60462002 is supported by National Natural Science Foundation of China.

REFERENCES

1. Steyskal, H. and J. F. Rose, "Digital beamforming in radar systems," *Microwave Journal*, 121–136, January 1989.
2. Litva, J. and T. K. Lo, *Digital Beamforming in Wireless Communications*, Artech, Boston, MA, 1996.
3. Thorner, J. E., "Approaches to sonar beamforming," *Southern Tier Technical Conference*, 69–78, 1990.
4. Van Veen, B. D. and K. M. Buckley, "Beamforming: A versatile approach to spatial filtering," *IEEE ASSP Magazine*, Vol. 5, Issue 2, 4–24, April 1988.
5. Cardoso, J. F. and A. Souloumiac, "Blind beamforming for non-Gaussian signals," *Radar and Signal Processing, IEE Proceedings F*, Vol. 140, Issue 6, 362–370, Dec. 1993.
6. Yao, K., R. E. Hudson, C. W. Reed, D. Chen, and F. Lorenzelli, "Blind beamforming on a randomly distributed sensor array system," *IEEE Journal on Selected Areas in Communications*, Vol. 16, Issue 8, 1555–1567, Oct. 1998.
7. Zhang, X. and Z. Bao, *Communication Signal Processing*, The publishing house of the National Defence Industry, Dec. 2000.
8. Eldar, Y. C. and A. V. Oppenheim, *Quantum Signal Processing, IEEE Signal Processing Magazine*, Vol. 19, Issue 6, 12–32, Nov. 2002.
9. Eldar, Y. C. and A. V. Oppenheim, "Orthogonal matched filter detection," *IEEE International Conference on Acoustics, Speech, and Signal Processing*, Vol. 5, 2837–2840, 7–11 May, 2001.

The Performance of QSP Beamformer with Array Errors

Shuangning Shi[1], **Yong Shang**[1], **Qinglin Liang**[1], **and Bin Liang**[2]

[1]School of Electronics Engineering and Computer Science, Peking University, 100871, China
[2]School of Electronics Engineering and Computer Science, Shihezi University, 832003, China

Abstract— There are errors and other uncertain factors in array processing, so the practical steering vectors deviate from the ideal steering vectors, which results in descent of the performance of the conventional beamformer. QSP beamformer is a robust beamformer based on the subspace approaches and it can achieve stable performance. The paper mainly discusses the performance of the QSP beamformer with array errors. The results of the simulation show that the effectiveness and output performance of the QSP beamformer is much better than that of the conventional beamformer.

1. INTRODUCTION

Array signal processing is widely used in the field of radar, communications, senor, medicine and so on. Beamforming [1, 2] is an important application of the array signal processing. Conventional beamforming can choose the optimum weight vectors and achieve the minimum-energy output without loss of the expected signal. In the condition of no error for the array, the conventional beamformer, such as MVDR Beamformer, has strong resolving power and high capability of disturbance rejection. In practical, the steering vector of the expected signal can not be accurately known, and the practical steering vector is not equal to the ideal steering vector, and then the performance of conventional beamformers, especially the output SINR of the array, will be degraded obviously [3–5]. To solve this problem, QSP beamformer, which is based on Quantum Signal Processing framework, is presented in this paper. The beamformer which is based on the subspace approaches is not sensitive to errors in the array and could obtain stable output performance no matter array errors exist or not.

The remainder of this paper is organized as follows. First we propose the basic signal model of array processing and mathematical expressions of the new beamformer. Then the simulations of QSP beamformer are carried out and the results are analyzed respectively. Finally, concluding remarks are drawn.

2. SIGNAL MODEL

Assume that there are J narrow-band stationary signals in far field $\{i_j(t),\ j = 1,\ 2,\ \ldots,\ J\}$, and the DOAs of the signals are $\{\theta_j,\ j = 1,\ 2,\ \ldots,\ J\}$ respectively. At the receiver antenna array is composed of M elements. Suppose that the additive white Gaussian noises at each array element are $\{n_k(t),\ k = 1,\ 2,\ \ldots,\ M\}$ with the same variance σ^2, then the steering vector of the signal whose DOA is θ_j can be denoted as $\mathbf{a}(\theta_j) = [1,\ \exp(-j\pi\sin(\theta_j)),\ \ldots,\ \exp(-j\pi(M-1)\sin(\theta_j))]'$. The received signal of the kth array element can be denoted as follows:

$$x_k(t) = \sum_{j=1}^{J} a(\theta_j) i_j(t) + n_k(t) \tag{1}$$

where K is the number of the snapshots, $k = 1,\ \ldots,\ K$; $\mathbf{n}(t) = [n_1(t),\ \ldots,\ n_M(t)]^T$.

In real environment, there are lots of errors in the receiver, such as the errors caused by array band response and errors which lie in the amplitude-phase characteristics of the elements. The errors make the practical steering vector different with the ideal steering vector. To simplify the condition, we denote all the errors as a Gaussian random vector $\mathbf{a}_e$, which is an addition for the ideal steering vector. The mean of $\mathbf{a}_e$ is 0 and the standard deviation is σ_a^2. So the practical steering vector is $\left[\mathbf{a}(\theta_j) + \mathbf{a}_{e_j}\right]$, and the practical sample data of the receiver after k snapshots can be denoted as

$$\mathbf{x}_k(t) = \sum_{j=1}^{J} \left(\mathbf{a}_k(\theta_j) + \mathbf{a}_{e_j}\right) \mathbf{s}_j(t) + \mathbf{n}_k(t) \tag{2}$$

3. QSP BEAMFORMER

The random vector errors mentioned above is mainly caused by the inaccurate steering vector $\mathbf{a}(\theta_j)$. The QSP beamformer based on signal subspace is proposed in [6] by the author. The new beamformer can overcome the unstable output performance caused by the inaccurate steering vector because it only needs the knowledge of the ideal DOAs.

The main principles of the QSP beamformer can be explained as follows: Assume that there are J steering vectors $\{\mathbf{a}(\theta_j),\ 1 \leq j \leq J\}$ in Hilbert space $\boldsymbol{\mathcal{H}}$, when a group of vectors $\{\mathbf{a}(\theta_j),\ 1 \leq j \leq J\}$ are given, we can construct a group of vectors $\{\mathbf{w}(\theta_j),\ 1 \leq j \leq J\}$ whose element is orthogonal to each other and which approach to the given group of vectors $\{\mathbf{a}(\theta_j),\ 1 \leq j \leq J\}$ under Least-Square constraint as close as possible. Furthermore, we introduce different impact factors $\{q_j,\ 1 \leq j \leq J\}$ for different signals, which satisfy $q_1 + q_2 + \ldots + q_J = 1$, then the vectors $\{\mathbf{w}(\theta_j),\ 1 \leq j \leq J\}$ should have the following equation minimum [7]:

$$\varepsilon_{LS} = \sum_{j=1}^{J} q_j \left\langle \mathbf{a}(\theta_j) - \mathbf{w}(\theta_j),\ \mathbf{a}(\theta_j) - \mathbf{w}(\theta_j) \right\rangle \tag{3}$$

Meanwhile subject to:

$$\left\langle \mathbf{w}(\theta_j),\ \mathbf{w}(\theta_{j'}) \right\rangle = c^2 \delta_{jj'}\ j \neq j' \tag{4}$$

where c is a constant value greater than zero, $\delta_{jj'}$ indicates that the vectors should be orthogonal to each other and the constant q_j denotes the impact factor of the jth signal. After calculation we can get:

$$\mathbf{w}(\theta_j) = c\mathbf{a}(\theta_j)\mathbf{Q}\left((\mathbf{Q}\mathbf{A}^*\mathbf{A}\mathbf{Q})^{1/2} \right)^{\dagger} \quad \{1 < j < J\} \tag{5}$$

where $(\cdot)^{\dagger}$ denotes the Moore-Penrose pseudo-inverse of $(\cdot)$ and $\mathbf{Q}$ is a diagonal matrix $\mathbf{Q} = diag(q_1,\ q_2, \ldots,\ q_J)$.

According to Equation (5), different performance outputs can be obtained when different impact factors chosen. For the signal of which the DOA is θ_j, if the value of q_j is larger, $\mathbf{w}(\theta_j)$ is closer to the steering vector $\mathbf{a}(\theta_j)$, which means we can get high-SNR output after the signal passed the beamformer. This approach is suitable to the high noise situation. If the value of q_j is smaller, $\mathbf{w}(\theta_{j'})$ will be closer to the steering vector $\mathbf{a}(\theta_{j'})$ because $q_{j'}$ gets larger. From Equation (5) we know that $\mathbf{w}(\theta_{j'})$ is orthogonal to $\mathbf{w}(\theta)$, so $\mathbf{w}(\theta_j)$ is nearly orthogonal to the steering vector $\mathbf{a}(\theta_{j'})$, and then we can get high-SIR output after the signal passed the beamformer. This approach is applied under the situation with high disturbance. Therefore the novel beamformer can achieve the optimal output gain and high SINR by configuring different impact factors flexibly under different circumstances.

4. SIMULATION RESULTS AND ANALYSIS

4.1. Simulation Situation

We utilize a four-element Uniform Linear Array (ULA) with the elements separated by half-wavelength interval and assume that the expected signal comes from the direction of 0 degree, and an interference signal comes from the direction of 15 degree, and SIR is 0 dB.The impact factor q_1 is given with 0.1. MVDR beamformer and QSP beamformer are adopted in the following simulation.

Simulation 1: In Fig. 1 curves, MVDR and QSP show the outputs of the two beamformers separately in the situation that the DOA of the signal is accurately known and there are no errors existing in the steering vector. We can see from Fig. 1 that the output performance of MVDR beamformer is good when the DOAs of all signals are stable and accurately known. As the SNR gets higher, the performance of the QSP beamformer is much worse than that of the MVDR beamformer. The curves MVDR R and QSP R in Fig. 1 show the outputs of the two beamformers separately in the situation where the DOAs of the signals have errors. The variance σ_a^2 is 0.01. It is obvious that the performance of MVDR beamformer is rapidly degraded when the errors exit in the steering vectors, especially when the SNR is high. Meanwhile, the performance of the QSP beamformer is only degraded slightly, and it is more stable than the performance of the MVDR beamformer.

Simulation 2: The parameters of this simulation are the same as that in Simulation 1 except that the σ_a^2 changes to be 0.1. In Fig. 2 curves, MVDR R and QSP R show the outputs of the

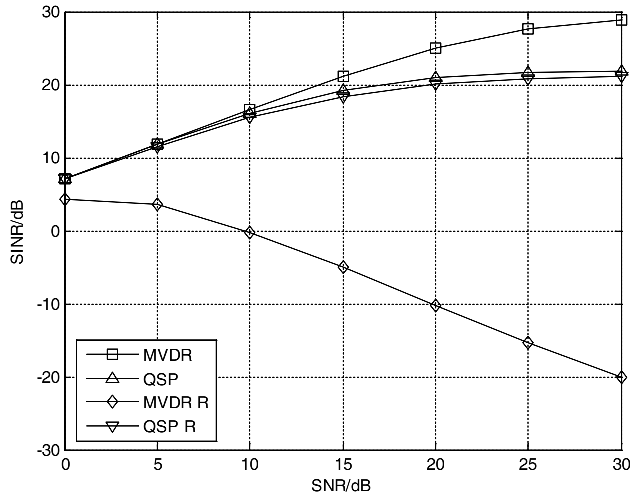


Figure 1: The output SINR of the beamformers.

two beamformers separately in this situation. Compared with Fig. 1, the output SINR of the QSP beamformer is also more stable than that of MVDR beamformer when the variance of the error in the steering vector increases, but the output performance of the QSP beamformer is degraded slightly.

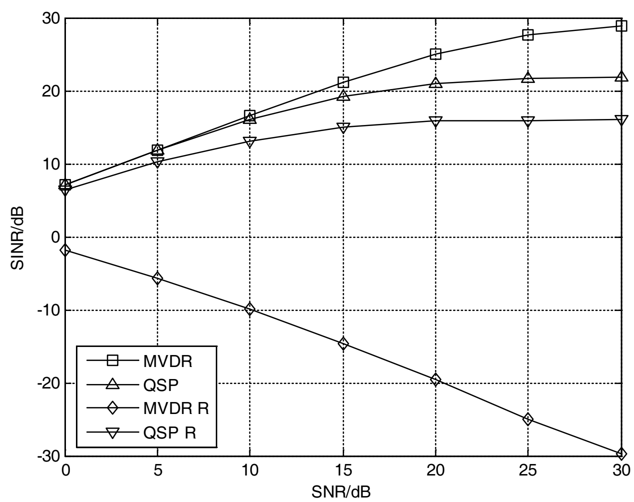


Figure 2: The output SINR of the beamformers.

In a word, in the ideal situation, MVDR beamformer can achieve optimum output performance, while in practical situation various errors make the performance of MVDR beamformer degraded, and some errors, such as random errors of the steering vector , will badly deteriorate the performance of MVDR beamformer. However QSP beamformer can overcome the effect of the errors and achieve stable output performance.

5. CONCLUSION

In real communication system, random errors on the steering vector will badly affect the output performance of conventional beamformer. In this situation, QSP beamformer can achieve stable output performance when the errors are limited in a certain range. In a word the new beamformer is robust and can be applied in the environment mentioned above properly.

ACKNOWLEDGMENT

This work was sponsored by the following project: (1) Project No. 60302006 is supported by National Natural Science Foundation of China; (2) Project No. 60462002 is supported by National Natural Science Foundation of China.

REFERENCES

1. Zhang, X. and Z. Bao, *Communication Signal Processing*, The publishing house of the National Defence Industry, Dec. 2000.
2. Van Veen, B. D. and K. M. Buckley, "Beamforming: A versatile approach to spatial filtering," *IEEE ASSP Magazine*, 4–24, April 1988.
3. Wang, A., Z. Bao, and L. Wang, "Remedies of the estimation errors and perturbation errors," *Journal of Xidian University*, Vol. 26, Issue 5, 614–618, Oct. 1999.
4. Carlson, B. D., "Covariance matrix estimation errors and diagonal loading in adaptive arrays [J]," *IEEE Trans. on AES*, Vol. 24, No. 7, 397–401, 1988.
5. Chou, C. and J. H. Lee, "Robust adaptive array beamforming under steering vector errors [J]," *IEEE Trans. on AP*, Vol. 45, No. 1, 169-175, 1997.
6. Shi, S., Y. Shang, and Q. Liang, "Beamforming based on quantum signal processing," *The 13th Academic Conference of CIE*.
7. Eldar, Y. C. and A. V. Oppenheim, "Quantum signal processing," *IEEE Signal Processing Magazine*, Vol. 19, Issue 6, 12–32, Nov. 2002.

Accurate Defect Mode Modelling in Photonic Crystals Using the Generalised Fictitious Source Superposition Method

L. C. Botten[1], **K. B. Dossou**[1], **S. Wilcox**[2], **R. C. McPhedran**[2], **C. M. de Sterke**[2]
N. A. Nicorovici[1], **and A. A. Asatryan**[1]

[1]CUDOS and University of Technology, Sydney, Australia
[2]CUDOS and University of Sydney, Australia

Abstract— We present an exact theory for modelling modes of generalised defects in 2D photonic crystals with a truly infinite cladding. We build on our fictitious source superposition method for simple defects and outline an elegant extension to model arbitrary defects. Results are presented for 2D and 3D, demonstrating the method's accuracy, efficiency, and ability to handle difficult cases of highly extended modes near cutoff.

While photonic band gaps—frequency ranges in which the propagation of light is suppressed through Bragg reflection—are at the foundation of the technological potential of many photonic crystal (PC) structures [1], it is the introduction of defects into an otherwise periodic medium (e.g., waveguides, cavities etc) that allows this potential to be realised. The introduction of defects gives rise to modes which are localised if their frequencies lie in the band gap of the surrounding PC.

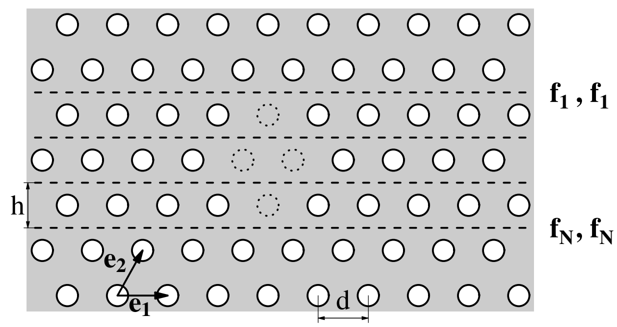


Figure 1: Schematic of the geometry with cylinders of radius a and refractive index n_i in a background of refractive index n_e. which shows a complex defect formed by the removal of four cylinders.

To date, the modelling of defect modes in imperfect periodic structures has been undertaken using techniques that assume a finite structure—either explicitly, or implicitly, as in supercell methods which periodically replicate a finite structure. While such methods work well for strongly confined modes, difficulties arise when the mode becomes extended or poorly confined, such as near cutoff. In such cases, the computational requirements of modelling a sufficiently large structure can be overwhelming or lead to inaccurate results. To handle problems of this type, we have developed an exact theory [2] known as *the fictitious source superposition (FSS) method* for computing defect modes in a genuinely infinite 2D lattice and have applied it to the study of the long-wavelength behaviour [3] of microstructured optical fibres (MOFs). Not only does our approach handle MOFs with an infinite cladding, but also it is computationally more efficient than other techniques when the size of the structure becomes large. While our original treatment [2] of the FSS method was a useful tool that helped resolve the controversy about the existence of a cutoff of the fundamental mode in a MOF [3], it was restricted to modelling only simple defects, i.e., a single defect in a single row of scatterers. The purpose of this paper is to extend the theory to accommodate arbitrary defects comprising compound defects (i.e., with a multiplicity of cylinders removed or modified) in multiple rows of an infinite photonic crystal [4].

The FSS method is based on three essential ideas: the first is the use of a fictitious source to tailor the exterior field so that a scatterer appears to have optical properties that are quite different to those it possesses in reality. To construct a defect, for example, the exterior field of an isolated cylinder must appear as if no radiation emanates directly from the cylinder. While this is straightforward in the case of an isolated cylinder, its extension to an infinite lattice requires the formulation of the defect mode as a superposition of quasiperiodic field problems—the second of our key ideas. To see how this works, we introduce exterior and interior expansions of the field in the vicinity of each scatterer j:

$$V_e = \sum_{n=-\infty}^{\infty} \left[a_n^j J_n(k_e\rho) + b_n^j H_n^{(1)}(k_e\rho) \right] e^{in\theta}, \qquad V_i \sum_{n=-\infty}^{\infty} \left[c_n^j J_n(k_i\rho) + q_n^j H_n^{(1)}(k_i\rho) \right] e^{in\theta} \tag{1}$$

where $(\rho, \theta) = \boldsymbol{r} - \boldsymbol{r}_j$. The $\{\boldsymbol{b}^j\}$ characterise outward radiation from cylinder j, while the $\{\boldsymbol{a}^j\}$ define the regular (nonsingular) field, comprising radiation incident on the cylinder from all exterior sources. Correspondingly, the regular part of the interior field is characterised by the $\{\boldsymbol{c}^j\}$, while the coefficients $\{\boldsymbol{q}^j\}$ represent the fictitious sources used to tailor the external field.

In each cylinder of the lattice, we embed a fictitious source $\boldsymbol{q}^j = \boldsymbol{q}\exp(i\boldsymbol{k}_0 \cdot \boldsymbol{r}_j)$, thus constructing a quasiperiodic field problem from the phased array of sources. The Rayleigh (field) identity $\boldsymbol{a} = \boldsymbol{S}^A\boldsymbol{b}$ then follows by noting that the regular part of the field in the vicinity of any cylinder is due to outgoing radiation from all other cylinders, with their interrelationship encapsulated in the Toeplitz matrix of array lattice sums $\boldsymbol{S}^A = \left[S_{nm}^A\right]$ where $S_{nm}^A = S_{n-m}^A$, with

$$S_l = \sum_{s \neq 0} H_l^{(1)}(k_e|s|d) e^{-il\arg(s)} e^{i\alpha_0 sd} \tag{2}$$

The continuity conditions satisfied by the field at each cylinder interface have the form $\boldsymbol{b} = \hat{\boldsymbol{R}}\boldsymbol{a} + \hat{\boldsymbol{T}}\boldsymbol{q}$ and these together with the field identity allow us to deduce that

$$\boldsymbol{b} = \boldsymbol{Z}\boldsymbol{q} \quad \text{where} \quad \boldsymbol{Z} = \left(\boldsymbol{I} - \hat{\boldsymbol{R}}\boldsymbol{S}^A\right)^{-1} \hat{\boldsymbol{T}}. \tag{3}$$

We may now construct the mode, removing the cylinders at a finite number of lattice points $\boldsymbol{r}_{s_m}$ where $s_m \in \mathcal{S}$, by introducing the compound source $\boldsymbol{q} = \sum_m \boldsymbol{q}_m e^{-i\boldsymbol{k}_0 \cdot \boldsymbol{r}_{s_m}}$ and noting that, in each cylinder, $\boldsymbol{q}^j = \sum_m e^{i\boldsymbol{k}_0 \cdot (\boldsymbol{r}_j - \boldsymbol{r}_{s_m})} \boldsymbol{q}_m$, and $\boldsymbol{b}^j = \sum_m \boldsymbol{Z}(k, \boldsymbol{k}_0) e^{i\boldsymbol{k}_0 \cdot (\boldsymbol{r}_j - \boldsymbol{r}_{s_m})} \boldsymbol{q}_m$. By integrating over the 2D Brillouin zone (BZ), it is clear that the averaged fictitious source for cylinder j $\langle \boldsymbol{q}^j \rangle = 0$, except for those $j \in \mathcal{S}$ for which $\langle \boldsymbol{q}^j \rangle = \langle \boldsymbol{q}^{s_m} \rangle = \boldsymbol{q}_m$. These sources are then available to tailor the outgoing field, thus generating the defect mode by effectively removing all cylinders $j \in \mathcal{S}$. To do so, we set

$$0 = \langle \boldsymbol{b}^{s_l} \rangle = \sum_m \bar{\boldsymbol{Z}}_{lm} \boldsymbol{q}_m \quad \forall s_l \in \mathcal{S}, \quad \text{where} \quad \bar{\boldsymbol{Z}}_{lm} = \langle \boldsymbol{Z}(k, \boldsymbol{k}_0) e^{i\boldsymbol{k}_0 \cdot (\boldsymbol{r}_{s_l} - \boldsymbol{r}_{s_m})} \rangle. \tag{4}$$

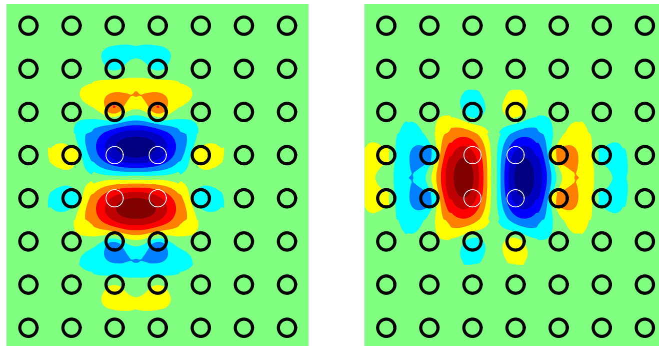
While we may solve this homogeneous system to calculate the wavenumber k of the defect mode, and subsequently construct the mode profile from the null space of $\bar{\boldsymbol{Z}}$, the calculation of the 2D superposition integral is a heavy computational burden, comparable to that experienced in supercell methods. To overcome this problem, we introduce our third key idea, namely a reformulation so as to require only a single integration in the superposition process.

To do this, we conceptualise the structure as a grating stack, introducing an array of fictitious sources into the layers in which defects are to be constructed. The layers containing the fictitious sources are sandwiched between semi-infinite photonic crystals which act as end mirrors since the entire structure is operated in a bandgap. The mirror properties are characterised by a Fresnel reflection matrix $\boldsymbol{R}_\infty$ [5], the use of which is crucial in that it encapsulates the second dimension of the BZ, thus eliminating one integration. Accordingly, following the treatment of Ref. [4] we reformulate the diffraction problem in terms of coupled grating layers and derive an expression

$$\boldsymbol{b} = \boldsymbol{Z}(k, k_{0x})\boldsymbol{q} \tag{5}$$

which relates the multipole source coefficients $\boldsymbol{b}$ to the fictitious sources $\boldsymbol{q}$. The defect mode is then generated by solving the homogeneous system $\langle \boldsymbol{Z}(k) \rangle \boldsymbol{q} = \boldsymbol{0}$, in which the average is computed by an integral over the 1D BZ, $k_{0x} \in [-\pi/d, \pi/d]$. While the theory models an infinite system, the truncation of the BZ integral to a finite sum in the numerical implementation leads to characteristics that resemble those of a supercell method. However, since this supercell is one-dimensional, it can be made arbitrarily large without incurring the computational penalty that afflicts the usual (2D) supercell methods, thus making the method both highly accurate and very efficient.

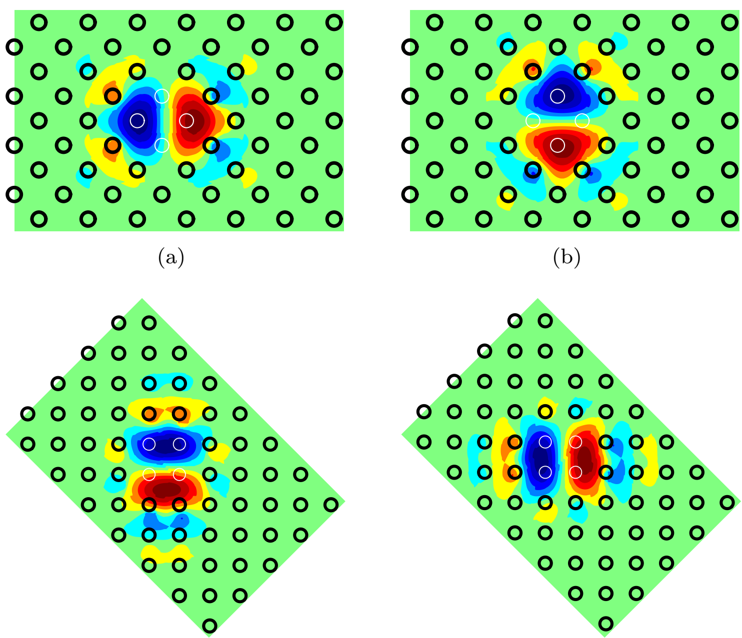
We turn now to the validation and demonstration of the technique, beginning with the calculation of cavity modes in a complex structure, the basic geometry of which derives from a square



(a) (b)

Figure 2: Square cavity in a PC with the grating layers perpendicular to Γ-X. Panels (a) and (b) display field plots for the degenerate pair of modes.

lattice of cylinders of normalised radius $a/d = 0.2$ and refractive index $n_i = 3$ in a background of refractive index $n_e = 1$. From this, we remove four cylinders in a square pattern and compute the modes in two ways: the first by removing two cylinders in two adjacent rows, and a second way in which we regard the structure as a triangular lattice, removing four cylinders that span three rows. Using the first method in the case of $E_{\|}$ polarisation, we find that the structure has two degenerate modes with a resonant frequency of $d/\lambda = 0.396354377$, that are shown in Figs 2(a) and (b).



(a) (b)

(c) (d)

Figure 3: The same cavity as in Fig. 2, except that the grating layers are now perpendicular to $\Gamma - M$. Panels (a) and (b) display fields (v_1 and v_2) for the degenerate pair of modes. For comparison with Figs 2(a), (b), we plot the linear combinations $(v_1 \pm v_2)/\sqrt{2}$ in (c), (d).

We next use the alternative approach of Fig. 3 in which the structure is treated as a triangular lattice with four cylinders spanning three rows removed. While this, of course, must yield the same degenerate pair of modes, the computational scheme is entirely different and, as such, the comparison of these results with those of Fig. 2 is a genuine test of the method's robustness and accuracy. In this case, we find the resonant frequency to be $d/\lambda = 0.396354309$, which is the same as previously, to seven significant figures. The two orthogonal modes computed by the FSS method are shown in Figs. 3(a) and (b), and by taking appropriate linear combinations, we recover in panels 3(c) and (d) the plots of Figs. 2(a) and (b).

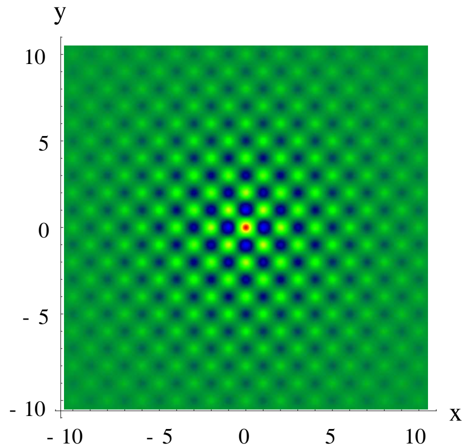


Figure 4: Electric field in an $E_{\|}$ polarised defect mode for a single cylinder defect of radius $a/d = 0.20$ and refractive index $n_{\text{defect}} = 2.7$ in a square symmetric lattice with $a/d = 0.2$ and $n_i = 3$.

While these results validate the method, they do not demonstrate the real strength of the FSS approach in handling the truly difficult cases in which the usual supercell methods fail. In Fig. 4 we show a poorly confined defect mode near the edge of the band gap which, because of its highly extended nature, makes it impossible for standard (supercell) methods to achieve the same accuracy. In this example, we consider the same square lattice as in previous examples but this time reduce the refractive index of the central cylinder to be $n_{\text{defect}} = 2.7$. Using truncated multipole expansions that include harmonics up to order $N_J = 6$, and with plane wave expansions including diffraction orders $-3, -2, \ldots, 2, 3$ [2], a Gaussian integration with $N = 40$ points allows our prototype FSS method, written in *Mathematica* and using *MathLink* to access a Fortran routine to compute the scattering matrices, to find the defect mode of Fig. 4, for a normalised frequency of $d/\lambda = 0.32083778154$, in less than 3 minutes on a standard 3.0 GHz Pentium 4 system running Microsoft Windows XP. Convergence tests have demonstrated that the result is accurate to the 11 significant figures shown. While the mode can be located by a supercell method, such as the commercial RSoft BandSOLVE software in comparable time, it is not possible for it to achieve nearly the same accuracy. A comprehensive discussion of these matters is presented in our earlier paper (Wilcox et al., [2]).

In conclusion, we have discussed a major extension of the Fictitious Source Superposition method so that it can deal with extended defects, rather than just simple point defects. In this extension of the theory, the advantages of the original treatment remain: it allows for the modelling of structures with a genuinely infinite cladding, making the use of a supercell unnecessary. We achieve this by starting from a truly periodic, and hence infinite system, and inserting fictitious sources in the inclusions in order to change the properties of a finite number of them. Though here the fictitious sources were used to remove some inclusions completely, it is equally possible to change the properties of these inclusions more generally, for example changing their refractive index, or radius. Our presentation will outline the theory and implementation of the method and will focus on a range of examples that validate its use and which demonstrate its effectiveness and accuracy. Our presentation will outline the theory and implementation of the method and focus on a range of examples that validate its use and which demonstrate its effectiveness and accuracy.

This work was produced with assistance from the Australian Research Council under its Centres of Excellence Program.

REFERENCES

1. Joannopouls, J. D., R. D. Meade, and J. N. Winn., *Photonic Crystals: Molding the Flow of Light*, Princeton University Press, Princeton, NJ, 1995.
2. Wilcox,'S., L. C. Botten, R. C. McPhedran, C. G. Poulton, and C. M. de Sterke, *Phys. Rev. E*, Vol. 71, 056606, 2005.
3. Wilcox, S., L. C. Botten, C. M. de Sterke, B. T. Kuhlmey, R. C. McPhedran, D. P. Fussell, and S. Tomljenovic-Hanic, *Optics Express*, Vol. 13, 1978–1984, 2005.
4. Botten, L. C., K. B. Dossou, S. Wilcox, R. C. McPhedran, C. M. de Sterke, N. A. Nicorovici, and A. A. Asatryan, "Highly accurate modelling of generalized defect modss in photonic crys-

tals using the FSS method," *Int. J. Microwave and Optical Technology*, (invited paper), Vol. 1, 133–145, 2006.

5. Botten, L. C., N. A. Nicorovici, R. C. McPhedran, A. A. Asatryan, and C. M. de Sterke, *Phys. Rev. E*, Vol. 64, 046603, 2001.

On Local Bianisotropic Metamaterials

E. O. Kamenetskii
Department of Electrical and Computer Engineering, Ben Gurion University of the Negev
Beer Sheva, 84105, Israel

Abstract— In attempts to combine the field locality and chirality in media properties, we come to non-classical aspects, which constitute a subject of a macroscopic quantum mechanics analysis. From classical laws these two notions are in an evident contradiction since manifestation of chirality in electromagnetics is usually due to effects of the field non-locality. Bianisotropics is conceived as a physical concept describing electromagnetic media which possess intrinsic mechanisms of magnetoelectric coupling. We propose the main idea of local bianisotropic metamaterials as a combination of two physical notions: (a) the near-field manipulation and (b) chirality.

1. INTRODUCTION

The concept of a bianisotropic medium was coined in 1968 by Cheng and Kong [1] defining a medium with the most general linear constitutive relations. In microwaves, bianisotropic media are conceived as artificial structures. The fact that small metallic resonant particles may show certain ME-like effects is well-known. It constitutes the subject of numerous theoretical and, not so numerous, experimental papers on microwave composite chiral and bianisotropic materials based on a collection of special-form metallic particles (see e.g., [2]). It is supposed that since particulate composite media are synthesized by artificially fabricated inclusions in a host medium, the designer is provided with large collections of necessary parameters: induced electric and magnetic moments with cross-polarization coupling. In an assumption that the inclusion sizes and spacings are small compared to the wavelength, a typical formal way of derivation of the macroscopic constitutive relations for bianisotropic composites is based on the quasistatic theories of polarization (see e.g., [3]).

It should be clear, however, that any quasistatic theories (similar, for example, to the quasistatic Lorentz theory used for artificial dielectrics [4]) are not applicable for metallic-inclusion bianisotropic composites. Actually, in a classical situation the "ME coupling" in composite media appears due to effects of non-local electromagnetic scattering from inclusions, but not because of the near-field cross-polarization effects. As it is discussed in [5], for interaction of the wave with such a "ME object" not just only the fields but also the various field gradients should play a role. It is stated in [6] that the separation between the macroscopic and microscopic electromagnetic descriptions is not quite as sharp in known artificial bianisotropic media as it is in pure dielectrics due to the fact that the cross-polarization coupling vanishes in the long-wavelength limit.

If one supposes that he has created an "artificial atom" with the local cross-polarization effect one, certainly, should demonstrate a special ME field in the near-field region. It means that using a gedankenexperiment with two quasistatic, electric and magnetic, point probes for the ME near-field characterization, one should observe not only an electrostatic-potential distribution (because of the electric polarization) and not only a magnetostatic-potential distribution (because of the magnetic polarization), one also should observe a special cross-potential term (because of the cross-polarization effect). This fact contradicts to classical electrodynamics. One cannot consider (classical electrodynamically) two coupled electric and magnetic dipoles — the ME particles — as local sources of the electromagnetic field [7]. So in a presupposition that an "artificial atom" with the near-field cross-polarization effect is really created, one has to show that in this particle there are special internal dynamical motion processes different from the classical motion processes.

Putting forth a physical concept of an electromagnetic composite material which possesses an intrinsic mechanism of local magnetoelectric coupling, we propose the main idea of bianisotropics as a combination of two physical notions: (a) the near-field manipulation and (b) chirality. From classical laws these two notions are in an evident contradiction since manifestation of chirality in electromagnetics is usually due to effects of the field non-locality. One comes to conclusion that the unified ME fields originated by local ME particles should appear (in the near-field region) with the symmetry properties distinguishing from that of the electromagnetic fields. This opens a new field of studies, which we call as the problem of local bianisotropic metamaterials (LBMMs).

2. NEAR-FIELD STRUCTURES

One of the main aspects attracted the concept of metamaterials was a possibility for the near-field manipulation [8]. In such a sense, metamaterials can be characterized as structures with tailored electromagnetic response. For bianisotropics, understanding of physics of the near field plays a very important role. Usually, in electrodynamics the following classification of the EM fields is used: the far-field EM fields are considered as the propagating waves and the near-field EM fields — as the evanescent (exponentially decaying) modes. The importance of phenomena involving evanescent electromagnetic waves has been recognized over the last years. The fact that evanescent waves are more confined than the single tone sinusoid waves and hence contain wider range of spatial frequencies indicates that it may be possible to have no theoretical limit of resolution for the near-field patterns. At present, the near-field manipulation becomes an important factor in new applications, such as near-field microscopy and new material structures. From a general point of view, the near-field can be defined as the extension outside a given structure (sample) of the field existing inside this structure (sample). Physically, there can be distinguished different categories of the near-fields [9]. For our purpose, we will analyze three types of the near-field EM structures: (a) Helmholtz-equation evanescent modes; (b) Laplace-equation quasistatic fields; and (c) quasistatic-oscillation fields. The quasistatic-oscillation near-field structures are the most important in our consideration.

3. MAXWELL EQUATIONS AND BIANISOTROPICS

Usually, a bianisotropic medium is classified as the medium characterized by constitutive relations:

$$\vec{D} = \overleftrightarrow{\varepsilon} \cdot \vec{E} + \overleftrightarrow{\xi} \cdot \vec{H}, \qquad \vec{B} = \overleftrightarrow{\zeta} \cdot \vec{E} + \overleftrightarrow{\mu} \cdot \vec{H}, \tag{1}$$

where $\overleftrightarrow{\varepsilon}$, $\overleftrightarrow{\mu}$, $\overleftrightarrow{\xi}$, and $\overleftrightarrow{\zeta}$ are the medium dyadics. From classical electrodynamics point of view, this medium is considered as the most general and exotic medium. It is supposed that studies of this medium provide deep insight into the nature of macroscopic Maxwell's equations, in general, and into the EM-field propagation mechanism, in particular. The question, however is: How one gets constitutive relations (1)? One can postulate these relations, as a case of the most general medium, and just only see for himself if such relations do not violate the Maxwell equations. This way of proof really works and the main physical model is based on a simple assumption that the medium is a structure composed with particles of intrinsically coupled electric and magnetic dipoles. Certainly, in his book, Post writes [10]: "Macroscopically, there is no reason to assume that these [electric and magnetic] dipoles cannot be glued together, pairwise in an almost rigid manner". As it is stated by Gronwald et al. [11], the constitutive relations for a general linear ME medium have to be just postulated as an axiom of classical electrodynamics. At the same time, it is pointed out [11] that derivation of these constitutive relations should be obtained after an averaging procedure from a microscopic model of matter. It is unclear, however, what kind of such a model is presumed by the authors.

As we all know, electromagnetic fields in a medium arise from the microscopic Maxwell equations written for the microscopic electric and magnetic fields, microscopic electric charge density and microscopic electric current density. A theory of electromagnetic processes in media is called as macroscopic electrodynamics or electrodynamics of continuous media. This is a phenomenological theory. It makes sense for average quantities: average positions and velocities of particles, which constitute a medium, average electric and magnetic fields. As a result of derivation of macroscopic Maxwell's equations one introduces the quantities: $\vec{D} \equiv \vec{E} + 4\pi\vec{P}$ and $\vec{H} \equiv \vec{B} - 4\pi\vec{M}$, where $\vec{P}$ is (electric) polarization and $\vec{M}$ is magnetization. There is, however, another way to derive the macroscopic Maxwell's equations. Suppose that there exist both the electric and magnetic microscopic charges. For artificial structures composed by local elements one can formally use such notions of "microscopic" magnetic charges and "microscopic" magnetic currents. For such a case, one easily obtains the macroscopic Maxwell equations from the microscopic equations, but physical meaning of the averaged quantities is not the same as in the above case. As a result of derivation of macroscopic Maxwell's equations one introduces now the quantities $\vec{D} \equiv \vec{E} + 4\pi\vec{P}^e$ and $\vec{B} \equiv \vec{H} + 4\pi\vec{P}^m$, where $\vec{P}^e$ and $\vec{P}^m$ are the electric and magnetic polarizations. The motion equations for both above cases are local equations: the average procedure for microscopic current densities takes place in scales much less than a scale of variation of any macroscopic quantity. An important thing is that no magnetoelectric couplings on the microscopic level are assumed in these motion equations. No classical laws describe interaction between linear electric and magnetic currents. The symmetry properties of magnetic charge and current densities under both spatial inversion and time reversal are opposite to those of electric charge and current densities.

In spite of the fact that there are no classical models on the microscopic level, which demonstrate the ME effect in media, on the macroscopic level of consideration, however, one can formally obtain the ME coupling in a frame of a classical description. It can be shown that biansotropic constitutive relations appear as the macroscopic effect of non-locality (effect of field space derivatives). In particular, from a formal analysis of average electric current $\langle \rho \vec{v} \rangle^e$ (a polar vector) and average magnetic current $\langle \rho \vec{v} \rangle^m$ (an axial vector), one can easily obtain the following relations:

$$\langle \rho \vec{v} \rangle^e = \frac{\partial \vec{P}^e}{\partial t} + c \overleftrightarrow{\delta} \cdot \nabla \times \vec{E} \tag{2}$$

and

$$\langle \rho \vec{v} \rangle^m = \frac{\partial \vec{P}^m}{\partial t} + c \overleftrightarrow{\nu} \cdot \nabla \times \vec{H}, \tag{3}$$

where $\overleftrightarrow{\delta}$ and $\overleftrightarrow{\nu}$ are pseudotensors, which are dependent on the medium properties. The second terms in right-hand sides of these equations show the non-locality effect. Based on Eqs. (2) and (3) and taking into account Maxwell's equations, one may obtain finally constitutive relations in a form of Eqs. (1).

4. LOCAL MAGNETOELECTRIC PARTICLES

The concept of bianisotropics, coined by Cheng and Kong [1], was aimed, in particular, to unify two separate branches of research on moving media and ME crystals. One may paraphrase this concept of bianisotropics in other words: possible unification of electromagnetic processes of dipole motions and symmetry breaking phenomena. It can be supposed that the unified ME fields originated from a point ME particle (when such a particle is created) will not be the classical fields, but the quantum (quantum-like) fields. This means that the motion equations inside a local ME particle should be the quantum (quantum-like) motion equations with special symmetry properties. The fundamental discrete symmetries of parity (P), time reversal (T) and charge conjugation (C), and their violations in certain situations, are central in modern elementary particle physics, and in atomic and molecular physics. As a basic principle, the weak interaction is considered as the only fundamental interaction, which does not respect left-right symmetry. Atoms are chiral due to the parity-violating weak neutral current interaction between the nucleus and the electrons. In crystalline solids, ME effect presumes symmetry breakdown in a structure of a medium. ME interactions with mutually perpendicular electric and magnetic dipoles in crystal structures arise from toroidal distributions of currents and are described by anapole moments [12].

Till now, no necessary microscopic justifications of local ME particles — structural elements of LBMMs — have been done, however. No ME couplings on the microscopic level are assumed in classical motion equations: no helical loops (recursion motions) are possible for bound electric charges and no interaction between linear electric and magnetic currents takes place. The dielectric polarization is parity-odd and time-reversal-even. At the same time, the magnetization is parity-even and time-reversal-odd. These symmetry relationships make questionable an idea of simple combination of two (electric and magnetic) dipoles to realize local ME particles for local bianisotropic metamaterials.

Following the results of recent studies, we can state now that spectral properties of magnetic-dipolar modes (MDM) in ferrite disks may put us into a proper way in realizing local ME particles. It was shown that magnetic dipole motion processes in a normally magnetized ferrite disk are characterized by handedness properties [13]. In microwaves, ferrite resonators with multi-resonance MDM [or magnetostatic (MS)] oscillations may have sizes two-four orders less than the free-space EM wavelength at the same frequency. A ferrite disk with the MDM oscillating spectra is a mesoscopic system in a sense that such a system is sufficiently big compared to atomic and lattice scales, but sufficiently small that quantum mechanical phase coherence is preserved around the whole sample. For a case of a MDM ferrite disk one has the quantized-like oscillating system which preserves the coherence. Our studies show a macroscopic effect of quantum coherence for MDM oscillations in normally magnetized thin-film ferrite disks [14].

In recent experiments it was shown that MDM oscillations in a normally magnetized ferrite disk are strongly affected by a normal component of the external RF electric field [15]. Since the RF electric field does not change sign under time inversion, the eigen electric moment should also be characterized by the time-reversal-even properties. One has special symmetry properties of the anapole moment [polar (electric) symmetry, i.e., the parity-odd, and time-reversal-even symmetry]. There are surface magnetic currents (resulting in the anapole moments) in a ferrite disk which appear due to symmetry breaking for magnetic-dipolar oscillating modes [13, 16]. This may explain the observed ME effect in ferrite disks with surface electrodes [17]. One can see that in a case of a ferrite disk + wire particle there are the ME modes characterizing by oscillations in metal-wire and ferrite-disk subsystems [17]. An experimental object (a ferrite disk + wire particle) can be modeled as a triple of vectors: an axial vector of a magnetic bias field $\vec{H}_0$, a polar vector of an electric moment $\vec{\mathrm{p}}^e$, and an axial vector of a magnetic moment $\vec{\mathrm{p}}^m$. Evidently, this triple of vectors is not invariant under the classical PT transformation but is invariant under the non-classical CPT transformation (Fig. 1).

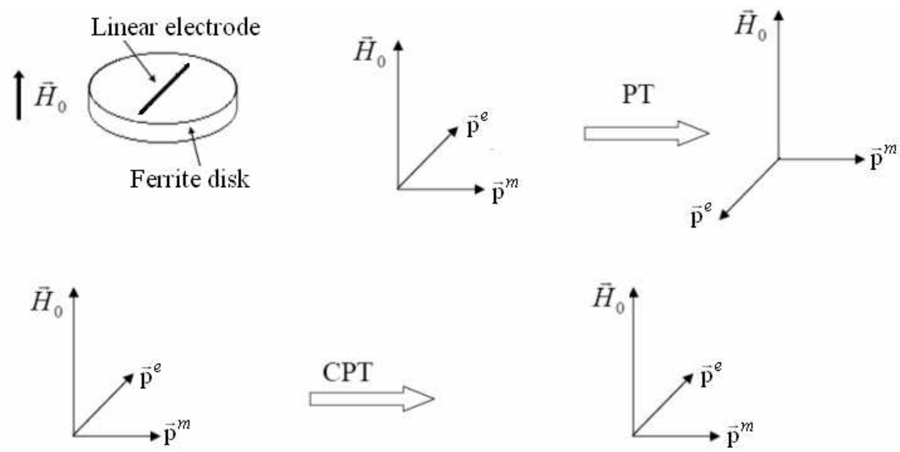


Figure 1: The CPT invariance of a system of two axial ($\vec{H}_0$, $\vec{\mathrm{p}}^m$) and one polar ($\vec{\mathrm{p}}^e$) vectors in a ferrite-disk local ME particle.

5. CONCLUSION

The problem of LBMMs is, in fact, the problem of unification of fundamental notions: electromagnetism and magnetoelectricity. This presumes possible unification of electromagnetic processes of dipole motions and symmetry breaking phenomena. Comprehensive solution of such a problem should be found through understanding a basic mechanism of a "junction" of electricity and magnetism in a point source — a local magnetoelectric particle.

An assumption of the local cross-polarization effect will inevitably lead to a special ME field surrounding a ME particle. This may follow from a simple consideration of free-space Maxwell equations when external sources are taken to be local dipoles $\vec{\mathrm{p}}^e$ and $\vec{\mathrm{p}}^m$. In this case solutions of Maxwell equations are

$$\vec{E}\left(\vec{r},\,\omega\right) = \left[k_0^2\vec{\mathrm{p}}^e + \vec{\nabla}\left(\vec{\mathrm{p}}^e \cdot \vec{\nabla}\right) + ik_0\nabla\times\vec{\mathrm{p}}^m\right]\frac{e^{ik_0|\vec{r}-\vec{r}_0|}}{|\vec{r}-\vec{r}_0|}, \tag{4}$$

$$\vec{H}\left(\vec{r},\,\omega\right) = \left[k_0^2\vec{\mathrm{p}}^m + \vec{\nabla}\left(\vec{\mathrm{p}}^m \cdot \vec{\nabla}\right) - ik_0\nabla\times\vec{\mathrm{p}}^e\right]\frac{e^{ik_0|\vec{r}-\vec{r}_0|}}{|\vec{r}-\vec{r}_0|}, \tag{5}$$

where $k_0 = \omega/c$. In these equations, no coupling between local dipoles $\vec{\mathrm{p}}^e$ and $\vec{\mathrm{p}}^m$ is presupposed. So no ME coupling between electric and magnetic fields can be presupposed as well. When one accepts experimental data of local ME coupling in ferrite disks with surface electrodes shown in [17], he certainly should take into consideration the symmetry properties analyzed in Fig. 1. This evidently results in non-classical symmetry properties of local ME fields.

When a local ME particle is placed inside a cavity, quantization of electromagnetic fields takes place [17]. Since the near field of a ME particle has symmetry breaking, the mapping of the ME-particle field into the cavity EM field will lead to symmetry breaking of the last one.

REFERENCES

1. Cheng, D. K. and J. A. Kong, *Proc. IEEE*, Vol. 56, 248, 1968; *J. Appl. Phys.*, Vol. 39, 5792, 1968.
2. *Advances in Complex Materials*, Eds. A. Priou, A. Sihvola, S. Tretyakov, and A. Vinogradov, Kluwer Academic Publ., Dordrecht, Boston, London, 1997.
3. Ishimaru, A., et al., *IEEE Trans. Antennas Propag.*, Vol. 51, 2550, 2003.
4. Collin, R. E., *Field Theory of Guided Waves*, IEEE, New York, 1991.
5. Raab, R. E. and O. L. De Lange, *Multiplole Theory in Electromagnetism*, Oxford, 2005.
6. Kamenetskii, E. O., *Phys. Rev. E*, Vol. 58, 7965, 1998.
7. Jackson, J. D., *Classical Electrodynamics*, 2nd ed., Wiley, New York, 1975.
8. Pendry, J. B., *Phys. Rev. Lett.*, Vol. 85, 3966, 2000; *Optics & Photonic News*, 33–37, September 2004.
9. Girard, C., et al., *Rep. Prog. Phys.*, Vol. 63, 893, 2000.
10. Post, E. J., *Formal Structure of Electromagnetics*, North-Holland, Amsterdam, 1962.
11. Gronwald, F., et al., *arXiv: physics*/0506219, 2005.
12. Carra, P., *J. Phys. A: Math. Gen.*, Vol. 37, L183, 2004.
13. Kamenetskii, E. O., *Phys. Rev. E*, Vol. 73, 016602, 2006; *J. Magn. Magn. Mater.*, Vol. 302, 137, 2006.
14. Kamenetskii, E. O., M. Sigalov, and R. Shavit, *J. Phys.: Condens. Matter*, Vol. 17, 2211, 2005.
15. Kamenetskii, E. O., A. K. Saha, and I. Awai, *Phys. Lett. A*, Vol. 332, 303, 2004.
16. Kamenetskii, E. O., *Europhys. Lett.*, Vol. 65, 269, 2004.
17. Saha, A. K., E. O. Kamenetskii, and I. Awai, *J. Phys. D: Appl. Phys.*, Vol. 35, 2484, 2002.

A Bianisotropic Left-handed Metamaterials Compose of S-ring Resonator

Xiangxiang Cheng[1,3], **Hongsheng Chen**[1,3], **Lixin Ran**[1,2]
Bae-Ian Wu[3], **Tomasz M. Grzegorczyk**[1,3], and **Jin Au Kong**[1,3]

[1]The Electromagnetics Academy at Zhejiang University, Zhejiang University, Hangzhou 310058, China
[2]Department of Information and Electronic Engineering, Zhejiang University, Hangzhou 310027, China
[3]Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Abstract— A realization of a kind of bianisotropic left-handed material based on the structure of S-ring resonator is proposed in this paper. By subtly shorting some horizontal metallic strips inside of the ring, magnetoelectric coupling is raised, leading to a bianisotropic effect in this artificial medium. The structure is theoretically analyzed from a microscopic viewpoint and its macroscopic constitutive parameters are presented, where a nonzero chirality component is shown. Due to the existence of the bianisotropy, we theoretically show the existence of cross polarization effects when wave incident onto a slab of this material, and numerical simulations on the real structure confirms the results predicted by our theoretical analysis.

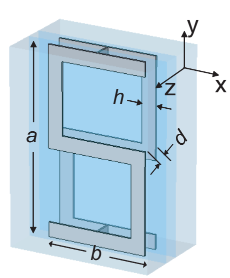
1. INTRODUCTION

Left-handed materials (LHM) were firstly introduced by Veselago [1], which has not been found in nature but was artificially realized as metamaterial and verified by power transmission experiment, prism refraction experiment, beam shifting experiment, T-junction experiment, focusing experiment, and so on [2–4].

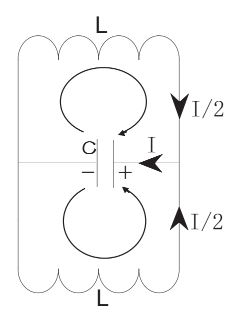
In this paper, we suggest a practical realization of a new kind of material based on a bianisotropic S-ring structure. The advantage of this structure is that each unit is only composed of one resonator but has its own double negative permittivity and permeability, combined with an additional chirality. The chirality is introduced by subtly introducing some vias inside of the ring. We study the negative refraction and cross polarization effects of a biaxial left-handed material with a heterodox constitutive relation whose chirality appears only in one direction. The position of the chirality element in the constitutive matrix is different with that in [5], where the chirality element appears in a off-diagonal position while in our case it appears in a diagonal position. The difference between our work and Ref. [6] is that the permittivity in our model is biaxial instead of uniaxial and can be negative. Since this structure has an LH pass band itself, negative refraction happens naturally and can be controlled by the chirality in the proposed bianisotropic metamaterial.

2. PROPOSE THE STRUCTURE

The configuration and dimension of our realization is shown in Fig. 1(a). When the top and bottom arms of the S-shaped inclusion on both sides of the dielectric substrate are shorted with some vias, as shown in the dashed circles in Fig. 1(a), both the effective capacitances between in the top and bottom horizontal strips disappear, leaving only the effective capacitance in the middle. Under a presupposition that the size of the particle is much smaller than the free space wavelength at



(a)



(b)

Figure 1: (Color online) The structure of (a) bianisotropic S-ring resonator (b) equivalent circuit. The top and bottom arms are shorted with vias marked with dashed circles.

resonant frequency, the equivalent circuit model, as shown in Fig. 1(b), can be used to predict the left-handed pass band of the structure. From the equivalent circuit analysis, we see that the total capacitance of the bianisotropic S-ring resonator is $C/2$, which is 1.5 times of the conventional S-ring resonator (see Fig. 5 in [7]). While the total inductance is almost the same for the same size of the two structure, the magnetic resonant frequency of the bianisotropic S-ring is $\sqrt{2/3}$ times of the original one. However, there is another significant difference that the modified S-ring resonator give us a chirality in the direction perpendicular to the x-y plane.

3. MICROSCOPIC ANALYSIS

In the limitation that the inclusion size is correspondingly small, the quasistatic analysis is valid. When an external magnetic field $\bar{B} = \hat{z}B_z^{ext}\exp(-i\omega t)$ is applied, an electromotive force $EMF = i\omega B_z^{ext}\frac{ab}{2}$ is induced along the metallic strips, equally in both upper half and lower half of the particle. This EMF is responsible for creating a total current flow I which produces a total magnetic moment. According to the equivalent circuit in Fig. 1(b), the particle has a magnetic resonance at the angular frequency $\omega_{m0}^2 \simeq \frac{2}{LC}$, where C is the total capacitance of the S-ring between the center of the metallic strips, as mentioned before, and L is the inductance of each half ring which can be approximated [8], where l_z is the periodicity in z direction for each S-ring unit cell. Hence, the magnetic moment $m_z = abI$ is

$$m_z = \alpha_{zz}^{mm}B_z^{ext}, \alpha_{zz}^{mm} \simeq \frac{a^2b^2}{L}(\frac{\omega_{m0}^2}{\omega^2} - 1)^{-1}. \tag{1}$$

At the same time, the quasistatic potential and charges Q of the middle bar on the front side of the substrate are the electrical images of the one on the back side of the substrate, which leads to a nonzero dipolar electric moment along z-direction, which is the same direction as the external magnetic field B. The electric dipole moment is induced by chirality and is given by:

$$p_z = i\alpha_{zz}^{em}B_z^{ext}, \alpha_{zz}^{em} = \frac{abdC\omega_{m0}^2}{2\omega}(\frac{\omega_{m0}^2}{\omega^2} - 1)^{-1}. \tag{2}$$

Here d is the real distance between the two parallel plates of the capacitance and the polarization along the two parallel metallic strips are obtained from $p_z = Qd = C \cdot \Delta V \cdot d$, where $\triangle V$ is the potential difference between the two plates of the capacitance.

Since the resonator is made of reciprocal media, reciprocity should also hold for the whole structure [5]. As a result, an magnetic moment can be induced when an external electric field along z-direction $\bar{E} = \hat{z}E_z^{ext}\exp(-i\omega t)$ is applied. The magnetic moment is also in z-direction, and can be expressed in a similar way as the electric moment in z-direction

$$m_z = i\alpha_{zz}^{me}E_z^{ext}, \alpha_{zz}^{me} = -\frac{abdC\omega_{m0}^2}{2\omega}(\frac{\omega_{m0}^2}{\omega^2} - 1)^{-1}. \tag{3}$$

On the other hand, when $\bar{E} = \hat{y}E_y^{ext}\exp(-i\omega t)$ is applied, an induced electric current I_w flows in the metallic strip on each side of substrates. We can imagine the whole particle acts like an induced Hertzian dipole modeled as two charge reservoirs of equal and opposite charge q and separated by a distance a when a is comparatively small to the wavelength. As a result of the time harmonic dependence of the external electric field, the imaginary induced dipole inside the particle has moment $p_y = qa$ and oscillates in time with angular frequency ω. The moment is also expressed in terms of current dipole moment $p_y = \frac{I_w a}{-i\omega}$ and polarizability:

$$p_y = \alpha_{yy}^{ee}E_y^{ext}, \alpha_{yy}^{ee} = \frac{a^2}{L_y\omega^2}(\frac{\omega_{e0}^2}{\omega^2} - 1)^{-1}, \tag{4}$$

where $\omega_{e0}^2 = \frac{1}{L_yC_y}$, L_y is the total inductance of the S-ring resonator, it includes mutual coupling between unit cells if the resonator forms an array and can be approximated [9]; and C_y is the total capacitance of the S-ring in the y-direction.

Finally, to summarize, the total electric and magnetic behavior of the particle is given:

$$p_y = \alpha_{yy}^{ee} E_y^{ext}, p_z = i\alpha_{zz}^{em} B_z^{ext}, m_z = \alpha_{zz}^{mm} B_z^{ext} + i\alpha_{zz}^{me} E_z^{ext}. \quad (5)$$

From here we can see the modified S-ring resonator does have a bianisotropic property and the chirality appears only in the $\hat{z}$ direction. For a periodically arrayed S-ring resonator, we can get the macroscopic susceptibilies [5]: $\chi_{eyy} = \frac{\epsilon_{rb}\alpha_{yy}^{ee}}{Vc}$, $\kappa_{zz} = \frac{\sqrt{\mu_0/\epsilon_0\epsilon_{rb}}\alpha_{zz}^{em}}{V_c}$ and $\chi_{mzz} = \frac{\mu_0\alpha_{zz}^{mm}}{V_c}$, where ϵ_{rb} is the relative-to-vacuum permittivity of the background isotropic medium, and V_c is the volume of each unit cell.

Figure 2 plots the analytical values for dispersive ϵ_y, μ_z and ξ_0 as functions of frequency under a certain set of given dimension parameters of the square bianisotropic S-ring resonator indicated in the caption of Fig. 2. From the calculated results we see that both the μ_z and ϵ_y obey Lorentz model. The resonant frequency of ϵ_y is around 0.8 GHz and the resonant frequency of μ_z is around 4.6 GHz. In the frequency band from 4.6 to 5.8 GHz, both μ_z and ϵ_y are negative. The most important thing is that there exists a nonzero chirality term, ξ_0, which has the same resonant frequency as μ_z around 4.6 GHz. Since in the static limit and in the very high frequency limit, the current flow around the ring becomes zero, therefore, ξ_0 approaches zero in these two cases.

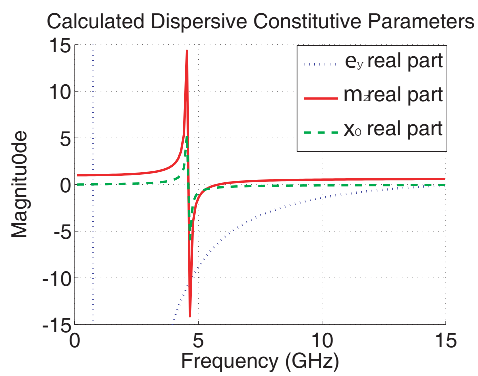


Figure 2: (Color online) Calculated values for dispersive ϵ_y, μ_z and ξ_0 for the bianisotropic S-ring resonator illustrated in Fig. 1(b) with the following structure parameters: $a = 5.2$ mm, $b = 2.8$ mm, $h = 0.4$ mm, $d = 0.5$ mm, $\epsilon_r = 9$, and $l_x \times l_y \times l_z = 4\,\text{mm}\times5.4\,\text{mm}\times1.5\,\text{mm}$.

4. MACROSCOPIC ANALYSIS

While an isotropic chiral medium with simultaneously negative permittivity and permeability is very difficult to realize in practice, an anisotropic chiral medium is much easier to achieve. Here we investigate the case in which the chirality element appears in one direction that is the same as the negative permeability appears. The homogenous bianisotropic biaxial material under consideration takes a constitutive relation as follows:

$$\begin{aligned} \overline{D} &= \overline{\overline{\epsilon}} \cdot \overline{E} + \overline{\overline{\xi}} \cdot \overline{H} = \epsilon_0(1+\overline{\overline{\chi}}_e)\cdot\overline{E} - i\sqrt{\epsilon_0\mu_0}\overline{\overline{\kappa}}\cdot\overline{H}, \\ \overline{B} &= \overline{\overline{\zeta}} \cdot \overline{E} + \overline{\overline{\mu}} \cdot \overline{H} = i\sqrt{\epsilon_0\mu_0}\overline{\overline{\kappa}}^T\cdot\overline{E} + \mu_0(1+\overline{\overline{\chi}}_m)\cdot\overline{H}, \end{aligned} \quad (6)$$

where $\overline{\overline{\chi}}_e$, $\overline{\overline{\chi}}_m$, and $\overline{\overline{\kappa}}$ are the macroscopic susceptibility tensors which give $\overline{\overline{\epsilon}} = \epsilon_0 \begin{pmatrix} \epsilon_x & 0 & 0 \\ 0 & \epsilon_y & 0 \\ 0 & 0 & \epsilon_z \end{pmatrix}$, $\overline{\overline{\mu}} = \mu_0 \begin{pmatrix} \mu_x & 0 & 0 \\ 0 & \mu_y & 0 \\ 0 & 0 & \mu_z \end{pmatrix}$, $\overline{\overline{\xi}} = \frac{1}{c}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -i\xi_0 \end{pmatrix}$, $\overline{\overline{\zeta}} = \frac{1}{c}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & i\xi_0 \end{pmatrix}$, where ϵ_0 and μ_0 are the permittivity and permeability of free space respectively, and c is the speed of light in free space. Note that ϵ_x, ϵ_y, ϵ_z, μ_x, μ_y, μ_z, and ξ_0 are all dimensionless quantities, thereinto, ϵ_y and μ_z are dispersive and can be negative at the same time to form a LH pass band under the appearance of chirality in z-axis.

As a time harmonic plane wave is incident from free space to the medium described in Eq. (6) with the interface perpendicular to x-axis and the incident wave vector $\overline{k_i} = k_{ix}\hat{x} + k_y\hat{y}$ lying in

x-y plane, so the chirality direction is perpendicular to the incident plane. Under the simplification that $\mu_x \approx \mu_y \approx 1$, the dispersion relation of this kind of bianisotropic metamaterial is obtained and we have wave numbers:

$$k_{\pm} = \frac{k_0}{\sqrt{\frac{(\kappa_{22}\nu_{11} + \kappa_{11}\nu_{22}) \pm \sqrt{(\kappa_{22}\nu_{11} - \kappa_{11}\nu_{22})^2 + 4\chi_{22}\kappa_{11}\nu_{11}\gamma_{22}}}{2}}} \tag{7}$$

for both characteristic waves, which propagate at different velocities, where $\kappa_{11} = \frac{\sin^2\phi}{\epsilon_x} + \frac{\cos^2\phi}{\epsilon_y}$, $\kappa_{22} = \frac{\mu_z}{\mu_z\epsilon_z - \xi_0^2}$, $\nu_{11} = \frac{1}{\mu_x}$, $\nu_{22} = \frac{\epsilon_z}{\mu_z\epsilon_z - \xi_0^2}$, $\chi_{22} = -\gamma_{22} = i\frac{\xi_0}{\epsilon_z\mu_z - \xi_0^2}$ and ϕ is the angle between $\overline{k}$ vector and x-axis inside the medium, which is equal to the phase refraction angle under this certain circumstance when $\overline{k}$ is lying in x-y plane and the interface is y-z plane. The medium exhibits a different refractive index for each polarization wave, each of which is elliptically polarized. For + polarization, the plus sign under the first square root in the denominator of Eq. (7) is withheld, so the wave vector should be smaller than $-$ polarization.

The k surface is determined by:

$$\epsilon_y k_y^4 + \epsilon_x k_x^4 + (\epsilon_x + \epsilon_y)k_x^2k_y^2 - \epsilon_y(\epsilon_z + \epsilon_x\mu_z)k_0^2k_y^2 - \epsilon_x(\epsilon_z + \epsilon_y\mu_z)k_0^2k_x^2 + \epsilon_x\epsilon_y C_1 k_0^4 = 0 \tag{8}$$

where $k\cos\phi = k_x$, $k\sin\phi = k_y$, $C_1 = \mu_z\epsilon_z - \xi_0^2$.

The continuity condition requires $k_0 \sin\theta_i = k_+ \sin\theta_+ = k_- \sin\theta_- = k_y$, k_0 is the wave number in free space and θ_i is incident angle. $k_{\pm}$ and $\theta_{\pm}$ are the wave numbers and phase refraction angles for both characteristic waves inside the structured medium. We can also obtain the phase refraction angles for different characteristic waves on the dispersion curves graphically since each possesses its own unique shape.

In the matematerial, since all field vectors are dependent on x and y only and do not have variation in the z direction, all field components in the bianisotropic medium can be expressed as functions of E_x once the wave number is determined. We can also obtain the amplitudes of Poynting vectors in the $\hat{x}$ and $\hat{y}$ directions as

$$S_x = \frac{\omega\epsilon_0 k_{x\pm}}{k_y^2}\left(\frac{\epsilon_x^2}{\epsilon_y} + \alpha_{\pm}\right)|E_x|^2, \tag{9}$$

$$S_y = \frac{\omega\epsilon_0}{k_y}(\epsilon_x - \alpha_{\pm})|E_x|^2, \tag{10}$$

in which $\alpha_{\pm} = \frac{\left(\epsilon_x k_{x\pm}^2 + \epsilon_y k_y^2 - \epsilon_x\epsilon_y\mu_z k_0^2\right)^2}{\mu_y\xi_0^2\epsilon_y^2k_0^4}$. Hence the group refraction angle $\theta_{g\pm} = \tan^{-1} S_y/S_x$ is given as $\theta_{g\pm} = \tan^{-1}\frac{k_y(\epsilon_x - \alpha_{\pm})}{k_{x\pm}(\epsilon_x^2/\epsilon_y + \alpha_{\pm})}$.

We assume that among the constitutive relation in Eq. (6), only three parameters: ϵ_y, μ_z, and ξ_0 are dispersive in the frequency band of our interest. When the metamaterial is a right-handed medium (RHM), whose parameters are all positive. When $|\xi_0|$ is non-zero but small, the k surface in Eq. (8) describes two elliptical curves, one is smaller (+ polarization) and the other is bigger ($-$ polarization) than the k surface for free space, as shown in Fig. 3(a). In LH frequency band, $\epsilon_y < 0$, $\mu_z < 0$, $\epsilon_x, \epsilon_z, \mu_x, \mu_y > 0$. Since ξ_0 has the same resonant frequency as μ_z, ξ_0 is also negative in this band. For weak coupling between the electric and magnetic fields, $|\xi_0|$ is small, the dispersion relations in Eq. (8) describe an elliptical curve and a hyperbolic curve as shown in Fig. 3(b).

Now consider how energy goes through the interface by observing the Poynting vector, in order to satisfy the radiation condition, according to Eq. (9) we should choose k_x values for both + and $-$ polarizations which can guarantee positive Poynting powers in x-direction. And Eq. (10) states that the condition for negative refraction will satisfy $\epsilon_x < \alpha_{\pm}$. Because of the existence of the weak but nonzero chiral factor, two polarization waves should be present simultaneously. In RH frequency band, the two waves are all positively refracted while in LH frequency band, they are oppositely refracted as illustrated in Fig. 3. That's to say, every negative refraction wave is accompanied by a positive refraction counterpart.

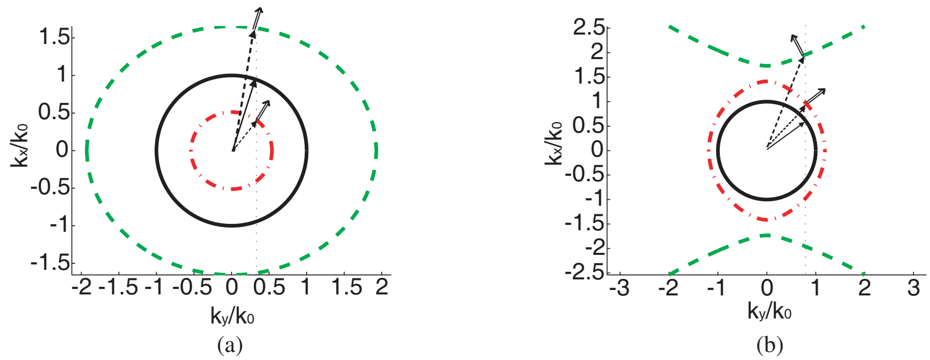


Figure 3: (Color online) The k surfaces for + (red dash-dot line) and − (green dashed line) polarization waves in the bianisotropic metamaterial. The black circle represents k surface for the free space. (a) In RH frequency band, the two waves are all positively refracted. The corresponding parameters are $\epsilon_x = 3$, $\epsilon_y = 2$, $\epsilon_z = 1$, $\mu_z = 1$, $\xi_0 = 0.8$. (b) In LH frequency band, the energy are divided into two flows and oppositely refracted with the parameters $\epsilon_x = 3$, $\epsilon_y = -2$, $\epsilon_z = 1$, $\mu_z = -2$, $\xi_0 = -1$.

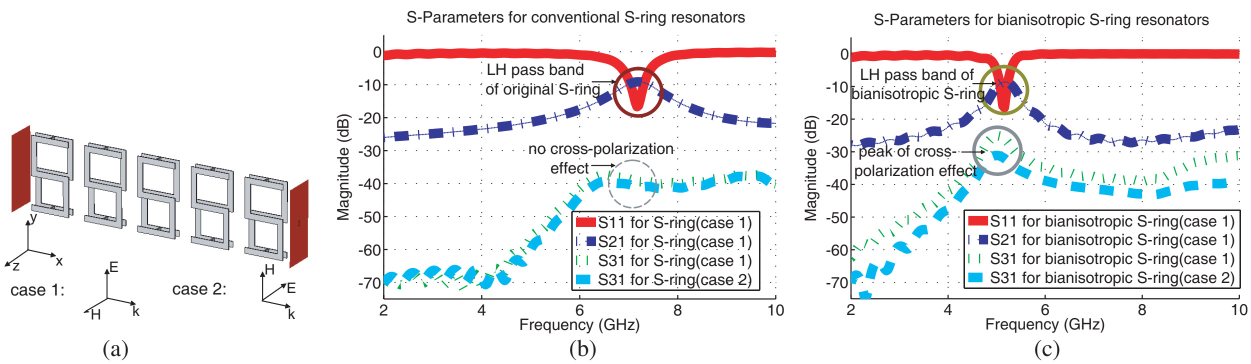


Figure 4: (Color online) S-parameter curves for both (b) conventional S-ring and (c) bianisotropic S-ring resonators in the two cases as their incidences shown in (a). $S11$ stands for energy of the reflected co-polarization wave, $S21$ stands for energy of the transmitted co-polarization wave, $S31$ stands for energy of the transmitted cross-polarization wave.

5. SIMULATIONS

The bianisotropic nature of the S-ring is further confirmed by numerical simulations. In the simulation, we choose the bianisotropic S-ring resonator as the inclusions of the bianisotropic slab, as well as the original S-ring resonator with exact the same size. The dimensions of the square S-ring are $a = 5.2$ mm, $b = 2.8$ mm, $h = 0.4$ mm, $d = 0.5$ mm, and the dielectric constant of the substrate is $\epsilon_r = 9$. A unit cell measures 4 mm×5.4 mm×2.5 mm and there are five cells in the wave propagating direction. We simulate two cases in which both TEM plane waves as incident propagate along x-axis and small probes are used to detect the co- and cross-polarization transmission power. In case 1, the E field is in y-axis and H field in z-axis which is able to produce both magnetic and electrical dipole moment for the proposed structure responsible for a double negative pass band, the peaks of $S21$ curves of both particles as shown in Fig. 4(b) and Fig. 4(c), which reinforces our point in Section 2 that the magnetic resonant frequency should be $\sqrt{2/3}$ times lower than the other. Furthermore Eqs. (2) and (3) demonstrate that the largest magnitude of the cross polarization effects will appear around the magnetic resonance angular frequency ω_{m0} ($f_{m0} \approx 5.16$ GHz) of the bianisotropic S-ring when external z-direction electrical or magnetic field is present. In case 2, the E field and H field change their positions, so no LH pass band is produced but there still should exists the bianisotropic effects and the biggest value should be at the same magnetic resonant frequency as in the first case. Unlike the fact in the theory that $S21$ is close to 0 dB, here we use small probes in the simulation, which yields a smaller amplitude of transmission power. As shown in Fig. 4, all the measured amplitude are 10 to 15 dB lower than the theory, but this does not alter the quantitative behavior of the cross-polarization behavior of the bianisotropic S-ring resonator.

From the S-parameters curves in Fig. 4 we can see, in both cases there exist strong coupling effects ($S31$) of the bianisotropic S-ring near the magnetic resonance compared to the absence of bianisotropy for the original S-ring resonator, which doesn't occur cross-polarization phenomenon as predicted. The peaks of cross-polarization effect curves ($S31$ for S-ring with chirality) in the two cases strongly substantiate the chirality of the bianisotropic S-ring resonator.

6. CONCLUSIONS

We demonstrate a new LHM structure based on the S-ring resonator, which possesses bianisotropic properties. We discuss the distinct dispersion relations in different frequency bands with the chiral factors of diverse values and analyze the novel model with quasistatic approach. Calculation and simulation results have been presented as evidences that the bianisotropy does exist in this kind of matematerial. This unique bianisotropic property contributes to many potential applications of the new proposed material.

ACKNOWLEDGMENT

This work was supported by the Chinese Natural Science Foundation under Grant No. 60531020 and 60671003, by the China Postdoctoral Science Foundation under Grant No. 20060390331, in part by the Office of Naval Research under Contract N00014–01–1–0713, and in part by the Department of the Air Force under Air Force Contract F19628–00–C–0002, and in part by the ZJNSF R105253.

REFERENCES

1. Veselago, V. G., *Sov. Phys. Usp.* Vol. 10, 509, 1968.
2. Smith, D. R., W. J. Padilla, D. C. Vier, S. C. Nemat-Nasser, and S. Schultz, *Phys. Rev. Lett.*, Vol. 84, No. 18, 4184, 2000.
3. Shelby, R. A., D. R. Smith, and S. Schultz, *Science*, Vol. 77, 292, 2001.
4. Ran, L., J. Huangfu, H. Chen, X. Zhang, K. Chen, T. M. Grzegorczyk, and J. A. Kong, *Progress in Electromagnetics Research*, Vol. PIER 51, 249, 2005.
5. Marqués, R., F. Medina, and R. Rafii-El-Idrissi, *Phys. Rev. B*, Vol. 65, 144440, 2002.
6. Cheng, Q. and T. J. Cui, *Phys. Rev. B*, Vol. 73, 113104, 2006.
7. Chen, H., L. Ran, J. Huangfu, X. Zhang, K. Chen, T. M. Grzegorczyk, and J. A. Kong, *Phys. Rev. E*, Vol. 70, 057605, 2004.
8. Chen, H., L. Ran, J. Huangfu, T. M. Grzegorczyk, and J. A. Kong, *J. Appl. Phys.*, Vol. 100, 024915, 2006.
9. Maslovski, S. I., S. A. Tretyakov, and P. A. Belov, *Micro. Opt. Tech. Lett.*, Vol. 35, 47, 2002.

The Bragg Scattering Properties on One-dimensional Composite Right/Left-handed Transmission Line

H. Li[1], **L. He**[1], **and Y. Zhang**[1,2]

[1]Pohl Institute of Solid State Physics, Tongji University, Shanghai 200092, China
[2]School of Electronics and Information, Tongji University, Shanghai 200092, China

Abstract— The transmission properties of one-dimensional composite right/left-handed (CRLH) transmission line (TL) are investigated theoretically and experimented in this paper. Bragg gap depends upon a change of scale length and is sensitive to disorder, so randomness will disappear this band gap. There is another type of photonic band gap corresponding to single negative, distinct from band gaps induced by Bragg scattering, the single negative gap is invariant upon a change of scale length and is insensitive to disorder.

1. INTRODUCTION

The concept of left-handed materials (LHM) was first proposed by Veselago, in 1968 [1]. It was then experimental realized by Shelby using the split-ring resonators (SRR) [2–4]. But the resonant structures such as SRRs has the defect of high loss and narrow band, so an extended transmission-line (TL) approach towards LHM was proposed [5–7]. LHM is also considered to be a more general model of composite right/left-handed (CRLH) transmission line, since it also includes right-handed (RH) nature that occur naturally in practical LHMS. The novel properties of CRLH have been demonstrated such as [8–10].

In this paper, the effect of disorder on single negative gap and Bragg gap is studied by designing the periodic and nonperiodic structure of one-dimensional composite CRLH-TL. The structure that exhibits such band gaps is explicitly designed, simulated and experimented.

2. THE TWO DIFFERENT GAPS IN CRLH-TL

A cell of CRLH TL is implemented by inserting series capacitor and shunt inductor into a conventional TL, which has an intrinsic series inductance and shunt capacitance. The equivalent circuit model of CRLH TL is shown in Fig. 1. The characteristics of the structure can be explained by a dispersion curve. In the lossless case ($\alpha = 0$), the ABCD matrix can be obtained from the equivalent model of Fig. 1 and the propagation constant can be determined using the standard procedure for 1-D periodic analysis of microwave networks [11, 12]. The resulting dispersion relation is

$$\cos(\beta_{\mathrm{bloch}} d) = \cos\theta + \frac{ZY}{2}\cos^2\frac{\theta}{2} + \frac{j}{2}\left(\frac{Z}{Z_0} + \frac{Y}{Y_0}\right)\sin\theta \tag{1}$$

Where $\beta_{\mathrm{bloch}} d$ is the Bloch propagation constant and $\theta = \beta_{TL} d = \omega\sqrt{LC}d$ is the propagation constant of the host TL, and $Z = \frac{1}{j\omega C}, Y = \frac{1}{j\omega L}$, L and C are loaded lumped element components. Using the Eq. (1), the dispersion diagram is shown in Fig. 2. From the Fig. 2, there are two band gaps in the center frequency of 1.82 GHz and 4.9 GHz. In order to study the characteristics of the two gaps, some experiments have been done by using of the 1-D CRLH-TL. The structure was

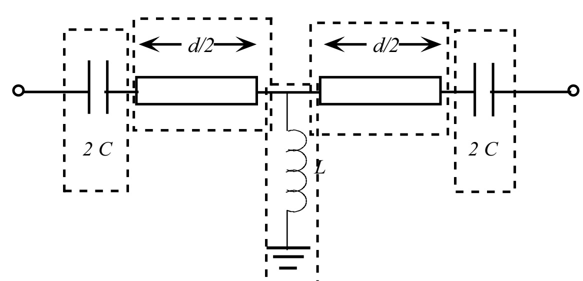


Figure 1: The CRLH-TL unit cell model. The structure consisting of a host TL medium with the distributed parameters L_0 and C_0 periodically loaded with discrete lumped element components, L and C.

made on Teflon substrate with thickness of 1 mm and a dielectric constant of $\varepsilon_r = 2.6$, the width of microstrip transmission line $w = 2.72$ mm. The length of unit cell is 20 mm with 16 units and the loading lumped-element components are $L = 1.8$ nH, $C = 3$ pF which is shown in Fig. 3(a). The S-parameters were measured between the frequencies of 0.05–8 GHz with Agilent 8722ES vector network analyzer. In Fig. 4, the solid curve is the measured result and the dot curse is the simulated result. There are three band gaps at center frequency about 0.5 GHz, 1.82 GHz and 5.5 GHz. The first gap is due to high-pass structure. The second band gap is between the LH passband and RH passband. The third band gap is opening at frequency about 4.6–5.6 GHz. It can be seen that the experimented result is in good agreement with the simulated result.

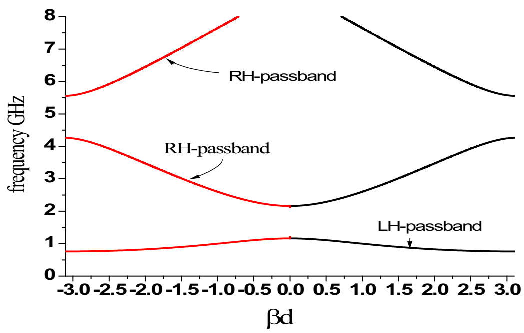
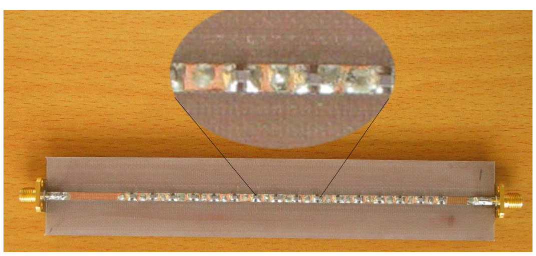
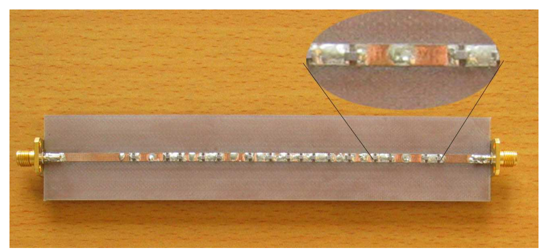


Figure 2: The calculated dispersion diagram of the Eq. (1), $C = 3$ pF, $L = 1.8$ nH.



The periodc structure of the CRLH-TL

(a)



The nonperiodic structure of the CRLH-TL

(b)

Figure 3: Input ports of the two structures are both at left.

3. CRLH-TL WITH NON-PERIODIC LENGTH OF UNITS

In order to find out a difference between the last two gaps in Fig. 4, the CRLH-TL with nonperiodic length of each unit-cell has been designed shown in Fig. 3(b). The length of each unit cell is random but the total length of nonperiodic structure is equal to the total length of periodic structure.

In Fig. 5, the solid curve is the measured result and the dot curse is the simulated result. With the random length of each unit cell and the total length of the structure fixed, the second band gap

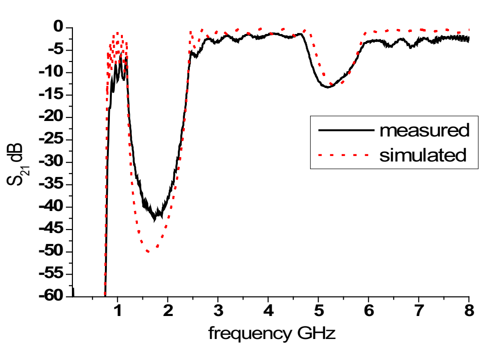


Figure 4: Simulated and measured transmission coefficient (S_{21}) for the periodic structure.

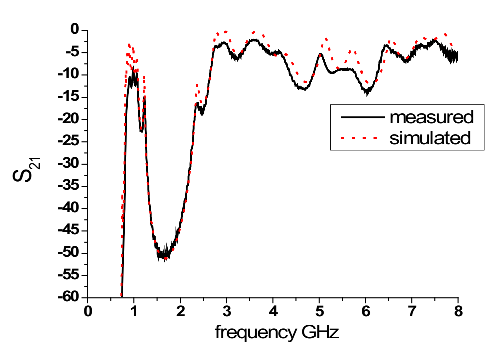


Figure 5: Simulated and measured transmission coefficient (S_{21}) for the nonperiodic structure.

between 1.4 and 2.3 GHz is invariant and is the same as the second gap in the periodic structure which the results are shown in Fig. 4. Whereas, there are several small gaps occur from 4.6–5.8 GHz. It can be to say, the third gap in periodic structure disappears with the random length of each unit cell. Fig. 6 shows the experimental dispersion diagram of the nonperiodic structure, which shows that the frequency of the second band gap is invariant but the third gap disappears. From the above discussion, it can be concluded that in the CRLH-TL, the third gap in the Fig. 4 is the Bragg gap and the second gap is single band gap which is different with the Bragg gap.

4. DISCUSSION

As is discussed in Section II and III, in the structure of the CRLH-TL, the gap between RH passband and LH passband is different from Bragg gap. In order to further confirm the different characteristics of the two gaps, the proposed structure in the Fig. 1 is used and the loading lumped-element components $L_0 = 3.75\,\mathrm{nH}$, $C_0 = 1.5\,\mathrm{pF}$ is chosen, it satisfies $\sqrt{\frac{L_0}{C_0}} = \sqrt{\frac{L}{C}} = 50\,\Omega$. Under this matching condition, it is found that the band gap corresponding to band gap (in the frequency of 1.15 GHz–2.43 GHz) disappear, but the band gap induced by Bragg scattering does not disappear simultaneously. The dispersion diagram can also be obtained by the following Eq. (1), which is shown in Fig. 7.

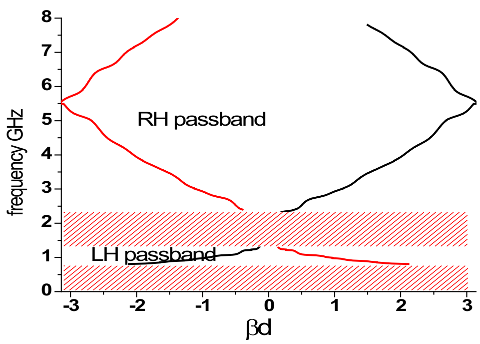


Figure 6: The experimental dispersion diagram of the nonperiodic structure.

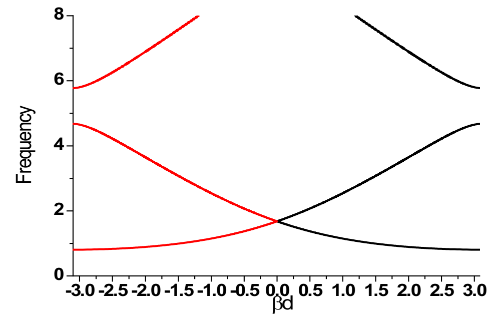


Figure 7: The dispersion diagram of the structure of Fig. 1, when the value of lumped element components satisfy $\sqrt{\frac{L_0}{C_0}} = \sqrt{\frac{L}{C}} = 50\,\Omega$.

Why the Bragg gap doesn't disappear under the matching condition $\sqrt{\frac{L_0}{C_0}} = \sqrt{\frac{L}{C}} = 50\,\Omega$ for the structure which is proposed in Fig. 1. Another structure in Fig. 8 is proposed. Similar to the conventional CRLH TL discussed in Fig. 1, in the lossless case ($\alpha = 0$), the ABCD matrix can be obtained from the equivalent model of Fig. 8 and the propagation constant can be determined using the standard procedure for 1-D periodic analysis of microwave networks. The resulting dispersion relation is

$$\cos(\beta d) = (1 + \frac{ZY}{2})\cos\theta + \frac{j}{2}(\frac{Z}{Z_0} + \frac{Y}{Y_0} + \frac{Z^2 Y}{4Z_0})\sin\theta \tag{2}$$

Using the lumped-element components $L = 3.75\,\mathrm{nH}$, $C = 1.5\,\mathrm{pF}$ too, $\sqrt{\frac{L_0}{C_0}} = \sqrt{\frac{L}{C}} = 50\,\Omega$. Under this matching condition, it is shown that the SNG and Bragg gap disappear simultaneously shown in Fig. 9 which is obtained from the above Eq. (2). In the Fig. 8, the capacitor and the inductor can be considered as an ideal point, so there doesn't exist reflection. For the model of Fig. 1, the depth of the Bragg gap is depended on the difference of impedance between capacitor and host line or between inductor and host line.

In conclusion, the gap between RH passband and LH passband comes from the local properties of each unit, but the third gap comes from the Bragg scattering. Bragg gap depends upon a change of scale length and randomness will disappear the band gap. The single negative band gap, distinct from band gaps induced by Bragg scattering, is invariant upon a change of scale length and is insensitive to disorder.

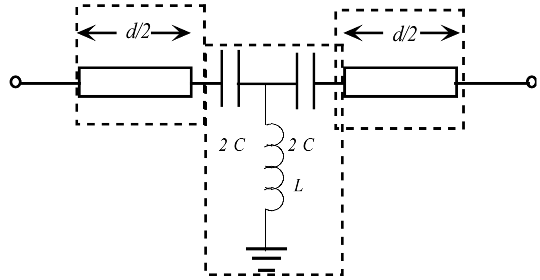


Figure 8: The unit cell model with exchanging the position of lumped capacitor and host line.

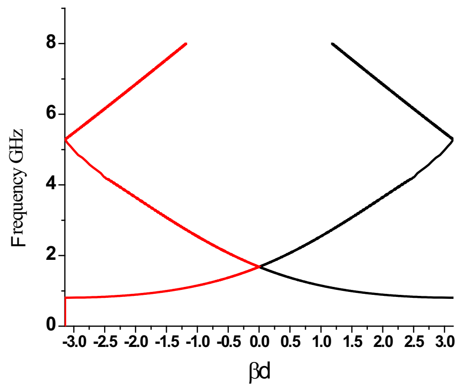


Figure 9: The dispersion diagram of the structure of Fig. 8, when the value of lumped element components satisfy $\sqrt{\frac{L_0}{C_0}} = \sqrt{\frac{L}{C}} = 50\,\Omega$.

ACKNOWLEDGMENT

This project was supported by the 973 project (2001CB610406), and the NSFC (50477048, 10474072).

REFERENCES

1. Veselago, V. G., "The electrodynamics of substances with simultaneously negative values of ε and μ," *Soviet. Physics. Uspekhi*, Vol. 10, No. 4, 509–514, 1968.
2. Pendry, J. B., A. J. Holden, W. J. Stewart, and I. Youngs, "Extremely low frequency plasmons in metallic microstructures," *phys. Rev. Lett.*, Vol. 76, No. 25, 4773–4776, June 1996.
3. Pendry, J. B., A. J. Holden, D. J. Robbins, and W. J. Stewart, "Magnetism from conductors and enhanced nonlinear phenomena," *IEEE Trans. Microwave Theory Tech.*, Vol. 47, 2075–2084, Nov. 1999.
4. Shelby, R. A., D. R. Smith, and S. Schultz, "Experimental verification of a negative index of refraction," *Science*, Vol. 292, 77–79, 2001.
5. Caloz, C. and T. Itoh, "Application of the transmission line theory of left-handed(LH) materials to the realization of a microstrip 'LH Line'," *IEEE AP- S Int. Symp. Dig.*, Vol. 2, 412–415, 2002.
6. Iyer, A. K. and G. V. Eleftheriades, "Negative refractive index matematerials supporting 2-D wave," *IEEE MTT-S Int. Microwave*, *Symp. Dig.*, 1067–1070, 2002.
7. Oliner, A. A., "A periodic-structure negative-refractive-index medium without resonant elements," *IEEE AP-S/URSI Int. Symp. Dig.*, 41, 2002.
8. Li, H. and Y. Zhang, "The Bragg gap on one-dimensional composite right/left-handed transmission line."
9. Li, J., L. Zhou, and C. T. Chan, "Photonic band gap from a stack of positive and negative index materials," *Phys Res Lett.*, Vol. 90, 083901.
10. Jiang, H. and Z. W. Chen, "Omnidirectional gap and defect mode of one-dimensional photonic crystals containing negative-index materials," *Applied Physics Letters*, Vol. 83, No. 26, 5386, 2003.
11. Colin, R. E., *Foundations for Microwave Engineering*, 2nd ed., Vol. 570, 552–556, McGraw-Hill, New York, 1992.
12. Polar, D. M., *Microwave Engineering*, 2nd ed., 162, 424–427, Wiley, New York, 1998.

A Microstrip Highpass Filter with Complementary Split Ring Resonators

C. Li, K. Y. Liu, and F. Li
Institute of Electronics, Chinese Academy of Science, China

Abstract— An accurate equivalent circuit model for the microstrip line with a complementary split ring resonator (CSRR), and the parameter extraction method, are presented. A highpass filter with steep rejection is designed based on the model. The measurement results confirm the efficient analysis and design procedure and the validity of the highpass filter configuration.

1. INTRODUCTION

Recently, split ring resonators (SRRs) and complementary split ring resonators (CSRRs) have gained growing interest for their potential applications. Originally, SRRs combined with metal wires are proposed to make left-handed materials (LHMs) [1]. LHMs are artificial periodical structures with both negative permittivity and permeability. The presence of SRRs leads to the negative permeability in a narrow band above resonance [2]. CSRRs can be made by etching the negative image of SRRs in the ground plane [3]. Hence they are the dual counterpart of the SRRs and exhibit negative permittivity upon their resonance. Since SRRs and CSRRs are both planar configurations, they open a way to develop novel planar microwave circuit and devices [4–7].

In this paper, we proposed an accurate equivalent circuit model for the CSRRs coupling to a microstrip line. The results obtained from the full wave simulation and the equivalent circuit models agree well over a wide frequency band. Based on the equivalent circuit model, a novel high pass microstrip filter with steep rejection is designed and verified by experiment.

2. EQUIVALENT CIRCUIT MODEL OF CSRRS:

Figure 1(a) shows a typical layout and its equivalent circuit for a microstrip line with CSRR etched in the ground plane. The substrate for simulation and later experiment has a relative dielectric constant of 2.65 and a thickness of 1.5 mm. The dimensions are as follows: $s = 0.8\,\text{mm}$, $r_{\text{out}} = 4.2\,\text{mm}$, $r_{\text{in}} = 2.2\,\text{mm}$, $gap = 0.6\,\text{mm}$, $w = 4.1\,\text{mm}$. The line width is chosen for the characteristic impedance of 50 Ohm for a typical microstrip line. The equivalent circuit of the CSRR-based microstrip line shown in Fig. 1(b) consists of a three-element LC tank circuit with two segments of 50 Ohm microstrip line at both sides. The LC tank circuit includes a series LC resonator (L and C_1) and a capacitance element (C_2) connected in parallel.

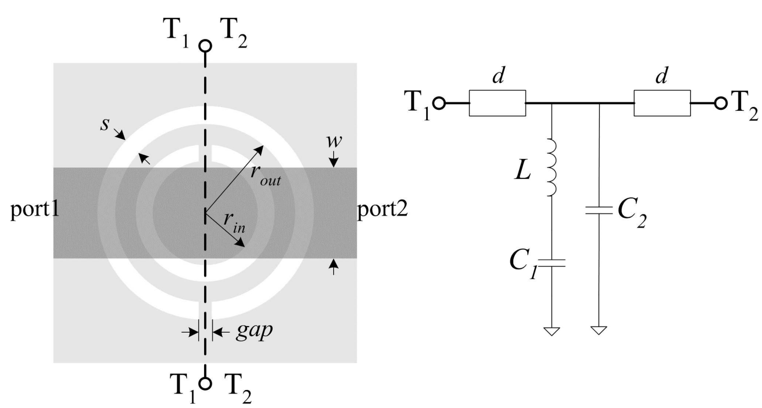


Figure 1: Layout and equivalent circuit of microstrip with CSRR.

Firstly, the performance of the CSRR-based microstrip line is obtained by full wave simulation. To determine the element values of the equivalent circuit, three independent equations are required. The first is given by the resonance condition of the whole tank circuit, which leads to the zeros of S_{11} at f_1. The second is given by the resonance condition of the series LC circuits (L and C_1), which leads to the zeros S_{21} at f_2. The third arises from the 3 dB insertion loss at f_3. Finally, the length of the microstrip line in the equivalent circuit can be obtained by fitting the phase of the S parameters. After some derivation, the element values in the equivalent circuit (Fig. 1) can be expressed by these three special frequency points as follows,

$$C_2 = \frac{Y_0(f_2^2 - f_3^2)}{\pi f_3(f_1^2 - f_3^2)} \tag{1}$$

$$C_1 = (\frac{f_1^2}{f_2^2} - 1)C_2 \tag{2}$$

$$L = \frac{1}{4\pi^2 f_2^2 C_1} \tag{3}$$

where Y_0 is the characteristic immittance of the ports for S parameters. Since $f_1 < f_3 < f_2$ (Fig. 2), it's easily determined from Eq. (1) that the value of C_2 is negative. For dimensions shown in Fig. 1(b), The extracted circuit parameters are as follows: $L = 5.3\,\mathrm{nH}$, $C_1 = 0.54\,\mathrm{pF}$, $C_2 = -1.3\,\mathrm{pF}$, $d = 3\,\mathrm{mm}$. Note that the value of C_2 is negative. Fig. 2 shows the predicted characteristics of the equivalent-circuit and the full wave simulation agree very well.

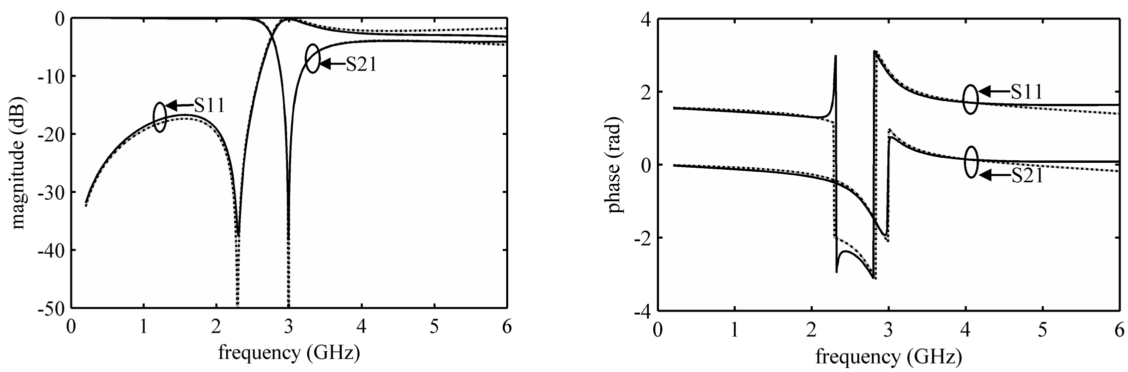


Figure 2: Full wave simulation (solid line) and equivalent circuit (dotted line) results of CSRR.

The shunt negative capacitance is undesirable for the design of high-pass and band-stop filters since it forbids the transmission at high frequency band. This can be avoided by widening the microstrip lines above the CSRR. Fig. 3 shows the layout and the equivalent circuits. The additional patches of the widened microstrip line provide the needed capacitance to offset the negative one.

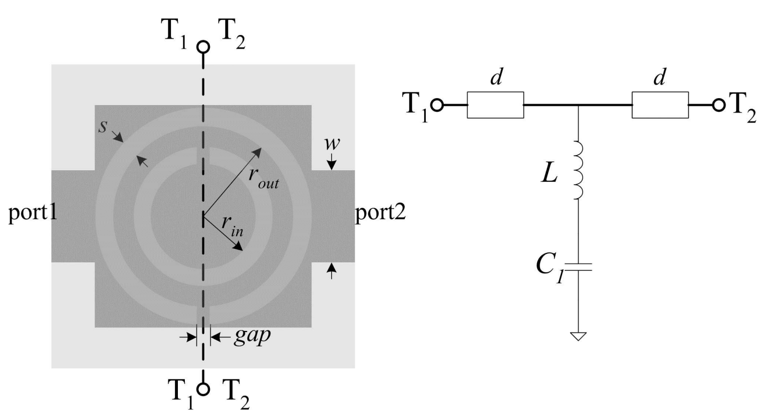


Figure 3: Offsetting the negative capacitance with extra patches (layout and equivalent circuit).

Figure 4 shows the corresponding results of the full wave simulation and the equivalent circuits. It's seen that the performance at high frequency band is improved with the patches. The modified structure serves as a parallel connected series LC resonator embedded between two segments of microstrip lines. The little difference of the resonant frequency (Fig. 4) may result from the coupling of the patches and the CSRR.

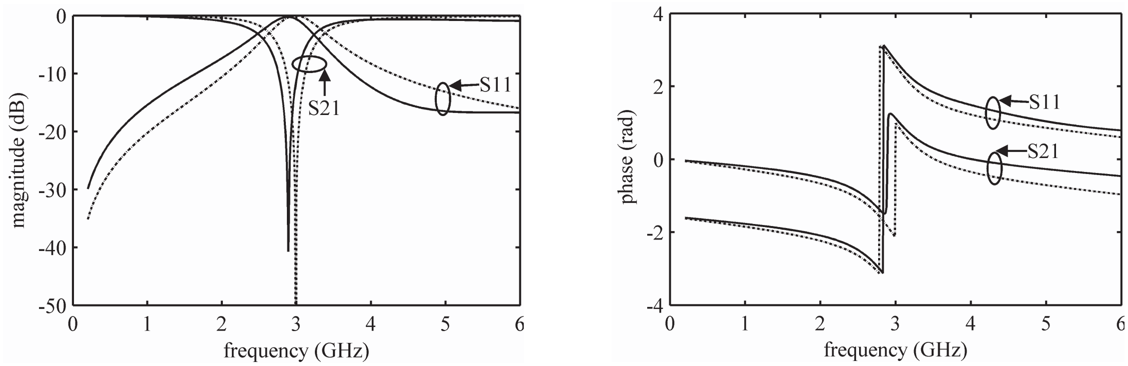
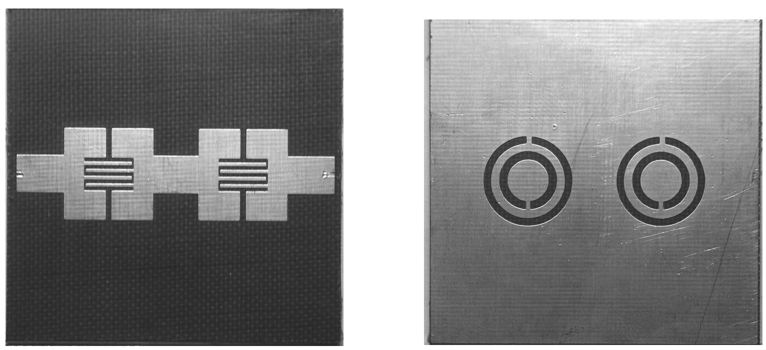


Figure 4: Full wave simulation and equivalent circuit results of CSRR with extra patches.

3. DESIGN OF NOVEL HIGHPASS FILTERS:

Figure 5 shows the photograph of the proposed highpass filter. It comprises of two CSRR sections with additional microstrip patches. The interdigital capacitors are introduced to prevent transmission at lower frequency. The length of the transmission line embedded between the two sections can be adjusted to optimize the response of the total structure which is similar to the elliptic function circuit. The total length of the filter is 2.4 mm. Fig. 6 shows the experimental performance of the proposed filter. The measured 3dB cutoff frequency is 3 GHz and the suppression is more than 25 dB below 2.6 GHz.



(a) Top view (b) Bottom view

Figure 5: The photograph of the highpass filter.

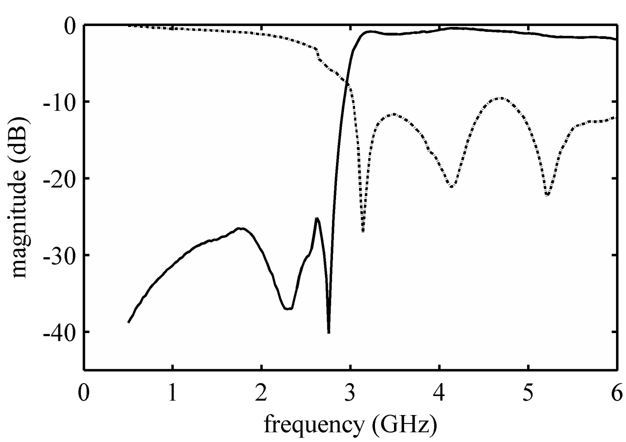


Figure 6: The measured S parameters of the highpass filter (The solid line is for S_{21}, The dotted line is for S_{11}).

4. CONCLUSION

An equivalent circuit model for the CSRR-based microstrip is developed with the parameter extraction method. A novel microstrip highpass filter with CSRRs is proposed. The numerical simulation and the measurement confirm the validity of the highpass filter configuration and the efficient analysis and design procedure. It's seen that the proposed highpass filter exhibits more steep rejection as compared to the conventional structure.

ACKNOWLEDGMENT

This work was supported by the National Natural Science Foundation of China under Grant No (60501018, 60271027), and the National Basic Research Program of China under Grant No 2004CB719800.

REFERENCES

1. Shelby, R. A., D. R. Smith, and S. Schultz, "Experimental verification of a negative index of refraction," *Science*, Vol. 292, 77–79, 2001.
2. Pendry, J. B., A. J. Holden, and D. J. Robbins, et al., "Magnetism from conductors and enhanced nonlinear phenomena," *IEEE. Trans. Microw. Theory Tech.*, Vol. 47, No. 11, 2075–2084, 1999.
3. Falcone, F., T. Lopetegi, J. D. Baena, et al., "Effective negative-e stopband microstrip lines based on complementary split ring resonators," *IEEE Microw. Wireless Compon. Lett.*, Vol. 14, No. 6, 280–282, 2004.
4. Baena, J. D., J. Bonache, F. Martin, et al., "Equivalent-circuit models for split-ring resonators and complementary split-ring resonators coupled to planar transmission lines," *IEEE. Trans. Microw. Theory Tech.*, Vol. 53, No. 4, 1451–1461, 2005.
5. Bonache, J. and I. Gil, "Complementary split ring resonators for microstrip diplexer design," *Electronics Letters*, Vol. 41, No. 14, 2005.
6. Bonache, J. and I. Gil, "Novel microstrip bandpass filters based on complementary split-ring resonators," *IEEE. Trans. Microw. Theory Tech.*, Vol. 54, No. 1, 265–271, 2006.
7. Garcia-Garcia, J., F. Martin, F. Falcon, et al., "Spurious passband suppression in microstrip coupled line band pass filters by means of split ring resonators," *IEEE Microw. Wireless Compon. Lett.*, Vol. 14, No. 9, 416–418, 2004.

High Directive Cavity Antenna Based on 1D LHM-RHM Resonator

Tao Jiang[1], **Yu Yuan**[1], **Dongxing Wang**[1], **Lixin Ran**[1,2], **and Jin Au Kong**[1,2]

[1]Electromagnetics Academy at Zhejiang University, Zhejiang University
Hangzhou 310027, China
[2]Department of Information and Electronic Engineering, Zhejiang University
Hangzhou 310027, China

Abstract— We propose a novel high directive antenna based on 1-D cavity resonator which is composed of a layer of air and a layer of anisotropic left-handed material. Resonance conditions for such a cavity resonator are analyzed, and in the cavity, the resonant frequency is independent of the lateral thickness. A cavity antenna can be made ultra-thin based on the aforementioned property.

1. INTRODUCTION

In 1968, Veselago predicted and investigated a medium (also named left-handed material, LHM) with simultaneously negative permittivity and permeability [1]. After the experiment verifying the characteristic of LHM is carried out in 2000 [2], much attention has been focused on fabrication and application of this kind of new medium. In 2002, Engheta theoretically analyzed the possibility of constructing a thin one-dimensional cavity resonators which combines both conventional material and metamaterial possessing negative permittivity and permeability [3], and this idea was experimentally verified [4, 5]. Later several kinds of subwavelength cavity antennas were put forward, whose ground plane reflective phase are negative [6–8]. In this paper, we propose an ultra-thin high-direcitive cavity antenna based on the 1-D resonator which is made of air and the anisotropic LHM.

2. THOERY ANALYSIS

Engheta [3] theoretically analyzed a thin one-dimentional (1-D) cavity resonator in which a slab of left handed material (LHM) act as a phase compensator, combined with another slab of conventional dieletric material. When the thicknesses of the LHM and the conventional material are small enough, the resonant frequency is only dependent on the ratio of thickness of the two slabs, therefore, the resonator can be very thin. Engheta's theory is based on the isotropic assumption, in this letter we consider a 1-D cavity resonator made of anisotropic LHM.

The cavity structure is schematically shown in the Figure 1. Between the perfectly electrical conducting (PEC) plates, there are two layers of media: a layer of air and a layer of anisotropic medium with the constitutive parameters $\overline{\overline{\varepsilon}} = [\varepsilon_x, \varepsilon_y, \varepsilon_z]$ and $\overline{\overline{\mu}} = [\mu_x, \mu_y, \mu_z]$. Their thicknesses are d_1 and d_2, respectively, and for the 1-D LHM, $\varepsilon_x < 0$, $\mu_y < 0$.

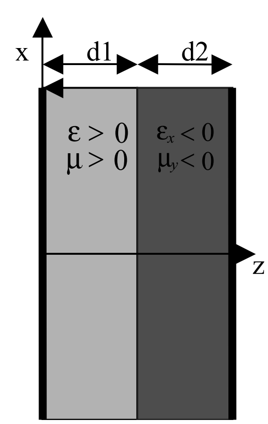


Figure 1: A two layer cavity resonator. The left part is air and the right part is left handed material.

We first consider the case that the electric field is along the x direction and the magnetic field is on the y-z plane, thus the wave vector is in the y-z plane. By utilizing the boundary conditions, the resonance condition is given as

$$k_{1z}\mu_y \tan k_{2z}d_2 + k_{2z}\mu_0 \tan k_{1z}d_1 = 0 \tag{1}$$

where k_{1z}, k_{2z} are the z components of the wave vectors $\vec{k}_1$ (the wave vector in the air) and $\vec{k}_2$ (the wave vector in the medium). For the oblique incidence, $k_{1z} = \sqrt{k_0^2 - k_y^2}$ and $k_{2z} = -\sqrt{\omega^2\varepsilon_x\mu_y - \frac{\mu_y}{\mu_z}k_y^2}$, where k_y is the tangential component of the wave vectors.

Similarly for the case that the magnetic field is along the y direction and the electrical field is in the x-z plane, the corresponding wave vector is in the x-z plane, the resonant conditon is

$$k_{1z}\varepsilon_x \tan k_{1z}d_1 + k_{2z}\varepsilon_0 \tan k_{2z}d_2 = 0 \tag{2}$$

for oblique incidence, $k_{1z} = \sqrt{k_0^2 - k_x^2}$ and $k_{2z} = -\sqrt{\omega^2\varepsilon_x\mu_y - \frac{\varepsilon_x}{\varepsilon_z}k_x^2}$, where k_x is the tangential component of the wave vectors.

Given a certain ω, for the normal incidence, i.e., $k_x = k_y = 0$ for both cases, when d_1 and d_2 are very small, make use of small argument approximation, from Equations (1) and (2) we get the resonance condition

$$\frac{d_1}{d_2} \approx -\frac{\mu_y}{\mu_0} \tag{3}$$

Thus the resonat frequency is only dependent on the ratio of d_1 and d_2, which implies that resonance is independent of the total thickness $d = d_1 + d_2$. Therefore, resonance can occur in ultra-thin cavity.

If a finite-sized monopole antenna oriented in the x direction is put inside the cavity, and a partially reflective plate instead of the PEC is placed on one side of the cavity, we can get the high directive emission.

3. SIMULATION RESULTS

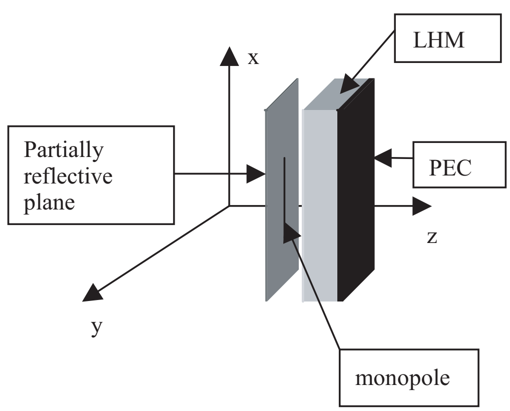


Figure 2: The schematic picture of experiment. A monopole is placed inside the cavity which contains both LHM and RHM, there are metallic and partially reflective plate on each side.

The 1-D LHM- RHM cavity antenna is illustrated in Figure 2. As is shown, inside the cavity a monoploe antenna is used as a source. On the left side of the cavity, a partially reflective plate (the reflectivity is about 0.97) is used so that the EM waves can transmit through it at the resonant frequency. The constitutive parameters of the LHM are as follows: $\varepsilon_y = \varepsilon_z = 4$, $\mu_x = \mu_z = 1$, ε_x and μ_y are described by the Drude model $\varepsilon_x = 1 - \frac{\omega_{ep}^2}{\omega(\omega+i\gamma_e)}\mu_y = 1 - \frac{\omega_{mp}^2}{\omega(\omega+i\gamma_m)}$, where $\omega_{ep} = 2\pi \times 2.45 \times 10^9$ rad/s, $\omega_{mp} = 2\pi \times 1.96 \times 10^9$ rad/s, $\gamma_e = \gamma_m = 2\pi \times 1.59 \times 10^6$ rad/s. The thicknesses of the air and the LHM (along the z direction) are $d_1 = 1$ mm and $d_2 = 3$ mm respectively, while the lateral size of the cavity (along the x and y directions) is 80 mm $\times$ 80 mm. The calculated resonant frequency is 10.655 GHz, according to the theoretical analysis.

We performed FDTD simulations and calculated far-field radiation patterns. Figures 3(a) and (b) give the far-field radiation patterns in the H-plane and E-plane, respectively at the frequency

10.46 GHz, which is very close to the resonant frequency according to theoretical prediction. The half power beamwidth (HPBW) in the E-plane is 23.2 degree and in the H-plane is 21.9 degree. A high directivity is obtained for the cavity antenna, and the directivity will be furthre improved if larger lateral size of the structure is used.

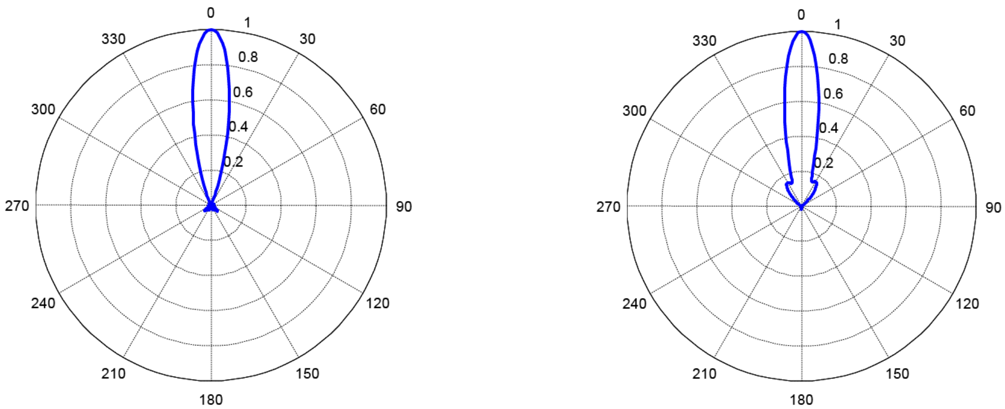


Figure 3: (a) Radiation pattern in E-plane, (b) Radiation pattern in H-plane.

4. CONCLUSIONS

In this paper, we proposed a novel cavity antenna based on 1-D LHM-RHM cavity resonator. The resonance condition for such cavity resonator is given analytically, and the FDTD simulation results for a 4 mm-thick cavity antenna agree with the theorectical prediction. This indicates that a high-directive cavity antenna with any thickness can be realized.

ACKNOWLEDGMENT

This work is supported by NSFC 60531020, 60671003, and in part by ZJNSF R105253.

REFERENCES

1. Veselago, V. G., "The electrodynamics of substances with simultaneously negative values of ε and μ," *Soviet Physics USPEKI*, Vol. 10, No. 509, 1968.
2. Smith, D. R., W. J. Padilla, D. C. Vier, S. C. Nemat-Nasser, and S. Schultz, "Composite medium with simultaneously negative permeability and permittivity," *Phys. Rev. Lett.*, Vol. 84, 4184, 2000.
3. Engheta, N., "An idea for thin subwavelength cavity resonator using metamaterials with negative permittivity and permeability," *IEEE Antennas Propag.*, Vol. 1, 2002.
4. Li, Y., L. Ran, H. Chen, J. Huangfu, X. Zhang, K. Chen, and T. M. Grzegorczyk, "Experimental realization of a one-dimensional LHM–RHM resonator," *IEEE Trans. on Microwave Theory and Tech.*, Vol. 53, No. 4, April 2005.
5. Wang, D., L. Ran, B.-I. Wu, H. Chen, J. Huangfu, T. M. Grzegorczyk, and J. A. Kong, "Multi-frequency resonator based on dual-band S-shaped left-handed material".
6. Ourir, A., A. de Lustrac, and J.-M. Lourtioz, "All-metamaterial-based subwavelength cavities (60λ) for ultrathin directive antennas," *Appl. Phys. Lett.*, Vol. 88, 084103, 2006.
7. Wang, S., A. P. Feresidis, G. Goussetis, and J. C. Vardaxoglou, "High-gain subwavelength resonant cavity antennas based on metamaterial ground planes," *IEE Proc.-Microw. Antennas Propag.*, Vol. 153, No. 1, February 2006.
8. Zhou, L., H. Li, Y. Qin, Z. Wei, and C. T. Chan, "Directive emissions from subwavelength metamaterial-based cavities," *Appl. Phys. Lett.*, Vol. 86, 101101, 2005.

Imaging of Objects through Lossy Layer with Defects

Xiangxiang Cheng[1,2], **Bae-Ian Wu**[3], **Tomasz M. Grzegorczyk**[1,3], **and Jin Au Kong**[1,3]

[1]The Electromagnetics Academy at Zhejiang University, Zhejiang University, Hangzhou 310058, China
[2]Department of Information and Electronic Engineering, Zhejiang University, Hangzhou 310027, China
[3]Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Abstract— The imaging process when a lossy layer is present between the target and the sensor is demonstrated. The lossy layer is inhomogeneous because of the existence of defects, whose dimensions are relatively small. We model the lossy layer as a homogeneous layer with effective parameters, such as permittivity, and use transmission line theory (TLT) to reconstruct the reflection coefficient without obstacle by supposititiously adding a LHM layer or an active medium with exactly opposite signs of permittivity and permeability of the obstacle layer and with the same thickness. Inversion of the transfer matrix in transmission line theory (TLT) is used to cancel the effect of the lossy layer as an obstacle in this inverse problem. The performance of this method is confirmed by some imaging examples.

1. INTRODUCTION

In an active remote sensing system, radar consisting of a transmitter system and a receiver system is utilized. The transmitter sends out a signal to the target of interest and the scattered signal in a specified direction is measured by the receiver, which is then converted into digital form for further processing and exploitation in order to obtain the imaging or the position of the target. Usually, there is nothing but atmosphere on the path between sensor and target, in which case the target information occupies a majority of the return signal despite inevitable noises. In this paper, we present the imaging process when there is an obstacle between the target and the sensor. We assume that the obstacle is a screen-like object and it is very lossy. We call it lossy layer, which might be a plasma medium, a dielectric with extraordinarily large loss, or even metal. We are trying to achieve objects image with the reflection coefficients received by the receiver sensor located behind the lossy layer. The frequency band we are considering extends from 0 to 600 GHz, so the center frequency f_c is equal to 300 GHz with a center wavelength $\lambda_c = 1\,\mathrm{mm}$. The thickness of the lossy layer so far is fixed at 0.1 mm ($\lambda_c/10$). The problem by now is confined in one dimensional (1D) imaging.

We introduce the basic approach of our imaging procedure in this paragraph. First we need to know some information of the lossy layer including the material property (permittivity and permeability) and the thickness, as well as the reflection coefficient when the obstacle and the target are both present which can be received at the receiver. Then we reconstruct the reflection coefficient when there is no obstacle by supposititiously adding a left-handed material (LHM) layer or an active medium, which is an all-calculation process. Finally we are able to obtain the target imaging using the reconstructed reflection coefficient.

We give a detailed description of the methodology how we reconstruct the reflection coefficient and why this method would work. The main challenge in the real world, however, is that there are all sorts of defects in the lossy layer which is equivalent to that the reflection coefficients received by the receiver contain noise. But the sizes of the defects are all very small compared to the center wavelength, for example, narrow and short slots, bubbles with small radius, tiny cracks, and so on. All different types of defects should be modeled properly in order to predict the effective parameters (such as permittivity) of the whole lossy layer with various defects. Two kinds of defects are investigated and the suitable models are proposed in sector 3 of this paper. Combining the methodology of reconstructing reflection coefficients from that which contains the effect of lossy layer in simulation and the model we propose for each type of defects, the performances of target imaging through lossy layer with these defects are present.

2. METHODOLOGY

Here we use the well-known transmission line theory (TLT) [1] to solve this 1-D problem. According to the TLT, for a TE wave normally incident onto a slab, as shown in Figure 1, if we know the reflection coefficient Γ_L at $z = d_0$ surface, we can calculate the E/H fields along with the generalized impedance $Z(z = d_0)$ on this certain surface, using Eq. (1) below:

$$\begin{aligned} E(d_0) &= E_i(e^{ikz} + \Gamma_L(d_0)e^{-ikz}) \\ H(d_0) &= \frac{E_i}{Z_0}(e^{ikz} - \Gamma_L(d_0)e^{-ikz}) \\ Z(z = d_0) &= \frac{E(d_0)}{H(d_0)} = Z_0\frac{1 + \Gamma_L(d_0)}{1 - \Gamma_L(d_0)} \end{aligned} \tag{1}$$

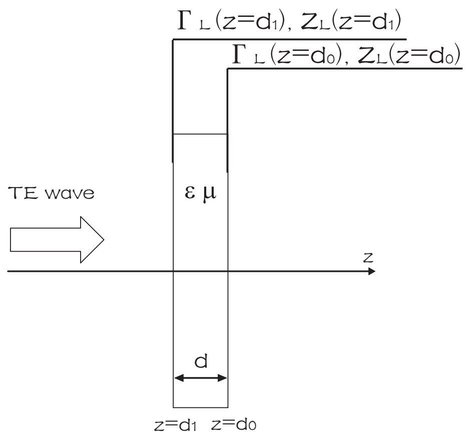


Figure 1: 1-D layers for transmission line theroy.

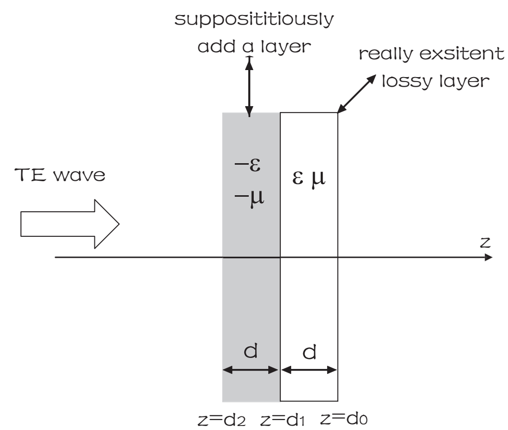


Figure 2: Methodology of reconstructing reflection coefficient.

At the same time, we also can obtain E and H fields at d_1 surface from the E and H fields at d_0 surface applying Eq. (2) below:

$$\begin{bmatrix} E(d_1) \\ H(d_1) \end{bmatrix} = \begin{bmatrix} \cos(kd) & -iZ\sin(kd) \\ -iY\sin(kd) & \cos(kd) \end{bmatrix} \begin{bmatrix} E(d_0) \\ H(d_0) \end{bmatrix} \tag{2}$$

where $k = \omega\sqrt{\epsilon\mu}$ is the wave vector inside the region $d_1 < z < d_0$, $Z = \sqrt{\mu/\epsilon}$ (for TE wave) is the characteristic impedance of the medium inside the region, $Y = 1/Z$ and d is the thickness of the layer. After that the generalized impedance and reflection coefficient at d_1 surface can both be calculated.

If the slab shown in Figure 1 is an obstacle in our imaging problem, the target is in the region where $z > d_0$. And usually we can get the total reflection coefficient at $z = d_1$ surface. Our idea of reconstructing the reflection coefficient without obstacle is illustrated in Figure 2. We imagine there is another layer with the same thickness closer to sensor.

In Figure 2 the white layer is the really existent obstacle, but the gray layer is added in our imagination, so it isn't real. We assume the supposititious layer has exactly opposite signs of permittivity and permeability of the obstacle layer and with the same thickness, the fields at d_2 surface can be obtained by the Eq. (3) because the fields at d_2 surface are known:

$$\begin{bmatrix} E(d_2) \\ H(d_2) \end{bmatrix} = \begin{bmatrix} \cos(kd) & -iZ\sin(kd) \\ -iY\sin(kd) & \cos(kd) \end{bmatrix} \begin{bmatrix} E(d_1) \\ H(d_1) \end{bmatrix} \tag{3}$$

where k is the wave number inside the obstacle slab, and d is its thickness.

By observing the two matrixes in Eq. (2) and Eq. (3), we find out they are inverse matrixes of each other. We call the matrix in Eq. (3) inverted matrix of the lossy layer. So we get that the fields at d_2 surface are exactly identical as at d_0 surface, which means we can recover the fields on d_0 surface by imagining such a layer in front of the obstacle. Therefore we can retrieve the reflection coefficient at $z = d_0$ surface without the effects of the obstacle. The purpose of adding one more layer is to cancel the effect of the obstacle slab ($d_1 < z < d_0$) and make it seems to disappear in the light of electromagnetic waves. So the layer supposititiously added is a LHM layer or an active medium with exactly opposite signs of permittivity and permeability of the obstacle slab and with the same thickness.

After we obtained the electrical and magnetic field without the influence of the obstacle slab, we can retrieve the reflection coefficient still at d_1 surface as below:

$$\begin{aligned} Z(d_0) &= \frac{E(d_0)}{H(d_0)} = \frac{\cos(kd)Z(d_1) + iZ\sin(kd)}{iY\sin(kd)Z(d_1) + \cos(kd)}, \quad Z(d_1) = \frac{E(d_1)}{H(d_1)} \\ R(z=d_1) &= \frac{Z(d_0) - Z_0}{Z(d_0) + Z_0} e^{i2k_0 d}, \quad Z_0 = \eta_0 = \sqrt{\frac{\mu_0}{\varepsilon_0}} \end{aligned} \tag{4}$$

So the $R(z = d_1)$ can be considered as the real reflection coefficient of the object we want to image in the region $z > d_0$.

3. MODEL

When the lossy layer is homogeneous, we can get perfect image using the reconstructed reflection coefficient. However, the lossy layer usually contains defects with all kinds of geometries whose dimensions are much smaller than the working wavelengths, so we should model these lossy layers with different defects in order to get their effective parameters as homogeneous layers. The purpose of modeling is also indicated in Figure 3.

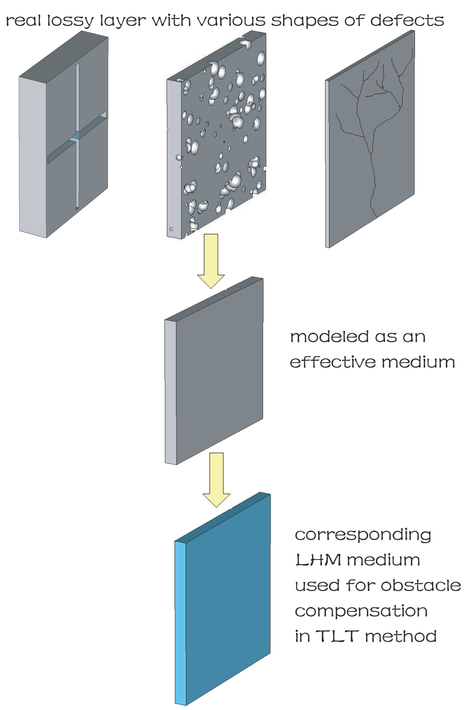


Figure 3: The purpose of modeling lossy layers when there are defects.

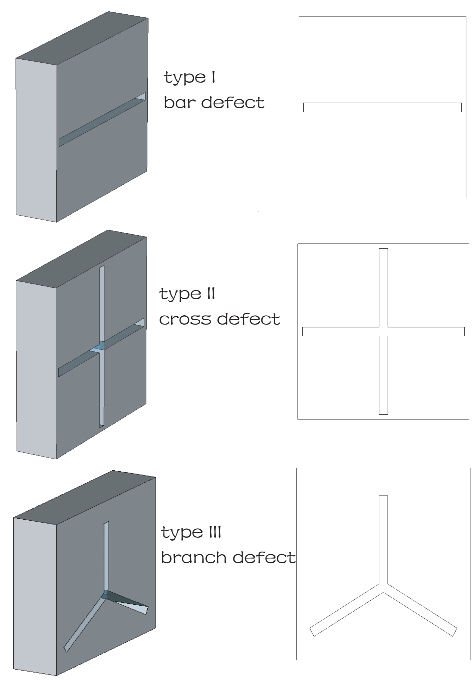


Figure 4: Unit cells for different types of defects in metal layer acting like FSS.

For dielectric materials, either lossy (lossy media) or dispersive (plasma media), they all have definite expressions or certain formulations to represent their material parameters (permittivity). The materials of lossy layer in our problem we look into are focus on metal and plasma media. Their permittivity can all be described by the Drude model.

For bulk metal materials, the free electron gas theory predicts their response to electromagnetic (EM) wave. The frequency dependent permittivity of metal is described by the Drude model as

$$\varepsilon(\omega) = 1 - \frac{\omega_p^2}{\omega^2} \tag{5}$$

where ω_p is the plasma frequency, which is related to the electron density n, electron mass m_e, and electron charge e as

$$\omega_p^2 = \frac{ne^2}{\varepsilon_0 m_e} \tag{6}$$

For most metals, the free electron density is at the order of 10^{22}cm^{-3}, and the plasma frequency ω_p is in the range of visible to ultraviolet (UV) wavelength [2].

Pendry has proposed an effective medium theory that a thin-wire photonic structure behaves like a low density plasma of very heavy charged particles with a plasma frequency in the GHz range [3], and according to Eq. (6) presented an expression of plasma frequency for his structure model. Because the effective electron density of wires is much less and the effective electron mass is much greater than those of electron in metallic film, it is evident that the plasma frequency of wires is much lower than that of the metal film (in the visible and near UV).

The same principle can be applied to our model in the following since the metal film contains defects which lower the electron density. Although the plasma frequency will not decrease so much as metal wires, the defects can always contribute to shifting the plasma frequency of the Drude model down in some degree. The key point is how we can determine the position of ω_p in the proposed Drude model.

A frequency selective surface (FSS) is a periodic metal surface which is basically an assembly of identical elements arranged in a one- or two-dimensional infinite array [4]. There are many types of FSS; in this paper, we mainly study three examples (Figure 4) which are more physical in our situation. We investigate the effective permittivity and permeability for the three kinds of metal layers, the retrieved results [5] of effective permittivity and permeability for the three defected layers are similar to each other. So we consider them with the same proposed model. During comparing the values coming from the Drude model with retrieved ones, we can use the aforementioned Drude model to fit the retrieved results. However the final model needs an additional coefficient for the original Drude model which we call modified Drude model as follow:

$$\varepsilon_{eff}(\omega) = \varepsilon\left(1 - \frac{\omega_p^2}{\omega(\omega + i\gamma)}\right) \tag{7}$$

Another try is based on the model of volume scattering medium. Unlike usual instance, instead of normal homogeneous dielectric media, the host medium is some metal-like lossy medium, or we can say a plasma medium but with very high plasma frequency and the small scatterers inside are air bubbles. We analyze the effective parameters of the lossy layer which has different defect extents (see Figure 5).

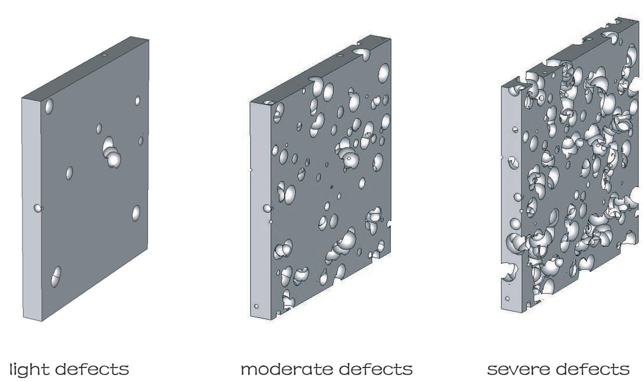


Figure 5: Lossy layers with different degrees of air bubble defects.

From the mechanism of how plasma frequency of the electron gas is given (Eq. (6)), when the air spheres are added to the plasma medium, it's like diluting the electrons in original plasma, the electron density n will become smaller which leads to a lower plasma frequency:

$$\omega_p^2 = \frac{(1 - f_s)n_0 e^2}{\varepsilon_0 m_e} = (1 - f_s)\omega_{p0}^2 \tag{8}$$

where n_0 is the original electron density, and ω_{p0} is the plasma frequency for the non-defected plasma medium. Eq. (6) tells how we predict the effective plasma frequency of the defected plasma layer when f_s (fractional volume of air spheres) is known.

4. PERFORMANCE

Here we present several imaging results which show the evidence of total performance of the through lossy layer imaging after applying our method. We use different models for their corresponding defects. Figure 6(a) is one example of the imaging results for the target positioning. A pec target is located at 25 mm away from sensor while an obstacle is 10 mm away. The obstacle is a metal film defected by a cross with thickness 0.1 mm. The original reflection coefficient data we use is from the S11 parameter in the simulation. And we are only using the data from 450∼600 GHz. In Figure 6(a) the green curve shows the result applying the effective permittivity values from the modified Drude model to compensate the original data, the model is shown in Eq. (7) and the parameter values in the model are $\varepsilon = 35$, $\omega_p = 394.0355\,\text{GHz}$, $\gamma = 0.1\,\text{GHz}$. On the other hand, the purple curve is the result using retrieved effective permittivity and permeability values. From Figure 6(a), we find the obstacle peak gets more suppressed and the power at target peak is enhanced more at the right position when the model is applied. So the parameter values from our model works better than the retrieval in our method. And we have an explanation which is that the retrieved results contain noise and our problem is an intrinsical inverse problem which is very sensitive to noise.

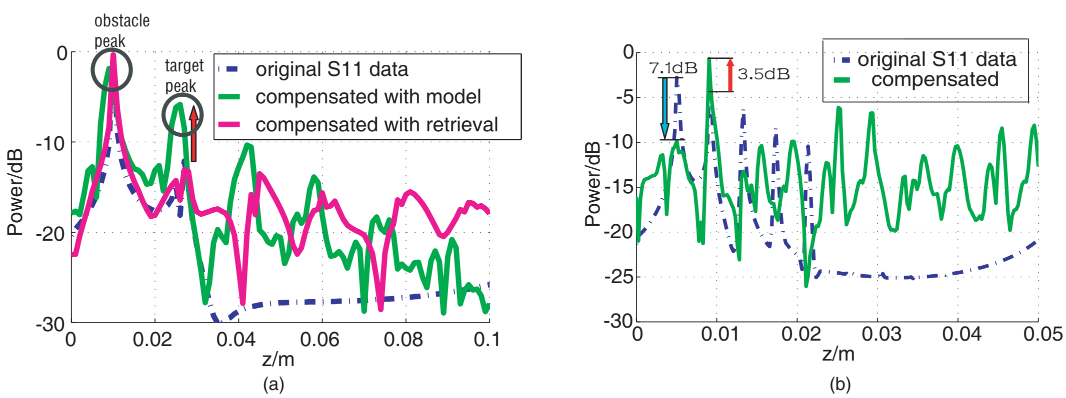


Figure 6: (a) Imaging of PEC target behind a metal film with cross defects using full frequency band 450∼600 GHz. (b) Imaging of PEC target behind the plasma layer without defects using the whole frequency band 0∼600 GHz.

Our initial goal is to detect the target distance away from the sensor. When there is a plasma layer ($\omega_p = 500\,\text{GHz}$, $\gamma = 10\,\text{GHz}$) undefected which is $f_s = 0$, we can see the big improvement of the imaging using our method. As Figure 6(b) shows, the blue curve shows the imaging results using reflection coefficient when a plasma layer without defects ($f_s = 0$) is present. The green curve is also using this data, but the effect of the plasma layer has been compensated. If there are no defects in the plasma slab, by using our method, we can suppress the peak which belongs to the obstacle slab by 7.1 dB and increase the power of the target peak by 3.5 dB as there is no obstacle. We can clearly see that the target is located at $z = 9\,\text{mm}$ while there is an obstacle at $z = 5\,\text{mm}$.

When air spheres are added into the plasma slab and the fractional volume occupied by air is known, we can predict the effective ω_p of the defected plasma layer using Eq. (8), and obtain its effective parameter values. We also compare the performance of the imaging method using parameter values from our model to those from retrieval, and the result is in agreement with the one in the foregoing case, which is when we are using model to predict the effective parameters, the imaging effects are better.

5. CONCLUSIONS

To summarize, inverted matrix is used to cancel the effect of the lossy layer as an obstacle in the process of objects imaging. Parameters from modeling are found to work better than retrieved parameters due to two reasons. First, retrieval can help us find the rough shape of effective permittivity and permeability and may provide us relatively accurate key point (such as plasma frequency), but it is noisy itself. Since the problem is an intrinsical inverse problem, when we invert the matrix under noisy condition, the noise is amplified and thus this method is very sensitive to noise. Secondly, permeability is always equal to one during our modeling and as input into the inverted matrix, since we find it has better effect than the other cases, while in retrieval the permeability is not unity although also some constant. Two typical kinds of defects are investigated and have been used to test our inverted matrix compensation method. The performance of the method is obvious as long as we can get the right parameters of the obstacle. If we can model lossy layer damaged by different types of defects properly, this method of imaging through lossy layer is very helpful for us to find the right position of objects behind obstacle.

ACKNOWLEDGMENT

This work is sponsored by the Office of Naval Research under Contract N00014-01-1-0713, the Department of the Air Force under Air Force Contract F19628-00-C-0002, and the Chinese National Foundation under Contract 60371010 and 60531020.

REFERENCES

1. Kong, J. A., *Electromagnetic Wave Theory*, EMW Publishing, Cambridge, Massachusetts, 2000.
2. Wu, D., N. Fang, C. Sun, X. Zhang, W. J. Padilla, D. N. Basov, D. R. Smith, and S. Schultz, "Terahertz plasmonic high pass filter," *Applied Physics Letters*, Vol. 83, No. 1, 201–203, 2003.
3. Pendry, J. B., A. J. Holdenz, D. J. Robbinsz, and W. J. Stewartz, "Low frequency plasmons in thin-wire structures," *J. Phys.: Condens. Matter*, Vol. 10, 4785–4809, 1998.
4. Munk, B. A., *Frequency Selective Surfaces Theory and Design*, John Wiley & Sons, Inc., 2000.
5. Chen, X., T. M. Grzegorczyk, B.-I. Wu, J. Pacheco Jr., and J. A. Kong, "Robust method to retrieve the constitutive effective parameters of metamaterials," *Physical Review E*, Vol. 70, No. 016608, 2004.

A Two-scale Model for Composite Rough Surface Scattering

W. Z. Yan[1], **L. X. Xu**[1], **Y. Du**[1], **F. Sheng**[1], **Z. N. Li**[1], **and J. A. Kong**[2]

[1]The Electromagnetics Academy at Zhejiang University, Hangzhou 310058, China
[2]Massachusetts Institute of Technology, Boston 02139-4307, USA

Abstract— A new two-scale model is proposed for wave scattering from a composite surface. For each single scale surface roughness, we use a weighted sum of the newly developed statistical integral equation model and the second-order small slope approximation. With these two unifying models collaborating at each scale, this proposed compound model may capture the actual scattering mechanisms and lead to more accurate predictions. It holds the potential to bridge the gap between the regions of validity of the Kirchhoff approximation and the small perturbation model.The chosen simulation parameters suggest that the new two-scale model holds the potential to expand the validity regions of both its large scale and small scale components. The new model may have promising applications for electromagnetic scattering from the ocean surface, whose entire roughness spectrum can be discomposed into small- and large-scale components.

For electromagnetic scattering from a composite rough surface, different forms of two-scale model (TSM) are usually employed [1, 2]. The surface perturbation is typically divided into the large scale part and the small scale one, which are treated by using different analytical models in these TSM models. To calculate the scattering coefficient, in [1], the Kirchhoff approximation (KA) was used for the large scale and the small perturbation method (SPM) for the small scale, while in [2], the SPM was replaced by the first-order small slope approximation (SSA1) for the small scale calculation, leaving KA representation for the large scale untouched.

However, the existence of a gap between the valid regions of the KA and the SPM models may render TSM modles of such combination inaccurate for configurations that are beyond the aforementioned valid regions. Meanwhile, in the literature, some researchers introduced a scale-dividing parameter to distinguish large and small scales in the surface spectrum. Values of the cut-off were often chosen in an ad hoc manner. On the other hand, both the second order SSA (SSA2) [3] and the recently developed statistical integral equation model (SIEM) [4] hold the potential to bridge the gap between the valid regions of the KA and SPM models, with varying degree of success. This observation motivates the study in which we use the weighted sum of the SIEM and SSA2 for the single scale computation for our new TSM. We shall briefly review these two models, and then present the new compound model. Finally some numerical simulations will also be provided.

Unlike the conventional integral equation model (IEM) and its various variations, the SIEM treats the local coordinates and related field terms statistically over the orientation distribution of the surface unit normal vector, which is characterized by joint probability distribution function. Furthermore, it incorporates rigorously the shadow function in the field calculation.

The scattering coefficient of the SIEM model is

$$\sigma_{qp}^0 = \sigma_{qp1}^0 - \sigma_{qp2}^0 + \sigma_{qp3}^0 \tag{1}$$

where

$$\begin{aligned}\sigma_{qp1}^0 &= \frac{k^2}{4\pi\cos\theta_i}\exp[-(k_{sz}+k_z)^2\sigma^2]\{\iint_{D_0} < f_{qp}(\theta_n,\phi_n)f_{qp}^*(\theta_n',\phi_n') > \exp[(k_{sz}+k_z)^2\psi(\rho)]\\ &\quad\cdot\exp[-i(k_{sx}-k_x)\rho\cos\phi - i(k_{sy}-k_y)\rho\sin\phi]\rho d\rho d\phi \end{aligned}\tag{2}$$

$$\sigma_{qp2}^0 = \frac{k^2}{2\cos\theta_i}\exp[-(k_{sz}+k_z)^2\sigma^2]| < f_{qp}(\theta_n,\phi_n) > |^2\int_0^R \rho J_0(k_{d\rho}\rho)\{\exp[(k_{sz}+k_z)^2\psi(\rho)]\}d\rho \tag{3}$$

$$\sigma_{qp3}^0 = \frac{k^2}{2\cos\theta_i}\exp[-\sigma^2(k_z^2+k_{sz}^2)]\sum_{n=1}^{\infty}\sigma^{2n}| < I_{qp}^n > |^2\frac{W^{(n)}(k_{sx}-k_x,k_{sy}-k_y)}{n!} \tag{4}$$

$J_0(\cdot)$ is the Bessel function of the zeroth order

$$k_{d\rho} = \sqrt{(k_{sx}-k_x)^2+(k_{sy}-k_y)^2} \tag{5}$$

$$\begin{aligned} < I_{qp}^n > &= (k_{sz}+k_z)^n < f_{qp} > e^{-\sigma^2 k_z k_{sz}} \\ &+ \frac{k_{sz}^n < F_{qp}(-k_x,-k_y) > + k_z^n < F_{qp}(-k_{sx},-k_{sy}) >}{2} \cos\theta_i \end{aligned} \tag{6}$$

$W^{(n)}(\alpha,\beta)$is the roughness spectrum of the nth power of the surface correlation function given by

$$W^{(n)}(\alpha,\beta) = \frac{1}{2\pi}\int \psi^n(x,y)e^{-i(\alpha x+\beta y)}dxdy \tag{7}$$

Details of the SIEM model can be found in [4].

The small slope approximation (SSA) was introduced by Voronovich [3] as a unifying theory which is based on the expansion of the S-matrix. It has been shown in previous literatures that the SSA has a wider valid region than the SPM. In [2], the SSA1 was used for the small scale calculation. However, it has been noticed that when the correlation length is large, the SSA1 is accurate only for very small slope, and quite sensitive to the RMS height [5]. The SSA2 has shown to be applicable to wider conditions than the SSA1 and is accurate over the whole scattering diagram of the SPM domain except the region with high values of the RMS slope. For surfaces with large kL and moderate to high RMS slopes, however, SSA2 still show appreciable discrepancy with numerical simulations. In this study, we choose SSA2 as a component of the single scale model to evaluate scattering due to its weakness.

The bistatic scattering coefficient for SSA1 is [3, 5, 6]

$$\begin{aligned} \sigma_{qp}^1(\vec{k}_s,\vec{k}) &= \frac{4}{\pi}(\frac{k_z k_{sz}}{k_z+k_{sz}})^2 \left|B_1(\vec{k}_s,\vec{k})\right|^2 \int e^{i\vec{k}_{dx}\cdot\vec{r}}D(\vec{r})d\vec{r} \\ &= 4l^2 e^{-k_{dz}^2h^2}(\frac{k_z k_{sz}}{k_z+k_{sz}})^2 \left|B_1(\vec{k}_s,\vec{k})\right|^2 \sum_{n=1}^{\infty}\frac{(k_{dz}^2h^2)^2}{n!n}e^{-k_{d\rho}^2h^2/(4n)} \end{aligned} \tag{8}$$

where $D(\vec{r})$ is given by

$$D(\vec{r}) = e^{-k_{dz}^2h^2(1-C(\vec{r}))} - e^{-k_{dz}^2h^2} \tag{9}$$

The SSA2 is much more involved:

$$\sigma_{qp}^2(\vec{k}_s,\vec{k}) = \sigma_{qp}^1(\vec{k}_s,\vec{k}) + \sigma_{qp}^{02}(\vec{k}_s,\vec{k}) + \sigma_{qp}^{12}(\vec{k}_s,\vec{k}) + \sigma_{qp}^{22}(\vec{k}_s,\vec{k}) \tag{10}$$

where

$$\sigma_{qp}^{02}(\vec{k}_s,\vec{k}) = \frac{2}{\pi}(\frac{k_z^2k_{sz}^2}{k_z+k_{sz}})Re\{B_1^*(\vec{k}_s,\vec{k})\int d\vec{r}\cdot e^{i\vec{k}_{d\rho}\cdot\vec{r}}D(\vec{r})\int d\vec{\xi}\cdot W(\xi)M_{qp}^*(\vec{k}_s,\vec{k},\vec{\xi})\} \tag{11}$$

$$\sigma_{qp}^{12}(\vec{k}_s,\vec{k}) = -\frac{2}{\pi}(\frac{k_z^2k_{sz}^2}{k_z+k_{sz}})Re\{B_1^*(\vec{k}_s,\vec{k})\int d\vec{r}\cdot e^{i\vec{k}_{d\rho}\cdot\vec{r}}D(\vec{r})\int d\vec{\xi}\cdot e^{i\vec{\xi}\cdot\vec{r}}W(\xi)M_{qp}^*(\vec{k}_s,\vec{k},\vec{\xi})\} \tag{12}$$

$$\sigma_{qp}^{22}(\vec{k}_s,\vec{k}) = \frac{1}{4\pi}(\frac{k_zk_{sz}}{k_z+k_{sz}})^2\int d\vec{r}\cdot e^{i\vec{k}_{d\rho}\cdot\vec{r}}D(\vec{r})\int d\vec{\xi}\cdot e^{i\vec{\xi}\cdot\vec{r}}W(\xi)\left|M_{qp}(\vec{k}_s,\vec{k},\vec{\xi})\right|^2 \tag{13}$$

All the parameters and functions $M_{qp}(\vec{k}_s,\vec{k},\vec{\xi})$, $B_1(\vec{k}_s,\vec{k})$ and $B_2(\vec{k}_s,\vec{k},\vec{\xi})$ are defined in [3].

The new TSM model that we shall propose for scattering from a composite rough surface is to use a weighted sum of the SIEM and SSA2 for either the large- or small scale. As usual the processes governing these two scale surfaces are considered to be independent. Putting it formally, we have

$$\sigma_{qp}^{TSM} = \sigma_{qp}^{small} + \sigma_{qp}^{large} \tag{14}$$

For each single scale, the SSA2 is expected to be truly reliable for smaller values of the RMS height with a reasonably small slope, and we resort to the SIEM for lager RMS height and slope. Thus we can further write the scattering coefficient of single scale surface as

$$\sigma_{qp}^{single} = a\cdot\sigma_{qp}^{SSA2} + (1-a)\cdot\sigma_{qp}^{SIEM} \tag{15}$$

where a is the weight given by

$$a = \exp[-(h/L-s_0)\cdot I(h/L-s_0)]\cdot\exp[-(h-h_0)\cdot I(h-h_0)] \tag{16}$$

Here $I(\cdot)$ is the step function, s_0 and h_0 are thresholds of the RMS slope and RMS height respectively.

To evaluate the model performance, we shall first compare the SIEM, the SSA2 and the method of moment (MoM) when kL is large and $k\sigma$ is small. We expect both the SIEM and SSA2 would show good performance for small RMS slope. In Fig. 1, the configuration of the surface falls into the usually acknowledged limit of the KA domain where the value of RMS slope is 3.8°. We can see from Fig. 1 that both the SIEM and SSA2 agree well with MoM. When we reduce kL and slightly increase $k\sigma$, such that the configuration is slightly on top of the SPM validity region, where the value of the RMS slope is 24.7°. The results are presented in Fig. 2. It is clear that the SSA2 shows $1 \sim 2\,\text{dB}$ discrepancy from the MoM for the majority of the scattered angle range. On the contrary, SIEM still shows very good agreement with MoM.

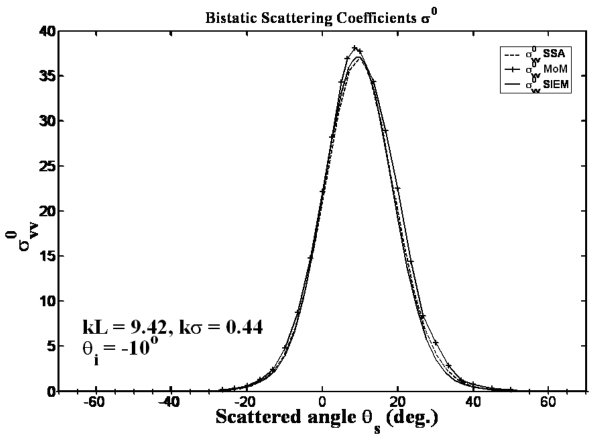


Figure 1: Bistatic scattering comparison between SSA2 and MoM.

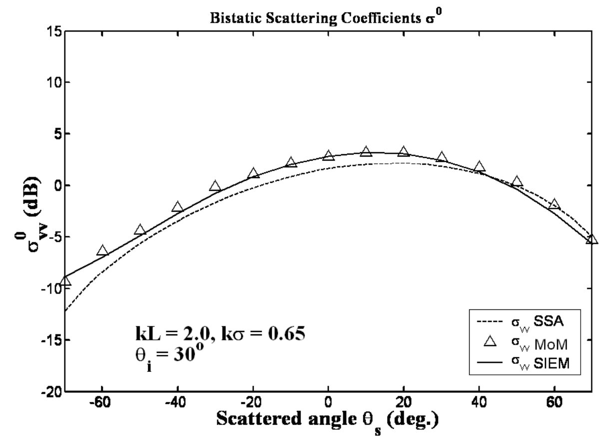


Figure 2: Bistatic scattering comparison between SSA2 and MoM.

Figure 3 shows the behaviors of backscattering coefficients predicted by proposed compound model for a single scale surface against MoM simulations. In the figure, we can observe that our prediction shows good agreement with the MoM at incident angles larger than 20°. It should also be noted that at small incident angles, especially near zero degree, the backscattering coefficients all exceed the MoM by around $3\,\text{dB}$. Such discrepancy is currently under investigation.

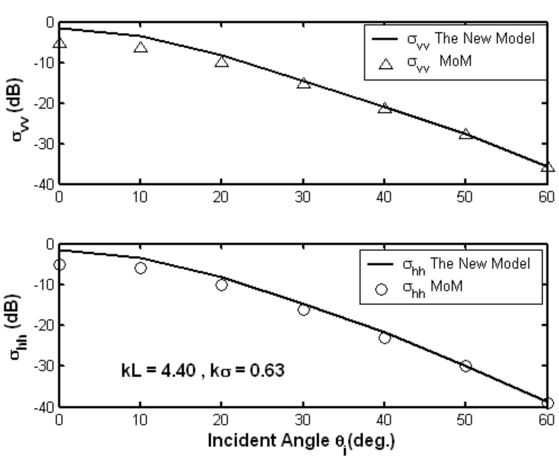


Figure 3: Compare backscattering coefficient from a single scale surface.

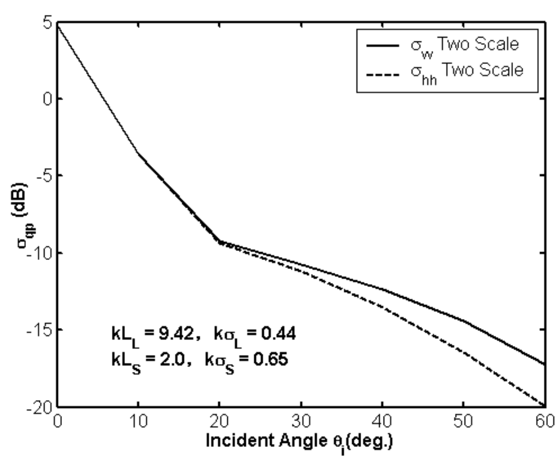


Figure 4: Backscattering coefficient of TSM (The subscripts L and S represent large and small scale respectively.

The backscattering coefficients calculated by our new TSM model for both vertical and horizontal polarization are shown in Fig. 4. The large- and small-scale surfaces fall into the KA region and SPM region respectively. Comparison with numerical simulations will be reported in a later paper.

ACKNOWLEDGMENT

This work was partly supported by China National Science Foundation under Grant No. 40571114 and Zhejiang Science Foundation under Grant No. Y106443.

REFERENCES

1. Ulaby, F. T., R. K., Moore, and A. K. Fung, *Microwave Remote Sensing*, Vol. 2, Artech House, Norwood, MA, 1982.
2. Soriano, G. and M. Saillard, "A two-Scale model for the ocean surface bistatic scattering", *Proceeding of IGARSS'03*, Jul. 2003.
3. Voronovich, A., "Small-slope approximation for electromagnetic wave scattering at a rough interface of two dielectric half-spaces," *Wave in Random Media*, Vol. 4, 337–367, 1994.
4. Du, Y., J. A. Kong, Z. Y. Wang, W. Z. Yan, and L. Peng, "A statistical integral equation model for EM scattering from gaussian rough surface," *Proceeding of PIERS 2005*, Hangzhou, Jul. 2005.
5. Soriano, G., C. A. Guerin, and M. Saillard, "Scattering by two-dimensional rough surfaces: comparison between the method of moments, Kirchhoff and small-slope approximation," *Waves Random Media*, Vol. 12, 63–83, 2002.
6. Gilbert, M. S. and J. T. Johnson, "A study of the higher-order small-slope approximation for scattering from a Gaussian rough surface," *Waves Random Media*, Vol. 13, 137–149, 2003.

A Modified Scheme of Sparse-matrix Canonical-grid Method for Rough Surface Scattering Using Interpolating Green's Function

M. Y. Xia, S. W. Huang, and G. H. Zhang
School of Electronics Engineering and Computer Science, Peking University, Beijing 100871, China

Abstract— A modified scheme of sparse-matrix canonical-grid (SMCG) method for numerical analysis of rough surface scattering is developed in this paper. It uses an interpolating Green's function to replace the Taylor series expansion in the original single-level SMCG method, which becomes intractable for high-order terms. This interpolating approach outpaces the multi-level SMCG method with reduced memory requirement and enhanced parallel computing performance, because it still employs the 2D FFT as the single-level version instead of the 3D FFT in the multi-level revision. The rough surface is modeled by triangular patches to suit RWG basis functions, and Galerkin's matching method is adopted to increase the convergence and accuracy compared with the previous point matching techniques. Numerical results show that the present approach is flexible, efficient and reliable.

Electromagnetic wave scattering from random rough surfaces have many important applications, especially in microwave remote sensing of winds over ocean and soil moisture. Though there are dozens of analytical models [1], their domains of validity are limited and largely remain uncertain. With the availability of high performance supercomputers, rigorous numerical approaches have become popular in the past decade [2, 3]. Because the simulated area must be sufficiently large to make the numerical results meaningful, the discretization of the governing surface integral equations would result in a huge dense matrix system, which challenges both the algorithms and computing platforms. One of the most efficient methods geared for this problem has been the sparse-matrix canonical-grid (SMCG) method. The single-level SMCG scheme [4, 5] is suitable only for small roughness, because the high-order terms of the Taylor series expansion are hard available. To extend the applicability to rougher cases, a multi-level approach was developed [6]. These SMCG versions are based on using rectangular pulse basis functions with point matching method. To increase the convergence and accuracy, an alternate multi-level SMCG version based on using single integral equation method with RWG basis functions was proposed [7, 8]. In the multi-level methods, 3D FFT is utilized instead of the 2D FFT in the single-level scheme. However, this incurs increased memory requirement and badly lowered FFT performance of the MPI-based parallel FFT software [9]. Because the FFTW [9] requires each dimension of a 3D array not less than the number of processors, we can not accelerate the computing by requesting more CPUs, as the number of expansion levels has to be increased accordingly. Though the accuracy is increased by using the RWG basis functions, the implementation of FFT based on these basis functions is more involved. To maintain the merits of the multi-level version but to avoid using the 3D FFT, we propose a simpler interpolating Green's function method to replace the Taylor series expansion about the z-direction grids in this work. In addition, to improve the condition number of the resulting matrix equation, we will use the Galerkin's matching method other than the point matching techniques. In the following, formulations for these modifications are outlined.

We consider a tapered wave incident upon a perfect-electric-conducting (PEC) rough surface, which induces a distribution of unknown current $\mathbf{J}_s$ on the surface. The electric field integral equation (EFIE) to be solved for $\mathbf{J}_s$ is:

$$j\omega\mu_0 \int_s \left[G(R)\mathbf{J}_s(\mathbf{r}') + \frac{1}{k_0^2}\nabla G(R)\nabla'_s \cdot \mathbf{J}_s(\mathbf{r}') \right]_{\tan} dS' = \left[\mathbf{E}^{inc}\right]_{\tan} \tag{1}$$

where 'tan' means taking the tangential components. Instead of expanding the Green's function as Taylor series, we approximate it using an interpolating function $T(z)$, which is taken to be the linear interpoalting function in this paper, and can be generalized to any interpolating function without modifications of the formulations:

$$G(R) = \frac{e^{-jk_0R}}{4\pi R} = \sum_{l=-L}^{L} \sum_{l'=-L}^{L} T(z - l\Delta z) G_{l-l'}(x - x', y - y') T(z' - l'\Delta z) \tag{2}$$

where Δ_z is the grid size, and L will be determined so that $L\Delta_z = |h(x, y)|_{\max}$, where $h(x, y)$ is the elevation of the rough surface. The new 2D integral kernel is

$$G_{l-l'}(x-x', y-y') = G(R_{l,l'}), \quad R_{l,l'} = \sqrt{(x-x')^2 + (y-y')^2 + [|l-l'|\,\Delta z]^2}. \tag{3}$$

To reduce the summation indexes, we rewrite (2) as

$$G(R) = \sum_{l=0}^{L}\sum_{l'=0}^{L} \gamma_l\gamma_{l'}(T_l G_{l-l'} T_{l'} + T_l G_{l+l'} T_{-l'} + T_{-l} G_{l+l'} T_{l'} + T_{-l} G_{l-l'} T_{-l'} \tag{4}$$

where $\gamma_0 = \frac{1}{2}$ and $\gamma_l = 1$ for $l \neq 0$, $T_l = T(z - l\Delta)$ and $T_{l'} = T(z' - l'\Delta z)$, and so forth. The interpolating expression (4) is exact if both $z/\Delta z$ and $z'/\Delta z$ are equal to integers and may have maximum errors if they are equal to half-integers. The error also varies with the horizontal distance $\rho_d = \sqrt{(x-x')^2 + (y-y')^2}$ as shown in Fig. 1. For the maximum error may occur, we chose $z - z' = \frac{1}{2}\Delta z$. It is suggestive by the figure that the error of the interpolating Green's function would be below 0.5% as long as the radial separation is greater than two grids; thus, it can approximate the far interactions in the SMCG methods.

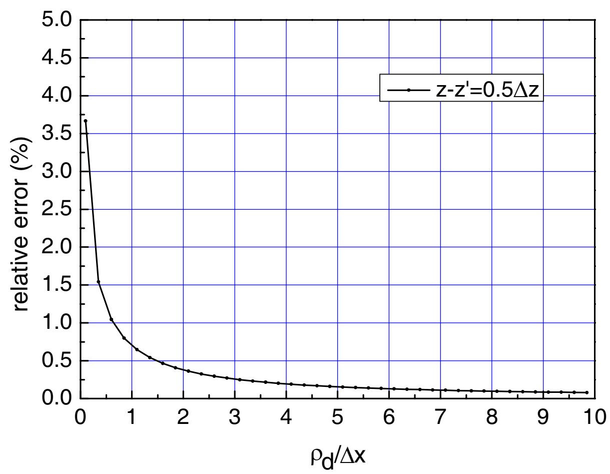


Figure 1: The relative error of the interpolating Green's function against radial separation.

The current distribution on the surface is divided as three types as illustrated in [7] and expressed as

$$\mathbf{J}_s(\mathbf{r}') = \sum_{i=0}^{N_x}\sum_{j=0}^{N_y}\left[I_1(i, j+\frac{1}{2})\tilde{\mathbf{f}}_{i,j+\frac{1}{2}}(\mathbf{r}') + I_2(i+\frac{1}{2}, j)\tilde{\mathbf{f}}_{i+\frac{1}{2},j}(\mathbf{r}') + I_3(i+\frac{1}{2}j+\frac{1}{2})\mathbf{f}_{i+\frac{1}{2},j+\frac{1}{2}}(\mathbf{r}')\right] \tag{5}$$

where $\mathbf{f}(\mathbf{r}')$ are the RWG basis functions with the subscripts indicate the posotion of the edge; $\tilde{\mathbf{f}}_{i,j+\frac{1}{2}}(\mathbf{r}') = 0$ for $i = 0$ and $\tilde{\mathbf{f}}_{i,j+\frac{1}{2}}(\mathbf{r}') = \mathbf{f}_{i,j+\frac{1}{2}}(\mathbf{r}')$ for $i \neq 0$, and $\tilde{\mathbf{f}}_{i+\frac{1}{2},j}(\mathbf{r}') = 0$ for $j = 0$ and $\tilde{\mathbf{f}}_{i+\frac{1}{2},j}(\mathbf{r}') = \mathbf{f}_{i+\frac{1}{2},j}(\mathbf{r}')$ for $j \neq 0$, because we must force $I_1(0, j+\frac{1}{2}) = 0$ and $I_2(i+\frac{1}{2}, 0) = 0$. Substituting (5) into (1) and testing it with the three types of basis functions, respectively, we obtain the moment equation in form of

$$\begin{bmatrix} Z_{11} & Z_{12} & Z_{13} \\ Z_{21} & Z_{22} & Z_{23} \\ Z_{31} & Z_{32} & Z_{33} \end{bmatrix} \begin{Bmatrix} I_1 \\ I_2 \\ I_3 \end{Bmatrix} = \begin{Bmatrix} V_1 \\ V_2 \\ V_3 \end{Bmatrix} \tag{6}$$

where Z_{IJ} are sub-matrices, and I_J is a sub-column vector. In the SMCG method, the interaction matrices are written as strong part and weak part. The strong part is numerically integrated and stored. The weak part is not stored and when it is multiplied with a column vector, the FFT technique applies. In view of the interpolating Green function (4), the multiplied result will be evaluated in the form of

$$\sum_{l}\sum_{l'}\int_{S}\int_{S'}F_l(x,\,y)G_{l-l'}(x-x',\,y-y')K_{l'}(x',\,y')\,dS'dS \tag{7}$$

where $K_{l'}(x',\,y')$ denotes a function related to the source position, while $F_l(x,\,y)$ related to the observation location. The integrals will be carried out using the three-point integration rule over a triangle, and all sampling points will be at the half-integer grids. Detailed implementation is similar to that described in [7]. We will present numerical demonstrations at the conference for simulated wind-driven ocean surfaces.

ACKNOWLEDGMENT

This work is supported by NSFC Project 60471006.

REFERENCES

1. Elfouhaily, T. M. and C. A. Guerin, "A critical survey of approximate scattering wave theories from random rough surfaces," *Waves in Random Media*, Vol. 14, No. 1, R1–R40, 2004.
2. Warnick, K. F. and W. C. Chew, "Numerical simulation methods for rough surface scattering," *Waves in Random Media*, Vol. 11, No. 1, R1–R30, 2001.
3. Saillard, M. and A. Sentenac, "Rigorous solutions for electromagnetic scattering from rough surfaces," *Waves in Random Media*, Vol. 11, No. 3, R103–R137, 2001.
4. Tsang, L., C. H. Chan, and K. Pak, "Backscattering enhancement of a two-dimensional random rough surface (three-dimensional scattering) based on Monte-Carlo simulation," *J. Opt. Soc. Am. A*, Vol. 11, No. 2, 711–715, 1994.
5. Pak, P., L. Tsang, and J. T. Johnson, "Numerical simulations and backscattering enhancement of electromagnetic waves from two-dimensional dielectric random rough surface with the sparse-matrix canonical-grid method," *J. Opt. Soc. Am. A*, Vol. 14, No. 7, 1515–1529, 1997.
6. Li, S. Q., C. H. Chan, M. Y. Xia, B. Zhang, and L. Tsang, "Multilevel expansion of the sparse-matrix canonical grid method for 2-D random rough surfaces," *IEEE Trans. Antennas Propagat.*, Vol. 49, No. 11, 1579–1589, 2001.
7. Xia, M. Y., C. H. Chan, S. Q. Li, et al., "An efficient algorithm for electromagnetic scattering from rough surfaces using a single integral equation and multilevel sparse-matrix canonical-grid method," *IEEE Trans. on Antennas & Propagation*, Vol. 51, No. 6, 1142–1149, 2003.
8. Xia, M. Y. and C. H. Chan, "Parallel analysis of electromagnetic scattering from rough surfaces," *Electronics Letters*, Vol. 39, No. 9, 710–712, 2003.
9. Frigo, M. and S. G. Johnson, "The design and implementation of FFTW3," *Proceedings of the IEEE*, Vol. 93, No. 2, 216–231, 2005. (also see http://www.fftw.org).

Buffered Block Forward Backward Method with Relaxation for 3D Scattering Problems

C. Brennan and M. Mullen

Dublin City University, Ireland

Abstract— The use of a relaxation parameter to improve the convergence rate of the buffered block forward-backward (BFBB) method is investigated. The implementation of the BFBB with relaxation is discussed and it is shown to be equivalent to the application of the standard forward-backward algorithm with relaxation on a suitably preconditioned augmented impedance matrix. A numerical example is provided illustrating the concepts introduced and supporting the conclusions drawn.

The computation of electromagnetic wave scattering from complex bodies is a field of research with an immense range of applications. In recent years there has been an increasing focus on full-wave formulations such as the surface integral equation with particular attention being paid to ways to circumvent the severe computational burden associated with their numerical solution when discretised using the method of moments [1]. Recognizing that full inversion of the dense matrices associated with the method of moments solution is impossible for practically sized problems research has increasingly focussed on iterative solutions that do not require the explicit storage of the impedance matrix or its inverse but rather iteratively refine a trial solution until some convergence criterion is satisfied. While Krylov-subspace based solvers have been favored for many years, due to their robust convergence properties, attention has been paid recently to the development of other iterative solvers, based on stationary methods such as Gauss-Seidel and others [2]. In particular the so-called forward backward method [3], also referred to as the method of ordered multiple interactions [4] has attracted much attention and has shown itself capable of providing solutions that, in some cases, converge more rapidly than Krylov methods for a range of 2D scattering problems [5][1]. The convergence of such methods are not guaranteed, and problems can be encountered when dealing with sharply varying (in particular re-entrant) geometries. These techniques were extended in [6] to scattering from 3D perfectly conducting bodies. This convergence properties of this new technique was recently analyzed in [7] where the name "buffered block forward backward" (BFBB) technique was introduced. This paper extends the BFBB by investigating the use of a relaxation parameter ω and its effect on the convergence of the method. We restrict our attention to scattering from perfectly conducting objects, though the ideas can be readily applied to scattering from imperfectly reflecting bodies also.

1. STATIONARY ITERATIVE SOLUTIONS TO METHOD OF MOMENTS MATRIX EQUATIONS

Application of the method of moments with N basis and testing functions to the electric field integral equation for a perfectly conducting body results in the following matrix equation.

$$\mathbf{Z}\mathbf{J} = \mathbf{V} \tag{1}$$

where $\mathbf{Z}$ is a $N \times N$ dense complex valued matrix containing interactions between basis and testing functions, while $\mathbf{V}$ is a $N \times 1$ vector storing information about the incident field. $\mathbf{J}$ is a $N \times 1$ vector containing the basis amplitides which must be determined. The Gauss-Seidel method [2] starts with a zeroth iterate solution vector $\mathbf{J}^{(0)}$ which is updated according to the rule

$$Z_{ii}J_i^{(k+1)} = V_i - \sum_{j<i} Z_{ij}J_j^{(k+1)} - \sum_{j>i} Z_{ij}J_j^{(k)} \text{ for } i = 1 \cdots N \tag{2}$$

This can be written compactly as

$$(\mathbf{D} + \mathbf{L})\,\mathbf{J}^{(k+1)} = \mathbf{V} - \mathbf{U}\mathbf{J}^{(k)} \tag{3}$$

[1]As mentioned in [5] these methods are equivalent to the the Symmetric Successive Over Relaxation method with relaxation parameter ω set to 1 [2], but this paper will continue to use term 'forward/backward method' as it has become commonly used within the computational electromagnetics community

where $\mathbf{D}, \mathbf{L}, \mathbf{U}$ are respectively the diagonal, lower triangular and upper triangular submatrices of $\mathbf{Z}$ with $\mathbf{Z} = \mathbf{D} + \mathbf{L} + \mathbf{U}$. We define the iteration matrix

$$\mathbf{M}_{GS} = -\left(\mathbf{D} + \mathbf{L}\right)^{-1} \mathbf{U} \tag{4}$$

The error at each Gauss Seidel iteration step $\epsilon^{(k)} \equiv \mathbf{J} - \mathbf{J}^{(k)}$ can be shown to evolve as

$$\epsilon^{(k)} = \mathbf{M}_{GS}^{k} \epsilon^{(0)} \tag{5}$$

and the sequence of solution vectors $\mathbf{J}^{(k)}$ will thus converge to the correct solution $\mathbf{J}$ if and only if the spectral radius $\rho\left(\mathbf{M}_{GS}\right)$ is less than one where

$$\rho\left(\mathbf{M}_{GS}\right) = \max\left\{|\lambda| : \lambda \in \mathbf{M}_{GS}\right\} \tag{6}$$

A variation of Gauss Seidel, which amongst the EM community has become known as forward-backward, is Symmetric Successive Overrelaxation. It consists of a forward sweep through the unknowns in the order $i = 1 \cdots N$ followed by a backward sweep in the order $i = N \cdots 1$. The use of a relaxation parameter ω to accelerate the convergence is discussed in [2]. Including this parameter the forward backward algorithm is

$$Z_{ii} J_i^{(k+\frac{1}{2})} = \omega \left(V_i - \sum_{j<i} Z_{ij} J_j^{(k+\frac{1}{2})} - \sum_{j>i} Z_{ij} J_j^{(k)} \right) + (1-\omega) Z_{ii} J_i^{(k)} \text{ for } i = 1 \cdots N \tag{7}$$

$$Z_{ii} J_i^{(k+1)} = \omega \left(V_i - \sum_{j<i} Z_{ij} J_j^{(k+\frac{1}{2})} - \sum_{j>i} Z_{ij} J_j^{(k+1)} \right) + (1-\omega) Z_{ii} J_i^{(k+\frac{1}{2})} \text{ for } i = N \cdots 1 \tag{8}$$

This can be written more compactly as

$$\left(\mathbf{D} + \omega\mathbf{L}\right) \mathbf{J}^{(k+\frac{1}{2})} = \omega\mathbf{V} - \omega\mathbf{U}\mathbf{J}^{(k)} + (1-\omega)\, \mathbf{D}\mathbf{J}^{(k)} \tag{9}$$

$$\left(\mathbf{D} + \omega\mathbf{U}\right) \mathbf{J}^{(k+1)} = \omega\mathbf{V} - \omega\mathbf{L}\mathbf{J}^{(k+\frac{1}{2})} + (1-\omega)\, \mathbf{D}\mathbf{J}^{(k+\frac{1}{2})} \tag{10}$$

The error evolves as

$$\epsilon^{(k)} = \mathbf{M}_{FB}^{k} \epsilon^{(0)} \tag{11}$$

where

$$\mathbf{M}_{FB} = \left(\mathbf{D} + \omega\mathbf{U}\right)^{-1} \left((1-\omega)\mathbf{D} - \omega\mathbf{L}\right) \left(\mathbf{D} + \omega\mathbf{L}\right)^{-1} \left((1-\omega)\mathbf{D} - \omega\mathbf{U}\right) \tag{12}$$

2. BUFFERED BLOCK FORWARD BACKWARD ALGORITHM

When considering three-dimensional objects we first divide the two-dimensional scatterer surface into M non-overlapping regions, numbered 1 to M. Associated with each such region is a grouping of basis functions whose domains lie in the region. The groups chosen define a unique decomposition of the impedance matrix $\mathbf{Z}$ into blocks or submatrices. Let $\tilde{\mathbf{Z}}_{ij}$ represent the submatrix containing interactions between basis functions in group i and j. A block version of the forward backward method with relaxation proceeds as follows.

$$\tilde{\mathbf{Z}}_{ii} \tilde{\mathbf{J}}_i^{(k+\frac{1}{2})} = \omega \left(\tilde{\mathbf{V}}_i - \sum_{j=1}^{i-1} \tilde{\mathbf{Z}}_{ij} \tilde{\mathbf{J}}_j^{(k+\frac{1}{2})} - \sum_{j=i+1}^{M} \tilde{\mathbf{Z}}_{ij} \tilde{\mathbf{J}}_i^{(k)} \right) + (1-\omega)\, \tilde{\mathbf{Z}}_{ii} \tilde{\mathbf{J}}_i^{(k)} \;\; i = 1 \cdots M \tag{13}$$

$$\tilde{\mathbf{Z}}_{ii} \tilde{\mathbf{J}}_i^{(k+1)} = \omega \left(\tilde{\mathbf{V}}_i - \sum_{j=1}^{i-1} \tilde{\mathbf{Z}}_{ij} \tilde{\mathbf{J}}_j^{(k+\frac{1}{2})} - \sum_{j=i+1}^{M} \tilde{\mathbf{Z}}_{ij} \tilde{\mathbf{J}}_i^{(k+1)} \right) + (1-\omega)\, \tilde{\mathbf{Z}}_{ii} \tilde{\mathbf{J}}_i^{(k+\frac{1}{2})} \;\; i = M \cdots 1 \tag{14}$$

where $\tilde{\mathbf{J}}_i^{(k)}$ is the k^{th} iterate estimate of the basis amplitudes in basis grouping i and similar interpretations apply to other sub-vector and sub-matrix quantities. while the forward backward method converges for many 2D problems this block version diverges when applied to even the simplest 3D scattering structure such as a flat plate. The solution as outlined in [6] is to introduce "buffer zones", that is a small part of its neighboring region. The interactions between each region and its buffer zone is explicitly included at each step of the iterative process. This has the effect of damping spurious diffractions that would otherwise arise and propagate through the solution. To keep what follows as general as possible we allow for the buffer zone for a region to depend on whether we are conducting the forward or backward sweep. In particular, region i has buffer zone $f(i)$ when considering the forward sweep, while it has buffer zone $b(i)$ when considering the backward sweep. The BBFB with relaxation algorithm becomes

$$\begin{bmatrix} \tilde{\mathbf{Z}}_{ii} & \tilde{\mathbf{Z}}_{if(i)} \\ \tilde{\mathbf{Z}}_{f(i)i} & \tilde{\mathbf{Z}}_{f(i)f(i)} \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{J}}_i^{(k+\frac{1}{2})} \\ \tilde{\mathbf{B}}_{f(i)} \end{bmatrix} = \omega\left(\begin{bmatrix} \tilde{\mathbf{V}}_i \\ \tilde{\mathbf{V}}_{f(i)} \end{bmatrix} - \begin{bmatrix} \tilde{\mathbf{L}}_i^{(k+\frac{1}{2})} \\ \tilde{\mathbf{L}}_{f(i)}^{(k+\frac{1}{2})} \end{bmatrix} - \begin{bmatrix} \tilde{\mathbf{U}}_i^{(k)} \\ \tilde{\mathbf{U}}_{f(i)}^{(k)} \end{bmatrix} \right) + (1-\omega) \begin{bmatrix} \tilde{\mathbf{Z}}_{ii} & \tilde{\mathbf{Z}}_{if(i)} \\ \tilde{\mathbf{Z}}_{f(i)i} & \tilde{\mathbf{Z}}_{f(i)f(i)} \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{J}}_i^{(k)} \\ \tilde{\mathbf{B}}_{f(i)} \end{bmatrix} \quad i = 1 \cdots M \tag{15}$$

$$\begin{bmatrix} \tilde{\mathbf{Z}}_{ii} & \tilde{\mathbf{Z}}_{ib(i)} \\ \tilde{\mathbf{Z}}_{b(i)i} & \tilde{\mathbf{Z}}_{b(i)b(i)} \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{J}}_i^{(k+1)} \\ \tilde{\mathbf{B}}_{b(i)} \end{bmatrix} = \omega\left(\begin{bmatrix} \tilde{\mathbf{V}}_i \\ \tilde{\mathbf{V}}_{b(i)} \end{bmatrix} - \begin{bmatrix} \tilde{\mathbf{L}}_i^{(k+\frac{1}{2})} \\ \tilde{\mathbf{L}}_{b(i)}^{(k+\frac{1}{2})} \end{bmatrix} - \begin{bmatrix} \tilde{\mathbf{U}}_i^{(k+1)} \\ \tilde{\mathbf{U}}_{b(i)}^{(k+1)} \end{bmatrix} \right) (1-\omega) \begin{bmatrix} \tilde{\mathbf{Z}}_{ii} & \tilde{\mathbf{Z}}_{ib(i)} \\ \tilde{\mathbf{Z}}_{b(i)i} & \tilde{\mathbf{Z}}_{b(i)b(i)} \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{J}}_i^{(k+\frac{1}{2})} \\ \tilde{\mathbf{B}}_{b(i)} \end{bmatrix} \quad i = M \cdots 1 \tag{16}$$

$\tilde{\mathbf{Z}}_{if(i)}$ is the submatrix of $\mathbf{Z}$ containing interactions between basis functions in group i and those basis functions in its forward buffer zone. A simple interpretation is applied to $\tilde{\mathbf{Z}}_{ib(i)}$ and other entries. The buffer zone unknowns, $\tilde{\mathbf{B}}_{f(i)}$ and $\tilde{\mathbf{B}}_{b(i)}$, hold approximate solutions for the unknowns in those buffer zones. Their role is to allow the currents in these buffer zones to accurately couple with the currents in group i, damping any spurious edge effects that would otherwise occur.$\tilde{\mathbf{L}}_i^{(k)}$ and $\tilde{\mathbf{U}}_i^{(k)}$ contain information about fields scattered from other groups to group i.

$$\tilde{\mathbf{L}}_i^{(k)} = \sum_{j=1}^{i-1} \hat{\tilde{\mathbf{Z}}}_{ij} \tilde{\mathbf{J}}_j^{(k)} \qquad \tilde{\mathbf{U}}_i^{(k)} = \sum_{j=i+1}^{M} \hat{\tilde{\mathbf{Z}}}_{ij} \tilde{\mathbf{J}}_j^{(k)} \tag{17}$$

where $\hat{\tilde{\mathbf{Z}}}_{ij}$ is obtained by taking $\tilde{\mathbf{Z}}_{ij}$ and putting to zero the interaction between any basis function in group j that is also contained in the appropriate buffer zone of group i ($f(i)$ for Equation (15) or $b(i)$ for Equation (16)). Such interactions have already been accounted for, when considering the interactions between group i and its buffer zone.

3. CONVERGENCE OF THE BBFB WITH RELAXATION

It is straightforward to verify that each forward sweep of the BBFB is equivalent to a forward sweep of an *unbuffered* block forward-backward method applied to the augmented matrix equation.

$$\mathbf{Z}_f \mathbf{J}_f = \mathbf{V}_f \tag{18}$$

By inspection of Equation (15) it is possible to show that

$$\mathbf{Z}_f = \left[\begin{array}{cc|cc|cc|c|cc} \tilde{\mathbf{Z}}_{11} & \tilde{\mathbf{Z}}_{1f(1)} & \hat{\tilde{\mathbf{Z}}}_{12} & 0 & \hat{\tilde{\mathbf{Z}}}_{13} & 0 & \cdots & \hat{\tilde{\mathbf{Z}}}_{1M} & 0 \\ \tilde{\mathbf{Z}}_{f(1)1} & \tilde{\mathbf{Z}}_{f(1)f(1)} & \hat{\tilde{\mathbf{Z}}}_{f(1)2} & 0 & \hat{\tilde{\mathbf{Z}}}_{f(1)3} & 0 & \cdots & \hat{\tilde{\mathbf{Z}}}_{f(1)M} & 0 \\ \hline \hat{\tilde{\mathbf{Z}}}_{21} & 0 & \tilde{\mathbf{Z}}_{22} & \tilde{\mathbf{Z}}_{2f(2)} & \hat{\tilde{\mathbf{Z}}}_{23} & 0 & \cdots & \hat{\tilde{\mathbf{Z}}}_{2M} & 0 \\ \hat{\tilde{\mathbf{Z}}}_{f(2)1} & 0 & \tilde{\mathbf{Z}}_{f(2)2} & \tilde{\mathbf{Z}}_{f(2)f(2)} & \hat{\tilde{\mathbf{Z}}}_{f(2)3} & 0 & \cdots & \hat{\tilde{\mathbf{Z}}}_{f(2)M} & 0 \\ \hline \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \\ \hline \hat{\tilde{\mathbf{Z}}}_{m1} & 0 & \hat{\tilde{\mathbf{Z}}}_{M2} & 0 & \hat{\tilde{\mathbf{Z}}}_{m3} & 0 & \cdots & \tilde{\mathbf{Z}}_{MM} & \tilde{\mathbf{Z}}_{Mf(M)} \\ \hat{\tilde{\mathbf{Z}}}_{f(m)1} & 0 & \hat{\tilde{\mathbf{Z}}}_{f(M)2} & 0 & \hat{\tilde{\mathbf{Z}}}_{f(m)3} & 0 & \cdots & \tilde{\mathbf{Z}}_{f(M)M} & \tilde{\mathbf{Z}}_{f(M)f(M)} \end{array}\right] \tag{19}$$

and

$$\mathbf{J}_f = \left[\tilde{\mathbf{J}}_1 \quad \tilde{\mathbf{B}}_{f(1)} | \tilde{\mathbf{J}}_2 \quad \tilde{\mathbf{B}}_{f(2)} | \cdots | \tilde{\mathbf{J}}_M \quad \tilde{\mathbf{B}}_{f(M)} |\right]^T \tag{20}$$

$$\mathbf{V}_f = \left[\tilde{\mathbf{V}}_1 \quad \tilde{\mathbf{V}}_{f(1)} | \tilde{\mathbf{V}}_2 \quad \tilde{\mathbf{V}}_{f(2)} | \cdots | \tilde{\mathbf{V}}_M \quad \tilde{\mathbf{V}}_{f(M)} |\right]^T \tag{21}$$

We have used horizontal and vertical lines to make explicit the blocks of $\mathbf{Z}_f, \mathbf{J}_f$ and $\mathbf{V}_f$ to be used during the unbuffered block forward sweep. In a similar manner it is possible to show that the backward sweep of the BBFB is equivalent to a backward sweep of an unbuffered block forward-backward method applied to the augmented matrix equation.

$$\mathbf{Z}_b \mathbf{J}_b = \mathbf{V}_b \tag{22}$$

In turn, it is straightforward to show that the unbuffered block forward and block backward sweeps applied to Equations (18) and (22) are equivalent to standard (point by point, rather than block by block) forward and backward sweeps applied to a preconditioned version of these equations. The preconditioner for Equation (18) is a block diagonal matrix whose diagonal blocks are the inverses of the diagonal blocks in matrix (19), with the equivalent block diagonal preconditioner being used for Equation (22). Let the preconditioned version of $\mathbf{Z}_f$ be equal to $\mathbf{Y}_f$ and the preconditioned version of $\mathbf{Z}_b$ be $\mathbf{Y}_b$ such that

$$\mathbf{Y}_f = \mathbf{D}_f + \mathbf{L}_f + \mathbf{U}_f \tag{23}$$
$$\mathbf{Y}_b = \mathbf{D}_b + \mathbf{L}_b + \mathbf{U}_b \tag{24}$$

where $\mathbf{D}, \mathbf{L}$ and $\mathbf{U}$ are diagonal, lower and upper triangular submatrices. The convergence of the BBFB with relaxation therefore depends on the eigenvalues of the iteration matrix

$$\mathbf{M}_{BBFB} = (\mathbf{D}_b + \omega \mathbf{U}_b)^{-1} ((1-\omega)\mathbf{D}_b - \omega \mathbf{L}_b) (\mathbf{D} + \omega \mathbf{L}_f)^{-1} ((1-\omega)\mathbf{D}_f - \omega \mathbf{U}_f) \tag{25}$$

In particular if the spectral radius, $\rho(\mathbf{M}_{BBFB})$ is less than one the solution will converge to the correct answer. In addition we expect that values of the relaxation parameter ω that lead to a smaller spectral radius should lead to a more rapid convergence.

4. RESULTS

We consider a simple example of plane wave scattering from a square flat metallic plate lying in the xy-plane. We deliberately choose a small problem so that we can explicitly compute the eigenvalues associated with the iteration matrix and verify the convergence criterion developed in the last section. For examples of the BBFB applied to larger problems readers should consult [6] or [7]. Referring to Figure 1 the incident field is propagating in the $-\hat{k}$ direction and the electric field is polarized in the $\hat{\imath}$ direction. Each side of the plate is of length 1λ (where $f = 300\,MHz$). The plate was discretised with 9 discretisations per side and rooftop basis functions were applied in the $\hat{\imath}$ and $\hat{\jmath}$ directions. For the purpose of applying the BBFB the forward direction was chosen

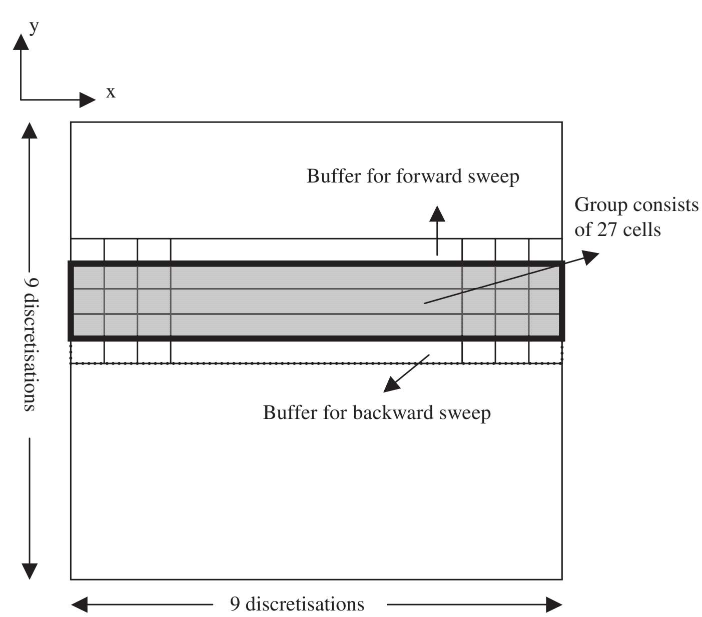


Figure 1: Set up for numerical example.

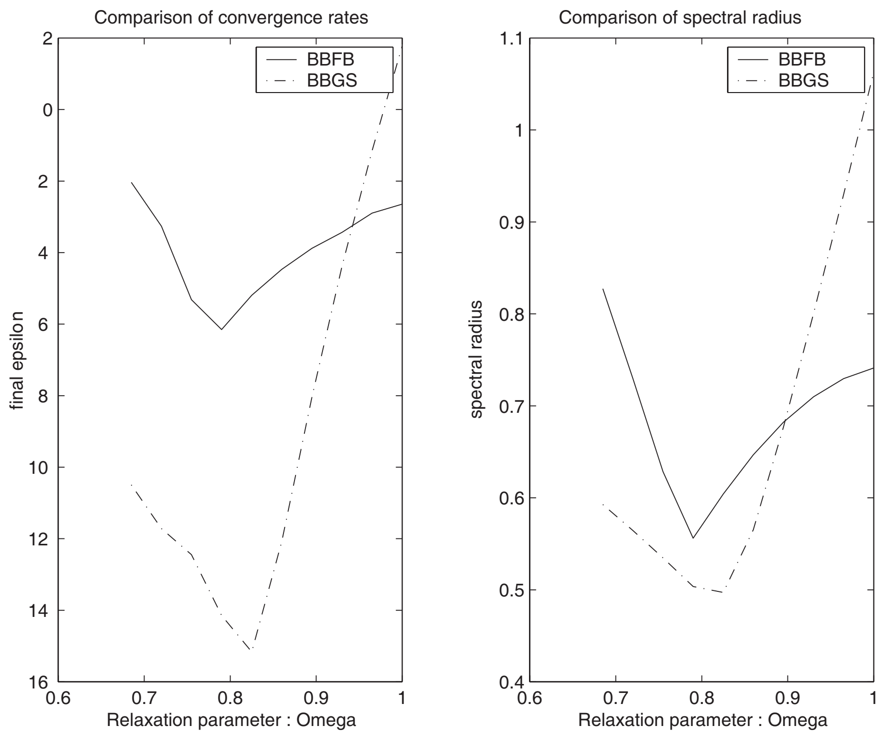


Figure 2: Convergence rates and spectral radius versus omega.

as $\hat{\jmath}$. Each group consisted of the 27 cells contained in three strips running in the $\hat{\jmath}$ direction. The buffer region consisted of the strip of cells immediately adjacent to the group, in the direction of the sweep. The BBFB algorithm with relaxation was applied with varying values of ω in the region $0.65 < \omega < 1.0$ and, each time, allowed to progress for 25 iterations. For each value of ω we noted the final error $\log_{10} \epsilon^{(k=25)}$ (It was possible to compute the exact solution $\mathbf{J}$ for this small problem by direct matrix inversion). The spectral radius of the iteration matrix $\mathbf{M}_{BFBB}$ was also computed for each value of ω. Figure 2 shows the spectral radius and final error values. We note that the BFBB converges in all cases as the spectral radius of the iteration matrix is less than 1 for all values of ω. The convergence rate is optimized by choosing a value of ω just under 0.8 where it is noted that the spectral radius is minimized, as expected. In order to gauge whether forward-backward schemes constitute an improvement over standard Gauss-Seidel we repeated the experiment, but instead used a buffered block Gauss-Seidel (BBGS) method to solve Equation(18). As each SSOR iteration is equivalent to two Gauss-Seidel steps we noted the final error after 50 iterations, in order to keep the computational costs comparable. We also computed the spectral radius of the corresponding iteration matrix. We note that the buffered block Gauss Seidel method does not converge in all cases, most notably when relaxation is not used ($\omega = 1$) where the spectral radius is greater than 1. However, when relaxation is applied the buffered block Gauss-Seidel method converges faster, provided $\omega < 0.9$.

ACKNOWLEDGMENT

The authors would like to recognize the support of this research by the Irish Research Council for Science, Engineering and Technology: funded by the National Development Plan.

REFERENCES

1. Harrington, R. F., "Matrix methods for field problems," *Proceeedings of the IEEE*, Vol. 55, No. 2, 136–149, 1967.
2. Golub, G. H. and C. F. van Loan, *Matrix Computations*, John Hopkins, 1996.
3. Holliday, D., L. L. DeRaad, and G. J. St. Cyr, "Forward-backward: A new method for computing low grazing angle scattering," *IEEE Trans. Ant. Prop.*, Vol. 44, No. 5, 722–729, 1996.
4. Kapp, D. A. and G. Brown,"A new numerical method for rough surface scattering calculations," *IEEE Trans. Ant. Prop.*, Vol. 44, No. 5, 711–721, 1996.

5. West, J. C. and J. M. Sturm, "On iterative approaches for electromagnetic rough surface scattering problems," *IEEE Trans. Ant. Prop.*, Vol. 47, No. 8, 1281–1288, 1999.
6. Brennan, C., P. J. Cullen, and M. Condon, "A novel iterative solution of the three dimensional electric field integral equation," *IEEE Trans. Ant. Prop.*, Vol. 52, No. 10, 2781–2785, 2004.
7. Brennan, C. and D. Bogusevschi, "Convergence analysis for buffered block forward-backward (BBFB) method applied to EFIE ," *Proceedings of IEEE AP-S International Symposium*, Albuquerque, United States, July 2006.

Application of the Stochastic Second-degree Iterative Method to EM Scattering from Randomly Rough Surfaces

Y. Du[1] **and J. A. Kong**[2]

[1]The Electromagnetics Academy at Zhejiang University, Zhejiang University, Hangzhou 310058, China
[2]Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Abstract— For a deterministic linear system, if the spectrum of the first degree iterative process lies within the circle centered at one half with radius one half in the complex plane, then the simple stationary second degree method due to Pillis can improve the asymptotic convergence rate over the first degree method. For linear systems with random matrices, however, Pillis' approach has to be modified. One recent modification is the Stochastic Second Degree (SSD) method. This paper presents an application of the SSD method to electromagnetic scattering from randomly rough surfaces. When in combination with the popular banded matrix canonical grid (BMIA/CAG) method for two-dimensional scattering for PEC case, with the Jacobi-Richardson shift preconditioning, the resulting SSD-BMIA method can improve convergence for the outer iteration over while maintaining identical convergence properties for the inner iteration of the BMIA method. The computational attractiveness of the BMIA method is well preserved. The cost for the convergence improvement is only an additional storage of 3 n-vectors, or in order in general, where is the number of total unknowns. Numerical illustrations are presented to illustrate the effectiveness of the proposed SSD-BMIA method.

In numerical study of electromagnetic wave scattering from randomly rough surfaces, a popular method is the banded matrix iterative approach/canonical grid (BMIA/CAG) approach due to Tsang et al. [1]. In this approach, the impedance matrix from Green' function is decomposed into a sum of a banded matrix, which recognizes the coherent mutual interaction for two points in the neighborhood of each other, and a Taylor expanded flat surface matrix, which is acted upon the observation that outside the neighborhood the Green's function connecting the two points on the rough surface is close to that of a flat surface, the considered canonical grid, and the Green' function is approximated by a Taylor expansion. The advantage of this approach is that, for the banded matrix part, its product with a surface current column vector can be more efficiently computed than that of a full matrix version, while for the Taylor expanded flat surface matrix part, its product with the vector can be speedily computed by the fast Fourier transform (FFT).

It is observed that the bandwidth b_w is usually much smaller than N, the order of the matrix [1]. The bandwidth depends on the surface rms height and correlation length besides surface length. Yet such dependence has not been explicitly quantitatively established in the literature. Accordingly, the bandwidth b_w is an adjustable parameter, and usually is adjusted in an ad hoc manner. For instance, in all the simulated cases in [1, 2], b_w is set to be one tenth of the surface length L, except the case $L = 8000\lambda$ where b_w is one half of one tenth of L.

This ad hoc setting may reduce the effiency that BMIA/CAG approach intends. For instance, if scheme A (defined in [2]) is to be used, and if conjugate gradient method (CG) is used for solving the inner iteration, then in calculating the product of the banded matrix and a vector, the computation needed is Nb_w, if b_w is proportional to N as implicitly assumed in [1, 2], then this computation is in the order of $O(N^2)$, which renders the performance gain of the product of the Taylor expanded flat surface matrix with a vector via FFT to be of little use. In general, the larger the bandwidth b_w, the faster the convergence of the outer iteration, but the more demanding the computation of the inner iteration when method such as CG is used. This clearly presents a tradeoff situation.

In numerical simulations, for a given realization, when b_w is specified, scheme A leads to a stationary first-degree iterative linear system because it corresponds to a so-called two-part splitting of the impedance matrix. It may converge very slowly. Methods have been proposed to improve the convergence rate of a general first-degree system [3–5]. These methods rely on using higher-degree systems, and most of them require explicit information of the spectrum of the iteration matrix B. The tradeoff is to balance the convergence rate and the complexity associated with higher-degree systems. The approach by Pillis [3] seems to be an interesting candidate because it uses a stationary second degree method, or so-called three-part splitting, that is simple in structure and closely mimics the first degree system, while the condition when the second-degree method is superior to a corresponding first-degree method is well established.

The key task in Pillis' approach is to find the optimal value for a structural parameter s which is assumed to be real constant. A procedure for performing such task is indicated in Theorem 4.2 of [3], where the range of s is decomposed into a finite union of mutually exclusive intervals. Such procedure may be applicable for any specific realization of the surface only if the spectrum of the iteration matrix B is completely known. This realization-wise treatment with stringent requirement on spectrum information, however, is not practical if one keeps in mind that the order of B is in general very large. A feasible way is to determine an optimal s for the whole set of realizations while only requiring a probability density of the spectrum rather than explicit knowledge. However, in this approach, the procedure in determining the optimal s breaks down. The reason is that decomposition of the range into a finite union of mutually exclusive intervals is impossible since one has no clue about the remote value any any more. To deal with such technical diculties, we have proposed an extension for the random spectrum case, which is called the stochastic second-degree method (SSD) [6]. In the following, we shall briefly review BMIA/CAG and SSD BMIA before numerical simulations.

Consider a tapered plane wave incident upon a one-dimensional rough surface with the Dirichlet boundary condition. The EFIE formulation is

$$\mathbf{Zu} = \mathbf{b} \tag{1}$$

where $\mathbf{Z}$ is the impedance matrix [1]. In the BMIA/CAG approach, $\mathbf{Z}$ is decomposed into a strong part and a weak part according to bandwidth b_w. In [2], three different schemes of BMIA/CAG are discussed. In this paper we choose scheme A for its relatively good convergence behavior. Its iterative sequence is

$$\left[\mathbf{Z}^{(\mathbf{s})} + \mathbf{Z}^{(\mathbf{w})(0)}\right] \mathbf{u}^{(k+1)} = \Delta\mathbf{Z}^{(\mathbf{w})(0)}\mathbf{u}^k + \mathbf{b} \tag{2}$$

and the corresponding iteration matrix B is defied as $\mathbf{B} = -\left[\mathbf{Z}^{(\mathbf{s})} + \mathbf{Z}^{(\mathbf{w})(\mathbf{0})}\right]^{-1} \Delta\mathbf{Z}^{(\mathbf{w})(0)}$. For scheme A to converge, the spectral radius of B must be less than one, i.e., $\rho(\mathbf{B}) < 1$. The convergence rate is defined as $R = -\log_{10}(\rho(\mathbf{B}))$.

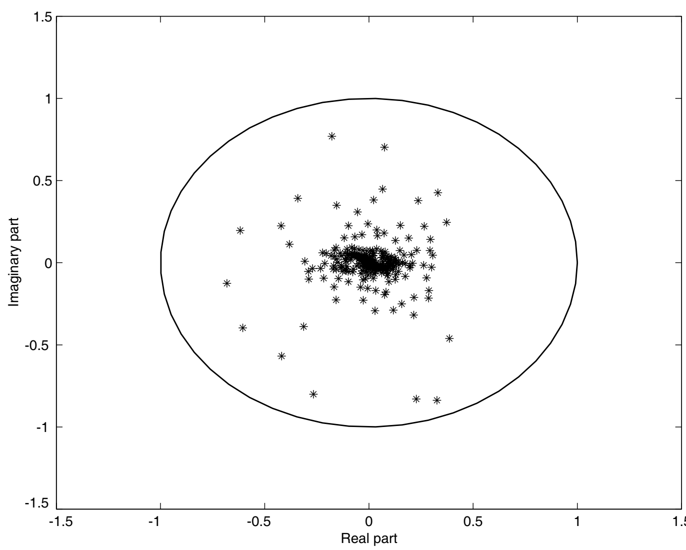


Figure 1: Spectrum of the iteration matrix B of the first-degree method.

The typical spectrum of the iteration matrix B in BMIA/CAG is shown in Fig. 1. This spectrum can be made to lie inside the circle $C[\frac{1}{2}, \frac{1}{2}]$, centered at one half, radius one half, by applying the following Jacobi-Richardson shift with $\mu = 1/2$:

$$\tilde{\mathbf{B}} = \mu\mathbf{B} + (1-\mu)\mathbf{I} \tag{3}$$

Upon applying SSD to the J-R shifted linear system, we have the following second-degree iterative sequence [6]

$$\mathbf{x_{k+2}} = \left(\tilde{\mathbf{B}} - \mathbf{I}\right)\left(\frac{1}{s+1}\mathbf{x_{k+1}} + \frac{s}{s+1}\mathbf{x_k}\right) + (1-s)\,\mathbf{x_{k+1}} + s\mathbf{x_k} + \tilde{\mathbf{b}} \tag{4}$$

If we further split the impedance matrix in the BMIA manner, the resulting SSD-BMIA iterative sequence is

$$\begin{aligned}\left[\mathbf{Z}^{(\mathbf{s})}+\mathbf{Z}^{(\mathbf{w})(\mathbf{0})}\right]\mathbf{x_{k+2}} \;=\; & \mu\left\{\Delta\mathbf{Z}^{(\mathbf{w})(\mathbf{0})}-\left[\mathbf{Z}^{(\mathbf{s})}+\mathbf{Z}^{(\mathbf{w})(\mathbf{0})}\right]\right\}\left(\frac{1}{s+1}\mathbf{x_{k+1}}+\frac{s}{s+1}\mathbf{x_k}\right)\\ & +\left[\mathbf{Z}^{(\mathbf{s})}+\mathbf{Z}^{(\mathbf{w})(\mathbf{0})}\right]\left((1-s)\,\mathbf{x_{k+1}}+s\mathbf{x_k}\right)+\mu\mathbf{b}\end{aligned} \tag{5}$$

Table 1: Four study cases.

case	rms height	correlation length	number of unknowns	bw	s
A	0.6	1.0	400	40	0.38
B	0.6	1.0	1000	65	0.36
C	0.8	1.5	400	80	0.35
D	0.8	1.5	1000	105	0.32

Table 2: Norms of absolute errors $\|\mathbf{Zx_n}-\mathbf{b}\|$ at outer iteration for one realization.

	Case A		Case B		Case C		Case D	
Outer Iteration	BMIA	SSD-BMIA	BMIA	SSD-BMIA	BMIA	SSD-BMIA	BMIA	SSD-BMIA
1	3.712131	5.969413	3.478387	9.031284	2.443272	5.764647	1.886743	8.960752
2	3.599623	3.020069	2.556504	4.053206	0.699749	2.615431	1.389548	3.894310
3	2.682279	1.626435	2.012388	1.772480	0.434127	1.123675	1.901385	1.636694
4	2.357035	0.893544	1.548403	0.790791	0.384351	0.465236	2.520178	0.660772
5	2.122297	0.532078	1.213952	0.359411	0.335911	0.187297	2.699463	0.273852
6	1.845258	0.338935	1.107822	0.215672	0.315709	0.080039	2.585077	0.104957
7	1.375795	0.186412	0.826674	0.140146	0.273459	0.036662	2.338471	0.047173
8	0.988561	0.134181	0.576072	0.095018	0.235526	0.019241	2.104365	0.026054
9	0.808514	0.090214	0.433918	0.069321	0.205354	0.010531	1.922375	0.011727
10	0.763027	0.059841	0.436098	0.050532	0.174260	0.005734	1.758164	0.006612
11	0.762864	0.036130	0.481578	0.036410	0.152071	0.003401	1.600559	0.003551
12	0.844536	0.023694	0.513724	0.025885	0.132400	0.002072	1.452021	0.002123
13	0.820338	0.015693	0.502202	0.017968	0.116800	0.001346	1.316409	0.001207
14	0.790330	0.009412	0.448761	0.012179	0.103582	0.000858	1.195237	0.000629
15	0.751307	0.005903	0.388786	0.008380	0.091654	0.000554	1.086476	0.000375
16	0.673698	0.003702	0.352597	0.005989	0.081462	0.000357	0.987959	0.000182
17	0.621855	0.002280	0.341132	0.004335	0.072073	0.000228	0.898305	0.000113
18	0.582980	0.001294	0.338236	0.003102	0.063932	0.000151	0.816712	0.000051
19	0.543754	0.001023	0.329532	0.002183	0.056602	0.000098	0.742550	0.000032
20	0.528871	0.000782	0.311811	0.001524	0.050135	0.000066	0.675165	0.000014

Now we are in a position to discuss the convergence properties and complexity of the SSD-BMIA approach. This approach is expected to be superior to BMIA in terms of convergence. This improvement comes at a minor additional cost on storage, but not computation. This is because, comparing the right sides of 5 for SSD-BMIA and 2 for BMIA/CAG, on the surface three additional matrix-vector products, namely, $\Delta\mathbf{Z}^{(\mathbf{w})(0)}\mathbf{x_k}$, $[\mathbf{Z}^{(\mathbf{s})}+\mathbf{Z}^{(\mathbf{w})(0)}]\mathbf{x_{k+1}}$ and $[\mathbf{Z}^{(\mathbf{s})}+\mathbf{Z}^{(\mathbf{w})(0)}]\mathbf{x_k}$ need to be computed for each outer iteration. Yet $\Delta\mathbf{Z}^{(\mathbf{w})(0)}\mathbf{x_k}$ is the value of $\Delta\mathbf{Z}^{(\mathbf{w})(0)}\mathbf{x_{k+1}}$ in previous outer iteration and can be simply stored for readily usage; $[\mathbf{Z}^{(\mathbf{s})}+\mathbf{Z}^{(\mathbf{w})(0)}]\mathbf{x_{k+1}}$ can be obtained from the last iteration of the inner iteration where conjugate gradient method is applied;

likewise $[\mathbf{Z^{(s)}}+\mathbf{Z^{(w)(0)}}]\mathbf{x_{k+1}}$ can be obtained from the last iteration of the inner iteration in previous outer iteration. As a result, no additional computation of matrix-vector products is required. The additional requirement on storage is 3 n-vectors, or in $O(N)$ order in general.

In the numerical simulations, the rough surfaces are Gaussian processes with Gaussian spectrum. The tapering parameter g is $L/4$, where L is the surface length. Ten unknowns are used per wavelength. The bistatic scattering coefficient is calculated using the average of 4000 realizations. The incident angle is 10 degrees while the scattering angles are from -90 degrees to 90 degrees. To examine the effect of roughness and number of unknowns on the minimal bandwidth b_w for which the iterative scheme A works, and on the resulting optimal parameter s, we consider two surfaces with increasing roughness, and for each surface we consider two surface lengths, resulting four study cases as shown in Table 1. The resulting bandwidth b_w, which is not considered minimal because identification of such value is a difficult task itself but which is some value that grows approximately in $O(\log N)$, is shown in the table for each case. Also shown is the optimal parameter s.

Figure 1 provides an example to visually confirm our assumption about a circularly symmetric spectrum for the iteration matrix B. It is true that a determination of the optimal parameter s depends on an explicit knowledge of the spectrum of B; however, obtaining such information is an extremely dicult task in random matrix theory. In the meanwhile, we expect that the SSD method may provide a gain on convergence rate of several times, but not too much larger, so the cost/complexity associated with accurately characterizing the probability density of the spectrum of B does not justify the slight to moderate gain improvement.

Table 2 compares the norm of absolute error $\|\mathbf{Zx_n}-\mathbf{b}\|$ for both BMIA/CAG and SSD BMIA at outer iteration n up to 20 for the four study cases for one realization. It is observed to be representative of the whole set of realizations. Table 3 compares the spectral radius, convergence rate, and corresponding convergence gain for the four cases.

Table 3: Convergence gain of SSD BMIA for four study cases.

	Case A		Case B		Case C		Case D	
	BMIA	SSD-BMIA	BMIA	SSD-BMIA	BMIA	SSD-BMIA	BMIA	SSD-BMIA
spectral radius	0.943	0.657	0.940	0.702	0.886	0.650	0.909	0.539
convergence rate	0.025	0.183	0.027	0.154	0.053	0.187	0.042	0.269
convergence gain		7.2		5.8		3.5		6.5

ACKNOWLEDGMENT

This work was supported in part by the China National Science Foundation under Grant No. 40571114 and Zhejiang Science Foundation under Grant No. Y106443.

REFERENCES

1. Tsang, L., et al., "Monte Carlo simulations of large-scale problems of random rough surface scattering and applications to grazing incidence with the BMIA/canonical grid method," *IEEE Trans. Antennas Propagat.*, Vol. 43, 851–859, 1995.
2. Tsang, L., et al., *Scattering of Electromagnetic Waves — Numerical Simulations*, John Wiley & Sons, Inc, New York, 2001.
3. Pillis, J. D., "Faster convergence for iterative solutions to systems via three-part splittings," *SIAM Journal on Numerical Analysis*, Vol. 15, 888–911, 1978.
4. Nichols, N. K., "On the convergence of two-stage iterative processes for solving linear equations," *SIAM J. Numer. Anal.*, Vol. 10, 460–469, 1973.
5. Parsons, B. N., "General k-part stationary iterative solutions to linear systems," *SIAM J. Numer. Anal.*, Vol. 24, 188–198, 1987.
6. Du, Y., "Fast convergence stationary second degree iterative method for linear systems with random matrices," in preparation.

Numerical Analysis of SPM and XPM Penalties of the Conventional IM-DD System with NRZ and RZ Format

H. Taga and H. W. Chang
Sun Yat-Sen University, No. 70 Lien Hai Rd., Kaohsiung 80424, Taiwan, China

Abstract— Transmission penalties due to the self-phase modulation (SPM) and the cross-phase modulation (XPM) are numerically studied for the conventional intensity modulation direct detection (IM-DD) system with the non return-to-zero (NRZ) and the return-to-zero (RZ) format. The nonlinear Schrodinger equation and the split-step Fourier method are utilized for the numerical simulations.

1. INTRODUCTION

The optical fiber communication system is wide-spread over the world, and it supports the vast traffic such as the internet every day. The commercially available optical fiber communication system utilizes intensity modulation direct detection (IM-DD) technology, and the digital bit patterns are the non return-to-zero (NRZ) format in general. For the long-haul system like the undersea cable system, the return-to-zero (RZ) format is also used commercially [1]. The merit of the RZ format is the reduction of the inter-symbol interference, because the consecutive marks are isolated each other. In addition to that, for the case of the wavelength division multiplexing (WDM) system, the RZ format is considered to be more tolerant to the cross phase modulation (XPM), because the pulse overlap between the neighboring channels is less than that of the NRZ case. On the other hand, the RZ format is considered to be more sensitive to the self phase modulation (SPM), because the pulse peak power is twice as much as that of the NRZ format.

In this paper, we consider the effects of the SPM and the XPM for the NRZ format and the RZ format. The numerical simulator calculates the nonlinear Schrodinger equation using the split step Fourier method [2]. The results show the difference of the SPM and the XPM for the NRZ format and the RZ format.

2. NUMERICAL METHOD

In order to evaluate the optical pulse propagation in the optical fiber, the nonlinear Schrodinger equation should be solved. A generic style of the nonlinear Schrodinger equation is:

$$\frac{\partial A}{\partial z} = -\frac{\alpha}{2}A - \frac{i}{2}\beta_2\frac{\partial^2 A}{\partial T^2} + \frac{1}{6}\beta_3\frac{\partial^3 A}{\partial T^3} + i\gamma\left|A\right|^2 A \tag{1}$$

where A denotes the electrical field of the signal light, z denotes the length of the optical fiber, T denotes the time, α denotes the loss coefficient of the optical fiber, β_2 denotes the second derivative of the propagation constant of the optical fiber (and it corresponds to the group velocity dispersion), β_3 denotes the third derivative of the propagation constant of the optical fiber (and it corresponds to the group velocity dispersion slope), and γ denotes the fiber nonlinearity coefficient.

In order to calculate the formula (1), the split-step Fourier method is generally used [2]. The right hand side of the formula (1) contains the linear terms and the nonlinear term. The first to the third terms are the linear terms, and the forth term is the nonlinear term. In the split-step Fourier method, the perturbation theory is applied, and the linear terms and the nonlinear term are calculated separately. At first, the linear terms are calculated assuming that the nonlinear term is constant. Then, the nonlinear term is calculated assuming that the linear term is constant. Repeating this procedure with a small step of the fiber length, the pulse propagation described by the nonlinear Schrodinger equation can be calculated numerically.

3. SYSTEM CONFIGURATION FOR THE SIMULATION

We evaluated the long-haul system using the simulator. Figure 1 shows a schematic diagram of the simulated system. There were sixteen modulated channels at the transmitter, and the channel separation was set to be 0.6 nm. The signal wavelengths were ranged from 1545.5 nm to 1554.5 nm. The bit-rate was 10 Gbit/s, and the pseudorandom 7th pattern was used to generate the modulated signal. Both the NRZ format and the RZ format were used for the simulation. For the optical fiber transmission line, the repeater span length was 50 km, and the loss of the fiber was 0.21 dB/km. The effective area of the fiber was 65 μm^2, and the nonlinear refractive index was 2.6×10^{-20}. The repeater had the output power of +9 dBm and the noise figure of 4.5 dB. The dispersion management [3] was adopted for the optical fiber transmission line, and ten spans of the fiber composed one block. In each block, the first span to the fifth span and the seventh span to the tenth span were the fibers with −2 ps/km/nm chromatic dispersion, and the sixth span was the fiber with +18 ps/km/nm chromatic dispersion. As a result, the whole system had zero chromatic dispersion at 1550 nm. Figure 2 shows the variation of the cumulative chromatic dispersion along the fiber length. At the receiver, the electrical bandwidth was limited to 7.5 GHz by the third order Bessel filter. The cumulative dispersion due to the transmission was fully compensated at the terminal. 50% of the cumulative dispersion was pre-compensated, and the residual 50% was post-compensated. The performance of the received signal was evaluated by the Q-factor [4].

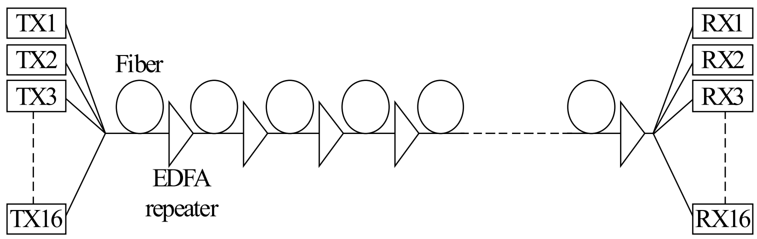


Figure 1: A schematic diagram of the simulated system.

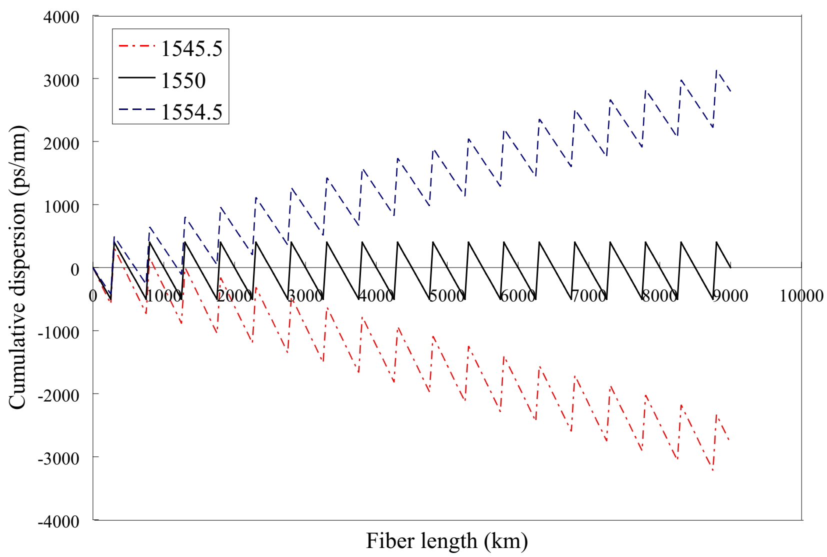


Figure 2: Variation of the cumulative chromatic dispersion.

4. RESULTS AND DISCUSSIONS

The transmission performance was evaluated as a function of the fiber length. In order to clarify the penalties due to the SPM and the XPM, the degradation of the Q-factor from the reference was calculated. As the reference, the Q-factor with the amplifier noise and the chromatic dispersion without the nonlinear degradation was used. Figure 3 shows the results of the numerical simulation. The horizontal axis shows the fiber length, and the vertical axis shows the penalty from the reference Q-factor in dB scale. In this figure, averaged penalty over sixteen signal channels was used.

As shown in Figure 3, the penalty due to the XPM is larger than that of the SPM in short distance, but the penalty due to the SPM becomes much larger than that of the XPM in long distance. The reason why the XPM penalty is larger than the SPM penalty in short distance is that the XPM is twice as much effective as the SPM [2]. On the contrary, the reason why the XPM penalty becomes smaller than the SPM penalty in long distance is that the neighboring bits causing the XPM are slipping out due to the difference of the group delay and the effect of the XPM becomes randomly varied over the entire bit pattern.

Comparing the NRZ format and the RZ format, the RZ format shows smaller penalties for both the SPM and the XPM. As expected, the RZ format shows smaller penalty in the XPM, and the reason is the smaller overlap between the optical pulses. On the other hand, the RZ format also shows smaller penalty in the SPM, even though the pulse peak power is twice as much as that of the NRZ format. The reason why the RZ format shows smaller penalty in the SPM can be deduced that each pulse of the RZ format is isolated. For the NRZ format, the waveform degradation of one pulse in the continuing marks impacts to the neighboring pulses, but it does not affect to the neighboring pulses for the RZ format.

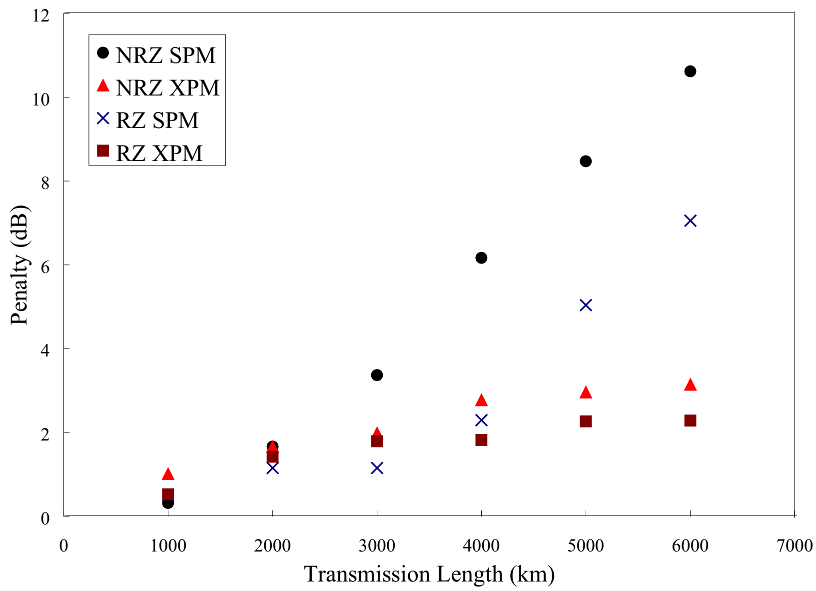


Figure 3: Transmission penalties due to the SPM and the XPM.

5. CONCLUSIONS

We have numerically studied the effect of the SPM and the XPM in the NRZ format and the RZ format. Even though its higher pulse peak power, the RZ format shows better performance against for both the SPM and the XPM for the long distance transmission. It can be said that the RZ format is suitable for the long-haul optical fiber communication system such as undersea fiber cable system.

REFERENCES

1. Bergano, N. S. and C. R. Davidson, "Wavelength division multiplexing in long-haul transmission systems," *IEEE J. of Lightwave Technol.*, Vol. 14, No. 6, 1299–1308, 1996.
2. Agrawal, G. P., *Nonlinear Fiber Optics*, Academic Press, San Diego, 1995.
3. Bergano, N. S., "Wavelength division multiplexing in long-haul transoceanic transmission systems," *IEEE J. of Lightwave Technol.*, Vol. 23, No. 12, 4125–4139, 2005.
4. Bergano, N. S., F. W. Kerfoot, and C. R. Davidson, "Margin measurements in optical amplifier systems," *IEEE Photon. Technol. Lett.*, Vol. 5, No. 3, 304–306, 1993.

Dispersion Relationships of AlGaInAs-InP DBR Gratings Using Floquet-Bloch Theory

Jiun-Jie Liau, **Jia-Huei Lin**, **Min-Xiu Jiang**, **and Nai-Hsiang Sun**
Department of Electrical Engineering, I-Shou University, Kaohsiung, Taiwan, China

Jerome K. Butler and Gary A. Evans
Department of Electrical Engineering, Southern Methodist University, Dallas, TX 75275-0338, USA

Abstract— Floquet-Bloch theory is used to calculate the dispersion relationship and the spectrum of first order contra-directional coupling in the AlGaInAs system. A structure of DBR gratings is calculated in this paper. Results from the Floquet-Bloch theory show the region of stopband. The maximum reflection efficiency of the structure is obtained.

1. INTRODUCTION

To date, the $\mathrm{Al}_x\mathrm{Ga}_y\mathrm{In}_{1-x-y}\mathrm{As}$ system is widely applied to long-wavelength semiconductor lasers due to its lager conduction band offset compared with that of the conventional $\mathrm{Ga}_x\mathrm{In}_{1-x}\mathrm{As}_y\mathrm{P}_{1-y}$ system. This material system results in better high-temperature performance; for example, it reduces carrier leakage from the quantum-well region under high temperature operation [1]. Recently, the $\mathrm{Al}_x\mathrm{Ga}_y\mathrm{In}_{1-x-y}\mathrm{As}$ system is applied to grating-outcoupled surface emitting (GSE) lasers with single wavelength of 1310 nm [2]. The geometry of a GSE laser consists of a ridge waveguide, a grating outcoupler and a first order distributed Bragg reflector (DBR). The first-order DBR gratings provide wavelength selective feedback to the laser cavity.

To design the first order DBR gratings, we need to calculate the dispersion relationship of AlGaInAs-InP waveguides. The Floquet-Bloch theory provides rigorous solutions with high accuracy for the structure of corrugated waveguides. In this paper, the Floquet-Bloch theory (FBT) is used to calculate the dispersion curves of the first Bragg gratings. Moreover, by using the Mahmoud-Beal's method and the FBT, the reflection and transmission characteristics of first-order DBR gratings can be obtained [3]. This study is helpful for us to analyze and design first-order gratings for the $\mathrm{Al}_x\mathrm{Ga}_y\mathrm{In}_{1-x-y}\mathrm{As}$ system.

2. PROBLEM FORMULATION AND RESULTS

Assume that a wave with time dependence of the form $\exp(j\omega t)$ propagates in the z direction. By applying the Floquet-Bloch theory, the forward propagated and backward propagated electric fields for TE mode are expressed as

$$E_y(x,z) = \sum_{n=-\infty}^{\infty} f_n(x)\cdot\exp(-jk_{kz}z) = E_{yf} + E_{yb} \tag{1}$$

$$E_y^b(x,z) = \sum_{n=-\infty}^{\infty} f_n(x)\cdot\exp(jk_{kz}z) = E_{yf}^b + E_{yb}, \tag{2}$$

respectively, where $k_{zn} = \beta_n + j\alpha = (\beta + nk) + j\alpha$ is the complex propagation constant of the n-th spatial harmonic, $K = 2/\Lambda$ is the grating wavenumber, Λ is the grating period, and n is the space harmonic order. Note that the electric fields in (1) and (2) consist of forward harmonics (E_{yf} and E_{yf}^b) and backward harmonics (E_{yb} and E_{yb}^b). The electric fields should satisfy the scalar Helmholtz equation in each layer.

Consider a temperature insensitive AlGaInAs/InP epitaxial laser structure. Fig. 1 shows the total electric field distributions for TE mode. The free space wavelength is 1310 nm. The refractive index profile is also shown in Fig. 1. For the grating structure, the duty cycle is 50%, and the tooth height is 0.1 μm.

Figure 2 shows the normalized wavenumber as a function of the normalized propagation constant. The grating period is 0.20175 μm. The first Bragg stopband occurs near the $\beta\Lambda = \pi$ region, while the normalized attenuation constant ($\alpha\Lambda$) becomes negative. Since there is no radiation and the dielectric material is lossless, the complex propagation constant indicates the occurrence of contra-directional coupling. The range of the stopband is from the wavelength of 1309.82 nm to the

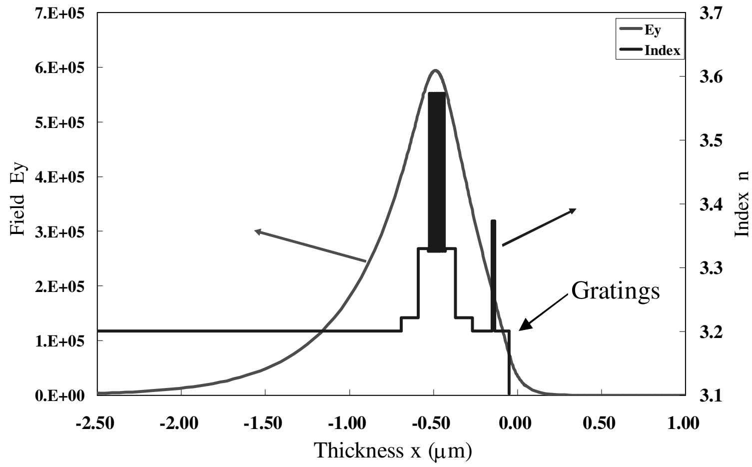


Figure 1: The electric field distribution for TE mode. The index profile is also shown.

wavelength of 1310.02 nm. The strongest coupling occurs at the middle of the stopband with the wavelength of 1309.92 μm. The maximum normalized attenuation constant ($\alpha\Lambda$) is -2.665×10^{-4}. Note that the attenuation is zero if the wavelength is not in the region of stopband.

The reflection and transmission spectrum are shown in Fig. 3. The grating length is 500 μm.

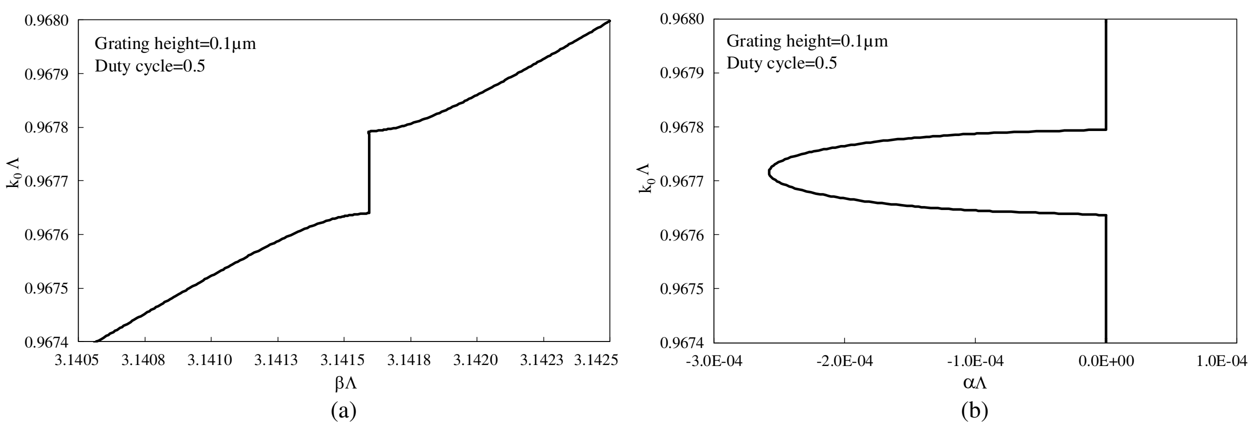


Figure 2: The dispersion curves around the first Bragg resonance. (a) The real part and (b) the imaginary part of the normalized propagation constants.

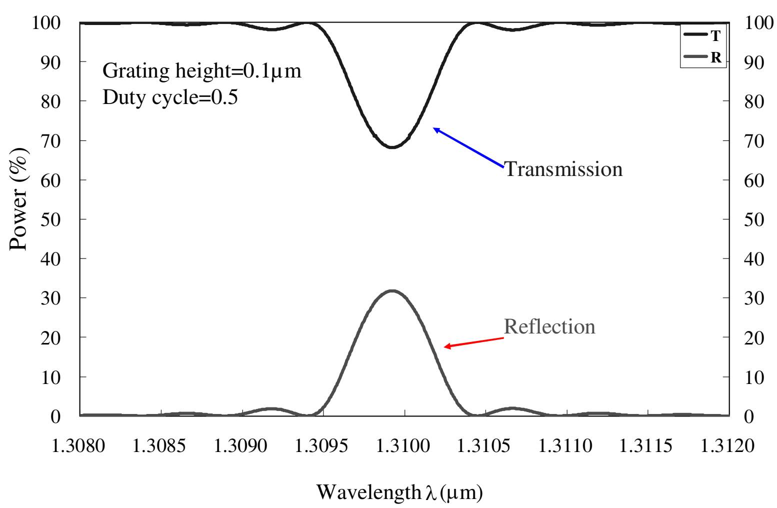


Figure 3: The reflection and transmission power efficiencies as a function of the wavelength.

Fig. 3 shows that the maximum resonance occurs at the wavelength of 1309.92 nm. When the strongest coupling occurs, 31.82% of the power can be reflected, while 68.18% of the power is transmitted through the grating. Note that the resonant region is from the wavelength of 1309.40 nm to the wavelength of 1310.45 nm. The range is slightly larger than that from the dispersion relationship (Fig. 2). As shown in Fig. 1, the power of the fundamental mode in the grating region is small, which yields less coupling strength and reduces the reflection efficiencies.

3. CONCLUSION

In this paper we used the Floquet-Bloch theory and the Mahmoud-Beal's method to calculate the dispersion relationships and the spectrum of a periodic dielectric waveguide for the $\mathrm{Al}_x\mathrm{Ga}_y\mathrm{In}_{1-x-y}\mathrm{As}$ systems. The dispersion curves of a DBR laser structure are analyzed. For the structure of Fig. 1, the maximum reflection power of 31.82% is obtained at the wavelength of 1309.92 with the grating period of 0.20175 μm.

REFERENCES

1. Selmic, S. R., T. M. Chou, J. P. Sih, J. B. Kirk, A. Mantie, J. K. Butler D. Bour, and G. A. Evans, "Design and characterization of 1.3-um AlGaInAs-InP Multiple quantum well lasers," *IEEE Journal on Selected Topics in Quantum Electronics*, Vol. 7, 340–349, 2001.
2. Masood, T., S. Patterson, N. V. Amarasinghe, S. McWilliams, D. Phan, D. Lee, Z. A. Hilali, X. Zhang, G. A. Evans, and J. K. Butler, "Single-frequency 1310-nm AlInGaAs-InP grating-outcoupled surface-emitting lasers," *IEEE Photon. Technol. Lett.*, Vol. 16, 726–728, 2004.
3. Chou, C.-C., N.-H. Sun, J. K. Butler, and G. A. Evans, "Radiation loss of grating-assisted directional couplers using the Floquet-Bloch theory," *Technical Digest CLEO/PR 2003*, 350, Taipei, Taiwan, 2003.

Dispersion Relationships of AlGaInP-InP Surface-emitting Lasers Using Floquet-Bloch Theory

Jiun-Jie Liau[1], **Shih-Han Chen**[1], **Yi-Ann Lin**[1], **Nai-Hsiang Sun**[1]
Jerome K. Butler[2], **and Gary A. Evans**[2]
[1]Department of Electrical Engineering, I-Shou University, Kaohsiung, Taiwan, China
[2]Department of Electrical Engineering, Southern Methodist University, Dallas, TX 75275-0338, USA

Abstract— Floquet-Bloch theory is used to calculate the dispersion relationship and the spectrum of contra-directional coupling in the AlGaInAs system. A structure of GSE gratings is calculated in this paper. The characteristics of reflection and radiation near resonance coupling are analyzed in this paper.

1. INTRODUCTION

Recently, the $Al_xGa_yIn_{1-x-y}As$ system [1] is applied to grating-outcoupled surface emitting (GSE) lasers with single wavelength of 1310 nm [2]. The characteristics of GSE lasers include narrow beam divergence, single wavelength emission and high output power. The geometry of a GSE laser consists of a ridge waveguide, a grating outcoupler and a first order distributed Bragg reflector (DBR). The second-order gratings perform an outcoupler to provide wavelength selective surface emission.

The dispersion relationship of AlGaInAs-InP waveguides is very important in designing the second order gratings. The Floquet-Bloch theory provides rigorous solutions with high accuracy for the structure of corrugated waveguides. In this paper, the Floquet-Bloch theory (FBT) is applied to calculate the dispersion curves of the second Bragg gratings. By using the Mahmoud-Beal's method and the FBT, the reflection, transmission and radiation characteristics of second-order gratings can be obtained [3]. This study is helpful for us to analyze the second-order gratings for the $Al_xGa_yIn_{1-x-y}As$ system.

2. PROBLEM FORMULATION

Figure 1 shows that the fundamental mode (e_y) is launched at region I towards the grating. This produces the reflected and transmitted fields. As shown in Fig. 1, the reflection and transmission

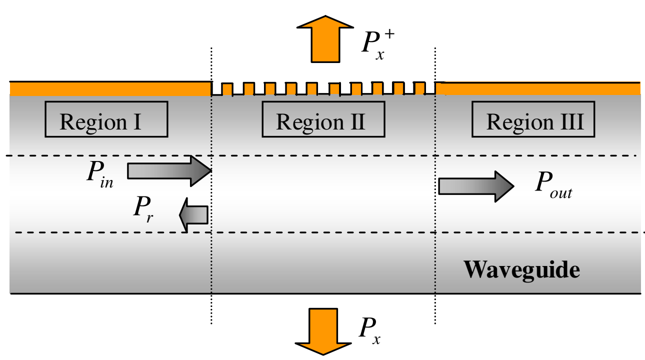


Figure 1: The basic geometry of a corrugated periodic waveguide.

power indicate the power in region III and region I, respectively. When the power radiates into the air (P_x^+) and the substrate (P_x^-) at the second-order Bragg condition, the expressions of radiation power in the air and substrate are given by

$$P_x^+ = \frac{1}{2}\mathrm{Re}\left[\int_0^L E_y \cdot H_z^* dz\right]_{\mathrm{air}}^{n=-1} \tag{1}$$

$$P_x^- = \frac{1}{2}\mathrm{Re}\left[\int_0^L E_y \cdot H_z^* dz\right]_{\mathrm{substrate}}^{n=-1} \tag{2}$$

where L is the length of the grating region, and E_y is the electric field in the grating region. Assume that a wave with time dependence of the form $\exp(j\omega t)$ propagates in the z direction. By using the Floquet-Bloch theory, the electric field for TE mode in region II are expressed as

$$E_y(x,z) = \sum_{n=-\infty}^{\infty} f_n(x) \cdot \exp(-jk_{kz}z), \tag{3}$$

where $k_{zn} + \beta_n + j\alpha = (\beta + nK) + j\alpha$ is the complex propagation constant of the n-th spatial harmonic, $K = 2/\Lambda$ is the grating wavenumber, Λ is the grating period, and n is the space harmonic order. Note that only the $n = -1$ harmonic is radiated at the second Bragg. Therefore, in (1) and (2), the $n = -1$ space harmonic of E_y and H_z are calculated.

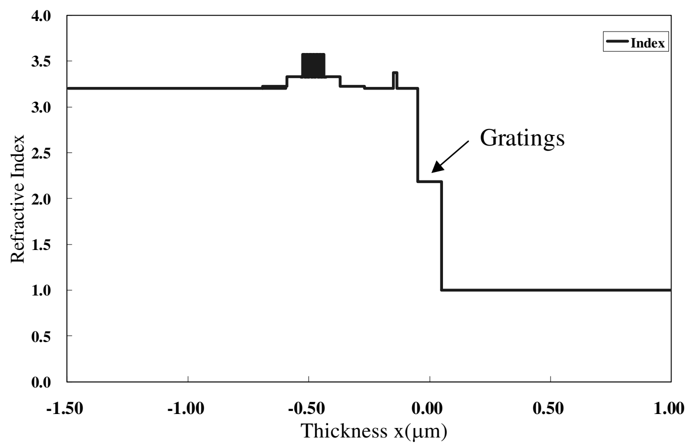


Figure 2: The index profile of a GSE Laser.

3. RESULTS

Figure 2 shows the index profile of a GSE laser for a AlGaInAs/InP system. For the grating structure, the duty cycle is 50%, and the tooth height is 0.1 μm. Fig. 3 shows the normalized wavenumber as a function of the normalized propagation constant. The grating period is 0.4035 μm. The second Bragg occurs at $\beta\Lambda = 2\pi$. Since the $n = -1$ harmonic is in the fast wave region, the attenuation constant is not zero over the calculated region. The dispersion curve in Fig. 2(a) contains a stopband. The range of the stopband is from 1310.20 nm to 1310.77 nm. The dielectric material is assumed to be lossless. Therefore, the non-zero attenuation constant indicates that the power radiates to the superstrate and the substrate near the region of the second Bragg. The strongest coupling occurs at $\beta\Lambda = 2\pi$ where the wavelength is 1310.31 μm, and the normalized attenuation $(\alpha\Lambda)$ is -4.9675×10^{-5}. Note that the strongest coupling corresponds to the minimum absolute value of the attenuation.

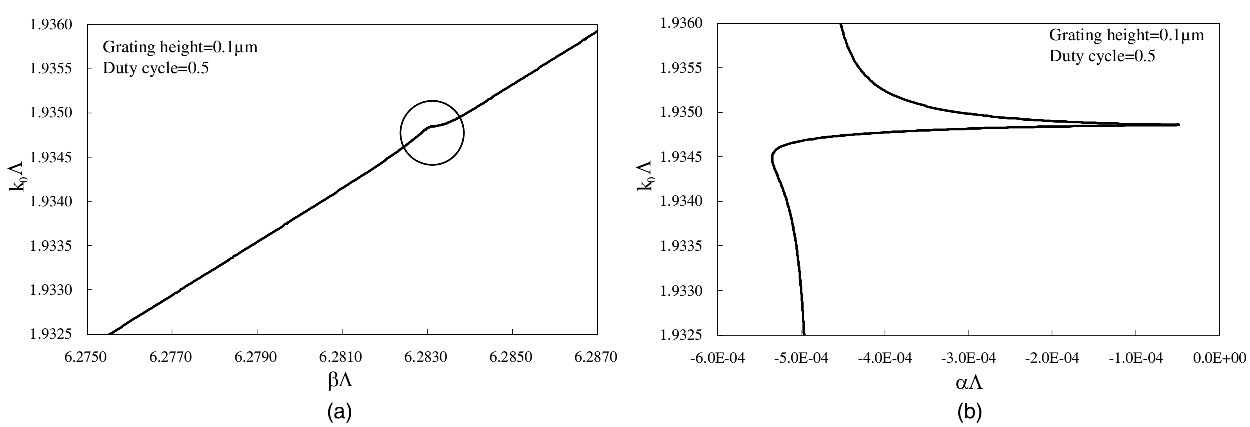


Figure 3: The dispersion curves around the second Bragg resonance. (a) The real part and (b) the imaginary part of the normalized propagation constants.

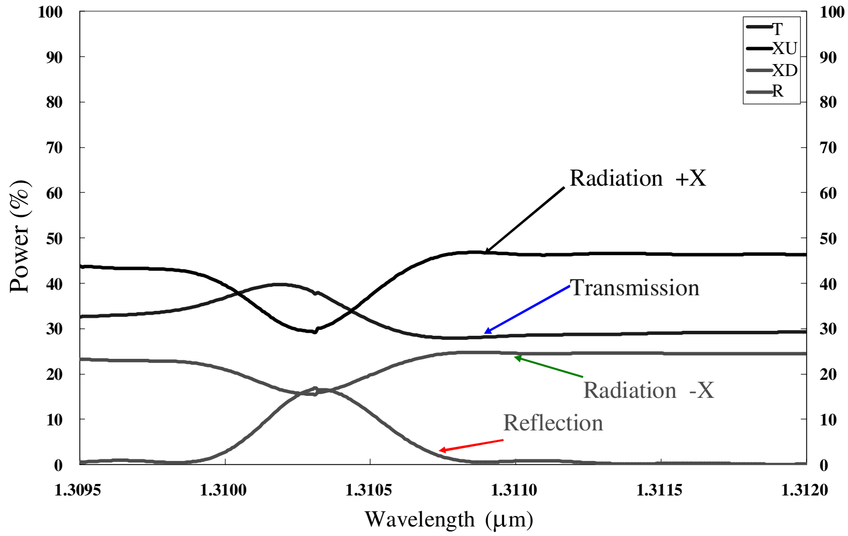


Figure 4: The transmission, reflection, power radiation to the superstrate ($+x$) and substrate ($-x$) as a function of the wavelength.

The reflection and transmission spectrum, and the radiated power in the superstrate ($+x$ direction) and the substrate ($-x$ direction) are shown in Fig. 4. The grating length is 500 μm. When the strongest coupling occurs, 17% of the power can be reflected, and 38% of the power is transmitted through the grating, while 29% and 16% of the power radiate to the superstrate and the substrate, respectively. The resonant region in Fig. 4 is from 1309.83 nm to 1310.89 nm. As shown in Fig. 4, the stopband represents the occurrence of contra-directional coupling, whereas no reflection occurs outside the stopband. That means the attenuation contributes reflection and radiation at resonance, whereas the attenuation represents radiation for off-resonance coupling. As shown in Fig. 4, at wavelength of 1309.5 nm, the transmission, radiation to the superstrate, and radiation to the substrate are 33%, 43% and 24%, respectively. The reflection is very small at the wavelength of 1309.5 nm.

4. CONCLUSION

In this paper we used the Floquet-Bloch theory and the Mahmoud-Beal's method to calculate the dispersion relationships and the spectrum of second order gratings for the $Al_xGa_yIn_{1-x-y}As$ systems. The dispersion curves of a GSE laser structure are analyzed. The characteristics of transmission, reflection and radiation at the second Bragg are discussed in this paper.

REFERENCES

1. Selmic, S. R., T. M. Chou, J. P. Sih, J. B. Kirk, A. Mantie, J. K. Butler, D. Bour, and G. A. Evans, "Design and characterization of 1.3 μm AlGaInAs-InP multiple quantum well lasers," *IEEE Journal on Selected Topics in Quantum Electronics*, Vol. 7, 340–349, 2001.
2. Masood, T., S. Patterson, N. V. Amarasinghe, S. McWilliams, D. Phan, D. Lee, Z. A. Hilali, X. Zhang, G. A. Evans, and J. K. Butler, "Single-frequency 1310-nm AlInGaAs-InP grating-outcoupled surface-emitting lasers," *IEEE Photon. Technol. Lett.*, Vol. 16, 726–728, 2004.
3. Liau, J.-J., C.-C. Chou, S.-C. Chang, N.-H. Sun, J. K. Butler, and G. A. Evans, "Analysis of second order gratings in grating-outpuled surface-emitting lasers," *OECC 2006*, 5P-16, Kaohsiung, Taiwan, 2006.

Avalanche Photodetector Design for the Mid-infrared Using GaSb-based Alloys

R. Bhatia[1], **M. Grzesik**[1], **S. Vangala**[1], **K. Vacarro**[2], **W. Goodhue**[1], **and C. Armiento**[1]

[1]University of Massachusetts Lowell, Lowell, MA, USA
[2]Air Force Research Laboratory, Sensors Directorate, Hanscom AFB, Massachusetts, USA

Abstract— Avalanche photodiodes (APDs) operating at wavelengths between 2–5 μm can be realized using alloys of InGaAsSb and AlGaAsSb lattice matched to GaSb substrates. This paper presents APD device models in this material system. Separated Absorption Multiplication (SAM) APD structures are based on an InGaAsSb absorption region and an AlGaAsSb multiplication region. A major challenge with modeling in these devices is the limited availability of material data, particularly the impact ionization coefficients. Preliminary efforts to measure the impact ionization coefficients, as well as progress in the growth of these alloys using MBE, will be discussed.

1. INTRODUCTION

The quaternary alloys based on GaSb are well suited for the development of Avalanche Photodetectors (APDs) operating in the 2–5 μm wavelength range. A challenge with APD development in small bandgap semiconductors is the large dark current that can result from tunneling. One approach used in shorter wavelength devices (such as InP-based APDs) is to separate the multiplication and light absorption functions using a Separate Absorption and Multiplication (SAM) APD. This concept has particular promise in APDs based on the small bandgap semiconductors that are used for long wavelength applications. The SAM-APD design requires that the *p*-*n* junction be formed in a large band gap material ($E_g > 1\,\text{eV}$) and a separate narrow band gap layer absorbs the wavelength of interest. Carriers absorbed in this region are then injected into the high field region (larger bandgap material) for avalanche multiplication.

The work described in this paper involves the design of SAM-APDs using two quaternary alloys that are lattice matched to GaSb: $\text{In}_x\text{Ga}_{1-x}\text{As}_y\text{Sb}_{1-y}$ and $\text{Al}_x\text{Ga}_{1x}\text{As}_y\text{Sb}_{1-y}$. Multiplication (i.e., impact ionization) takes place in the larger band gap $\text{Al}_x\text{Ga}_{1-x}\text{As}_y\text{Sb}_{1-y}$ layer and the $\text{In}_x\text{Ga}_{1-x}\text{As}_y\text{Sb}_{1-y}$ layer is used to absorb incident photons in the 2–5 μm wavelength range. The operating wavelength range of the APD is determined by the alloy composition of the $\text{In}_x\text{Ga}_{1-x}\text{As}_y$ Sb_{1-y} absorber layer. The composition of the $\text{Al}_x\text{Ga}_{1-x}\text{As}_y\text{Sb}_{1-y}$ alloy should be chosen to maximize the difference between the impact ionization rates of electrons (α) and holes (β). As long as the multiplication is initiated by the carrier with the larger impact ionization coefficient, the multiplication-induced excess noise will be minimized.

In this paper, we report a SAM APD model that was developed for this material system that is based on ATLAS, a device modeling tool (Silvaco Inc.) that is typically used for more conventional semiconductors such as Si and GaAs. We also report progress on the growth of these GaSb-based compounds using Molecular Beam Epitaxy (MBE) as well as efforts to measure the impact ionization rates in these materials.

2. APD DESIGN REQUIREMENTS

An optimized SAM-APD design requires high quantum efficiency, low dark current, and high multiplication with low excess noise (associated with the avalanche multiplication process). These requirements translate into particular material requirements: [1]

1. High quality, low defect density material and junctions to reduce tunneling currents.
2. Maximum difference between the ionization rates of holes and electrons — for lowest excess noise.
3. Appropriate layer doping and thicknesses that enable reverse bias breakdown due to impact ionization rather than tunneling.
4. Doping levels and layer thicknesses chosen to allow the electric field region to extend into the absorber layer at applied voltages near the breakdown voltage.

Achieving these APD design guidelines can be challenging in narrow bandgap (< 0.7 eV) materials such as GaSb and InGaAsSb, which typically suffer from soft breakdown and large leakage currents. These are often caused by band-to-band tunneling and traps that arise from native defects. The relatively immature state of the processing technology in these materials can also result in large dark currents caused by surface leakage currents due to poor junction passivation techniques.

The current state of MBE growth, summarized in Section 6, has imposed some limitations on realistic doping levels to use in our model. At the current time, the achievable doping levels in MBE-grown GaSb are between $10^{16}\,\mathrm{cm}^{-3}$ to $10^{19}\,\mathrm{cm}^{-3}$. We have accounted for these doping limitations in choosing values in our simulations. Our models also assume that step p-n junctions and heterojunctions are modeled as abrupt to simulate devices grown by MBE.

3. MATERIAL PARAMETERS USED FOR DEVICE SIMULATIONS

Modeling of APDs based on GaSb and its related alloys is difficult due to limited knowledge of material parameters, particularly for quaternary alloys. One aspect of the design challenge arises because the material properties of theses alloys depend on their composition. Designers must often estimate material values of alloys based on a simple linear extrapolation of its binary constituents. Although this may not be accurate, it is often the only option available to a designer.

Table 1 lists some of the important material parameters used in our device simulations. The comment column identifies the approach for determining these values. Many of these values were determined by linear interpolation of binary data.

Table 1: Material parameters used for device simulations at room temperature.

Material-> ***Properties***	**GaSb**	**$Al_xGa_{1-x}As_ySb_{1-y}$ x= 0.55 y=0.047**	**$In_xGa_{1-x}As_ySb_{1-y}$ x=0.15 y=0.17**	***Comments***
Band Gap (eV)	0.726 [5]	1.25	0.55	Linear Interpolation
Permittivity (F/cm)	15.7	13.55	15.9	Linear Interpolation
Density of States (cm^{-3}) Conduction Band Valence Band	2.07×10^{17} 1.82×10^{19}	3.03×10^{19} 5.88×10^{18}	1.7×10^{17} 6.42×10^{18}	Calculated from effective mass data
Electron Affinity (eV)	4.06 [2]	3.83 [5]	4.3 [5]	Linear Interpolation
Electron Lifetime (s) **Holes Lifetime (s)**	10^{-7} 10^{-7}	10^{-7} 10^{-7}	10^{-7} 10^{-7}	Assumed
Ionization Rates	$\alpha= 5.46\times10^6 \exp(-3.7\times10^5/E)^1$ $\beta= 2.7\times10^7 \exp(-4.47\times10^5/E)^1$	$\alpha= 1.04\times10^5 \exp(-4.1\times10^5/E)^2$ $\beta= 1.91\times10^5 \exp(-4.12\times10^5/E)^2$	$\alpha= 2.41\times10^6 \exp(-4.45\times10^5/E)^2$ $\beta= 1.98\times10^6 \exp(-3.69\times10^5/E)^2$	E-Electric Field AlGaAsSb data based on ternary data of [8]. GaSb data calculated from [9]. InGaAsSb data from [10]

An example of key material parameters are the impact ionization coefficients. These parameters are critical to predicting the multiplication process in the APD as well as the breakdown voltage of the device. The APD design should allow for the gain (or multiplication process) to be initiated by the carrier (electron or hole) with the larger coefficient. This will yield low excess noise and an optimized time response. At the present time there is very little data on the ionization coefficients in these alloys. In order to model the SAM APD, the published ionization rates of the nearest ternary alloy were used. Efforts are underway in our group, to measure ionization rates for various compositions of $\mathrm{Al}_x\mathrm{Ga}_{1-x}\mathrm{As}_y\mathrm{Sb}_{1-y}$. This will provide more accuracy in the modeling of the long wavelength SAM-APD.

4. LONG WAVELENGTH SAM-APD MODEL DEVELOPMENT

A two-dimensional schematic of a mesa SAM-APD is shown in Figure 1. The device model developed in ATLAS consists of n-type GaSb substrate (Region 1) with an electron carrier concentration of $10^{17}\,\mathrm{cm}^{-3}$. Region 2 is a 5 µm thick n-type $\mathrm{In_{0.15}Ga_{0.85}As_{0.17}Sb_{0.83}}$ absorber with an electron carrier concentration of $4 \times 10^{16}\,\mathrm{cm}^{-3}$. This composition corresponds to a band gap of 0.55 eV and an operating wavelength of approximately 2.25 µm. The wavelength sensitivity is adjustable from 2 to 5 µm depending on the composition of the $\mathrm{In}_x\mathrm{Ga}_{1-x}\mathrm{As}_y\mathrm{Sb}_{1-y}$ absorber region. At the top of the mesa is a 2 µm thick $\mathrm{Al_{0.45}Ga_{0.55}As_{0.05}Sb_{0.95}}$ layer with a band gap of approximately 1.25 eV that forms the p-n junction (multiplication region). The $\mathrm{Al_{0.45}Ga_{0.55}As_{0.05}Sb_{0.95}}$ layer is modeled as two regions. Region 3 is 0.5 µm thick with $n = 6 \times 10^{16}\,\mathrm{cm}^{-3}$. Region 4 is 1.5 µm thick with $p = 5 \times 10^{17}\,\mathrm{cm}^{-3}$. In order to function as a SAM APD, the AlGaAsSb p-n junction must be in close proximity to the AlGaAsSb-InGaAsSb interface [6]. This implies that n-AlGaAsSb must be the lighter doped region. The following lattice matching condition is used in our models as given by Adachi [4];

$$\mathbf{Al}_{\boldsymbol{x}}\mathbf{Ga}_{\mathbf{1}-\boldsymbol{x}}\mathbf{As}_{\boldsymbol{y}}\mathbf{Sb}_{\mathbf{1}-\boldsymbol{y}}: \; y = 0.0396x/(0.4426 + 0.0318x) \text{ for } (0 \leq x \leq 1) \tag{1}$$

$$\mathbf{Ga}_{\boldsymbol{x}}\mathbf{In}_{\mathbf{1}-\boldsymbol{x}}\mathbf{As}_{\boldsymbol{y}}\mathbf{Sb}_{\mathbf{1}-\boldsymbol{y}}: \; y = 0.3835 - 0.3835x/(0.421 + 0.216x) \text{ for } (0 \leq x \leq 1). \tag{2}$$

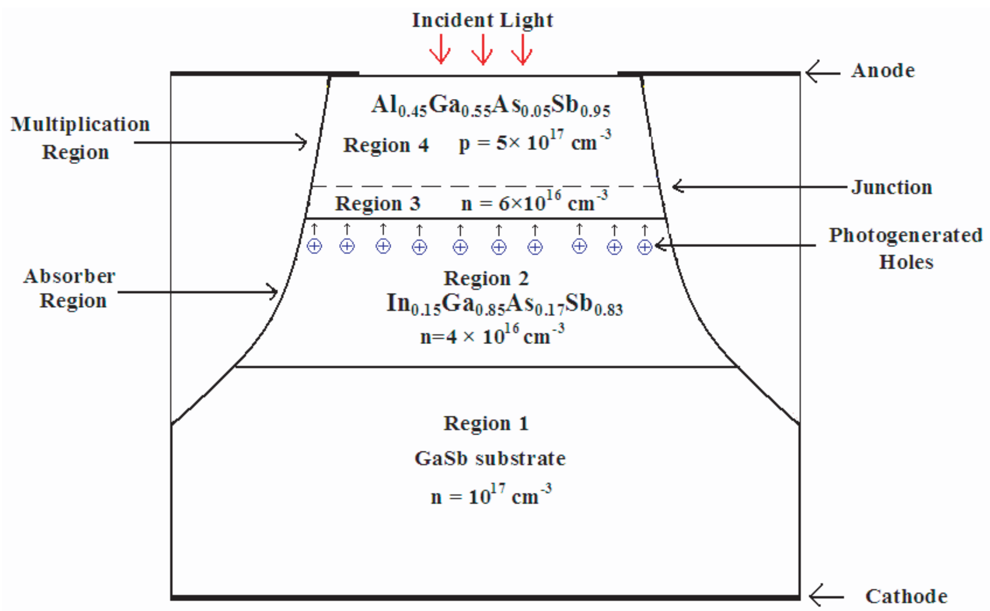


Figure 1: Model of Separate Absorption and Multiplication APD (SAM APD).

The model can predict the measurable characteristics of APD operation such as currentvoltage characteristics, photocurrent and quantum efficiency. In addition, internal operation such as the electric field can be calculated. A 2-D electric field profile along the mesa edge near the junction is shown in Figure 2(a) indicating lower electric fields at the edges. If the device has higher electric field strength at the mesa walls, it would result in premature edge breakdown and subsequent poor performance. A one-dimensional plot of electric field strength along a cutline (passing from the center of the APD) is shown in Figure 2(b). Note that the electric field, which peaks at the p-n junction, is contained within the AlGaAsSb multiplication layer at zero bias. At a reverse bias of 12 V however, the electric field extends into the InGaAsSb absorber layer. The ability to calculate the electric field allows the designer to choose layer doping and thicknesses to adjust the voltage that the electric field punches through into the absorption layer, thereby collecting photocurrent.

In addition to doping and layer thicknesses, energy band discontinuities at the heterojunction interfaces can affect the behavior of SAM-APDs. Discontinuities of the conduction or valence bands at a heterojunction interface may hinder carrier flow and significantly reduce photocurrent. The model can calculate the energy band diagram as shown in Figures 3(a) and (b). The band offset values of the GaSb/InGaAsSb hetero interface are taken as 0.3 eV for the conduction band

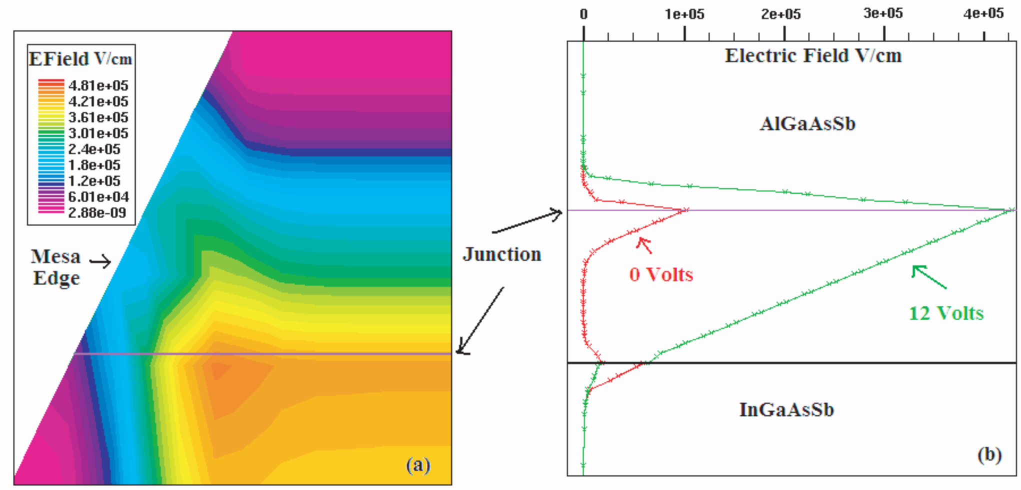


Figure 2: (a) Electric field profile along the mesa edge near p-n junction. (b) Onedimensional plot of Electric Field along the centre of the APD.

and 0.1 eV for the valence band. Band offsets for the InGaAsSb/AlGaAsSb interface are taken as 0.73 eV for the conduction band and 0.05 eV for the valence band [7]. Figure 3(a) indicates that a barrier is present for hole flow in the valence band but no barrier exists for electrons in the conduction band. This barrier hinders photogenerated holes from being injected into the p-n junction, thereby reducing the photocurrent and the gain of the device. In this APD model, the effect of the barrier has been reduced by decreasing the width of the n-doped $Al_xGa_{1-x}As_ySb_{1-y}$ region to about 0.5 µm to allow the electric field to penetrate into the InGaAsSb region as shown in Figure 2(b). The field enables holes to gain sufficient energy to cross over the barrier into the AlGaAsSb layer and multiply by the impact ionization process. At reverse bias close to breakdown voltages, the hole barrier is significantly reduced to allow photogenerated holes in the InGaAsSb to penetrate into the AlGaAsSb region as shown in Figure 3(b). The design should limit the magnitude of the electric field in the InGaAsSb layer to insure that significant tunneling currents do not occur. Stillman et al. have explained this effect in greater detail [1].

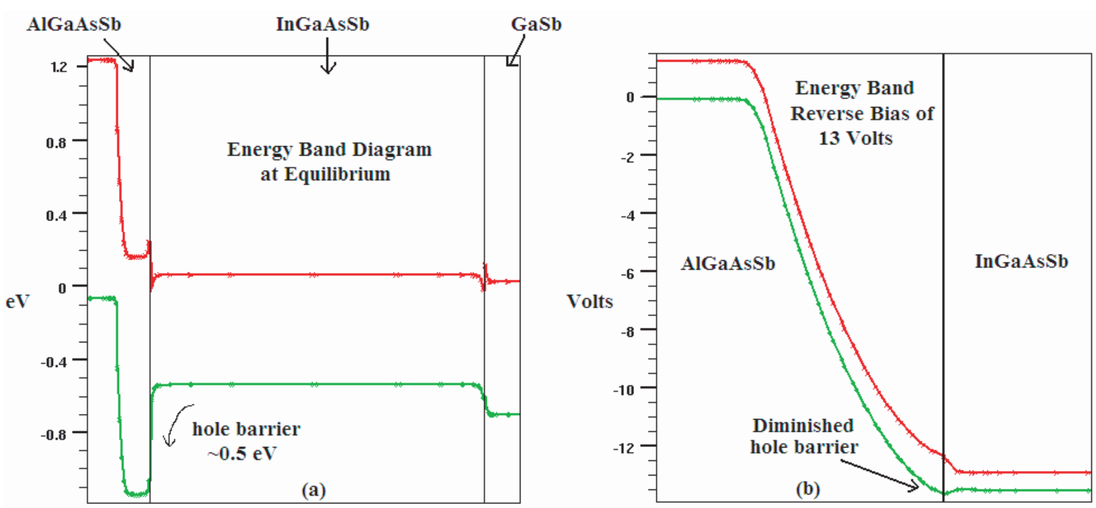


Figure 3: (a) Energy band diagram of SAM APD at Equilibrium. (b) Energy band diagram at the AlGaAsSb/InGaAsSb heterojunction under reverse bias.

5. MODELING RESULTS

Device simulation obtained the I-V characteristics shown in Figure 4. Breakdown occurs at approximately 13 Volts. There is uncertainty in the correct breakdown voltage since the impact ionization coefficients in the AlGaAsSb material system are not precisely known for any composition. These results were generated by and using the impact ionization coefficients from the closest ternary data. However, due to the large band gap of the AlGaAsSb material (1.25 eV), we expect a breakdown of around 12–13 V for high quality growth. Also the doping profile effects breakdown voltage with step *p-n* junctions giving higher breakdown voltages due to their narrow depletion widths. For graded doping profiles the depletion width is wider and spreads out closer to InGaAsSb region reducing the breakdown voltage. Our breakdown values are higher than the experimental I-V data of the SAM APD fabricated by Sulima [6], which yielded breakdown voltages of 6–8 Volts using a narrower band gap composition of $Al_{0.28}Ga_{0.72}As_{0.014}Sb_{0.986}$ as the multiplication region.

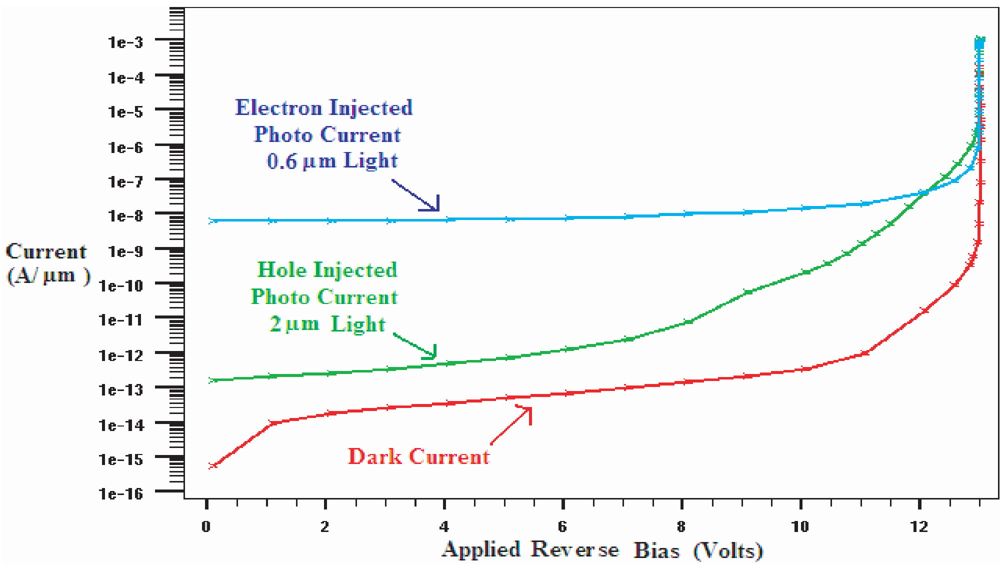


Figure 4: Reverse I-V dark current and photocurrent.

Figure 4 shows photocurrents obtained from a 2 µm light source with an intensity of 0.01 W/cm^2 that passes through the AlGaAsSb region and is absorbed by the narrow band gap InGaAsSb layer. This is the hole injected photocurrent and the effect of the barrier is clearly seen with very low photocurrent generated at low voltages. However, at around 7–8 Volts the electric field starts penetrating into the InGaAsSb region diminishing the barrier (Figure 3(b)) and the photocurrent starts increasing. The quantum efficiency value exceeds 100% at applied voltages above 12 Volts, indicating avalanche multiplication. At 12.5 Volts the quantum efficiency is approximately 300% for this device. The electron injected photocurrent for a 0.6 µm light source with the same intensity is also shown in Figure 4. This is primarily generated by absorption in the AlGaAsSb region and the relatively large photocurrent even at low voltages demonstrates that there is no barrier for electrons. Since, this is a two-dimensional model, the photocurrent is calculated in A/µm. For a three-dimensional device with a 200 µm diameter the currents will be about two orders of magnitude higher.

6. EXPERIMENTAL ACTIVITIES

MBE Material Growth

The UML effort includes development of Molecular Beam Epitaxy (MBE) growth of the materials discussed above. A Riber 3200 MBE system equipped with As and Sb cracker cells is employed for the growth of AlGaAsSb and InGaAsSb quaternary alloys lattice matched to GaSb. Substrate temperatures of 500°C are used with Ga/Sb flux ratios between 1–7, and fixed Al/As and In/As flux ratios. In-situ reflection high-energy electron diffraction (RHEED) is used to measure growth rates, alloy composition, and layer thicknesses (on wafers where RHEED oscillations can be observed). In the case of AlGaAsSb, growth rates for AlGaSb and GaSb are used to determine the initial Al composition, while Vegard's law is used to estimate the As composition in GaSb lattice matched quaternary alloy [4]. For InGaAsSb, the growth rates of GaSb and InAs are used for initial compositional determination [11].

X-ray rocking measurements performed on one of the AlGaAsSb layers grown on a GaSb substrate showed a substrate and alloy peak separation with in 120 arcsec suggesting good material quality with an Al composition of 12%. Cross-sectional transmission microscopy (TEM) images show no grading at GaSb/AlGaAsSb interfaces. Quaternary alloys with a Ga composition of about 75% were grown on semi-insulating GaAs substrates for doping studies indicated high p-type background doping. This problem is currently under investigation. With proper growth conditions one is expected to be able to reduce background doping levels to $5 \times 10^{15}\,\mathrm{cm}^{-3}$ for InGaAsSb and $2 \times 10^{16}\,\mathrm{cm}^{-3}$ for AlGaAsSb material systems. GaTe and Be sources will be used as n- and p-type dopants for the proposed APD device structures.

Measurement of Impact Ionization Coefficients

In order to measure the impact ionization coefficients, specialized APDs must be fabricated which allow for pure injection of electrons and holes into the electric field region in the same device. PIN diodes have been fabricated in $Al_{0.55}Ga_{.45}As_{.02}Sb_{.98}$ with a 1 µm thick I-region. The devices were defined using conventional photolithography and etched using a citric acid-based solution. Benzocyclobutene was used to passivate the devices and Ti/Au was used for both the p-type and n-type contacts. Initial attempts at ionization rate measurements were made using the structure shown below. Electrons were generated in the p-region using a 543 µm HeNe laser. Holes were generated in the GaSb layer using a 1.55 µm laser. Problems with this structure occurred due to a very small hole-induced photocurrent which preventing accurate measurement of the ionization coefficients. The ATLAS model was able to predict that this problem was caused by a barrier at the GaSb/AlGaAsSb heterojunction interface that hindered the diffusion of photo-induced holes from the GaSb layer into the electric field region. Efforts are now aimed at fabricating thinned AlGaAsSb homojunction APDs (with a removed GaSb substrate) that can be illuminated from both sides to achieve pure electron and pure hole injection.

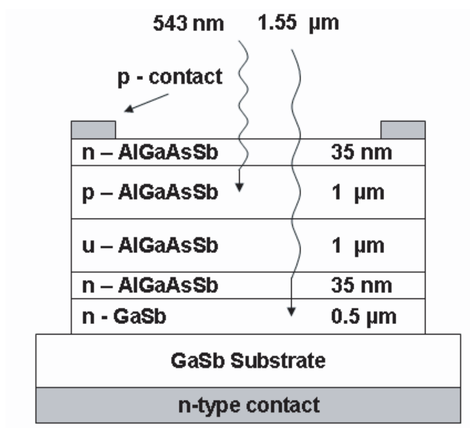


Figure 5: Initial APD Design used for measurement of impact ionization rates (top illumination version).

7. CONCLUSIONS

A basic model of a Long Wavelength APD Device based on quaternary alloys of GaSb has been demonstrated. Larger bandgap $Al_{0.45}Ga_{0.55}As_{0.05}Sb_{0.95}$ is used as multiplication region whereas the small band gap $In_{0.15}Ga_{0.85}As_{0.17}Sb_{0.83}$ alloy is used as the absorber for operation in the 2 µm wavelength range. These initial efforts demonstrate that a capable model of APDs in this material system can be developed. With improved information on material parameters, it is expected that optimized APDs for operation in the 2–5 µm wavelength range can be developed.

ACKNOWLEDGMENT

This work was supported by Air Force Office of Scientific Research (Dr. Pomrenke).

REFERENCES

1. Stillman, G. E., L. W. Cook, G. E. Bulman, N. Tabatabaie, R. Chin, and P. D. Dapkus, “Long wavelength (1.3 to 1.6 µm) detectors for fiber-optical communications,” *IEEE Transactions on Electron Devices*, Vol. ED-29, No. 9, September 1982.
2. Dutta, P. S., H. L. Bhat, and V. Kumar, “The physics and technology of gallium antimonide: An emerging optoelectronic material,” *Journal of Applied Physics*, Vol. 81, Issue 9, 5821–5870, May 1, 1997.

3. Wang, C. A., "Progress and continuing challenges in GaSb-based III-V alloys and heterostructures grown by organometallic vapor-phase epitaxy," *Journal of Crystal Growth (J. Cryst. Growth)*, ISSN 0022-0248 , 2004.
4. Adachi, S., "Band gaps and refractive indices of AlGaAsSb, GaInAsSb, and InPAsSb: Key properties for a variety of the 2–4 μm optoelectronic device applications," *Journal of Applied Physics*, Vol. 61, Issue 10, 4869–4876, May 15, 1987.
5. Vurgaftman, I., J. R. Meyer, and L. R. Ram-Mohan, "Band parameters for III-V compound semiconductors and their alloys," *Journal of Applied Physics*, Vol. 89, Issue 11, 5815–5875, June 1, 2001.
6. Sulima, O. V., M. G. Mauk, Z. A. Shellenbarger, J. A. Cox, J. V. Li, P. E. Sims, S. Datta, and S. B. Rafol, "Uncooled low-voltage AlGaAsSb/InGaAsSb/GaSb avalanche photodetectors," *IEE Proceedings-Optoelectronics*, Vol. 151, Feb. 2004.
7. Afrailov, M. A. and M. Ozer, "Electrical and photoelectrical properties of isotype N+-GaSb/n-GaInAsSb/N+-GaAlAsSb double heterojunctions," *Phys. Stat. Sol. (C)*, Vol. 2 No. 4, 1393–1398, March 2005.
8. Law, H. D., L. R. Tomassaetta, and J. S. Harris, "Ionization coefficients of $Ga_{0.72}Al_{0.28}Sb$ avalanche photodetectors," *Appl. Phys. Lett.*, December 1978.
9. Hildebrand, O., W. Kuebart, and M. H. Pilkuhn, "Resonant enhancement of impact in $Ga_{1-x}Al_xSb$," *Applied Physics Letters*, Vol. 37, Issue 9, November 1, 1980.
10. Andreev, I. A., M. P. Mikhailova, S. V. Mel'nikov, "Avalanche multiplication and ionization coefficients of GalnAsSb," *Soviet Physics Semiconductors*, 1991.
11. Reddy, M. H. M., J. T. Olesberg, C. Cao, and J. P. Prineas, "MBE-grown high-efficiency GaInAsSb mid-infrared detectors operating under back illumination," *Semicond. Sci. Technol.*, Vol. 21, March 2006.

High-quality Ultra-uniform Quantum Dot (QD) Fabrication Techniques for High-performance Terahertz Quantum Cascaded Laser

Xuejun Lu
Department of Electrical and Computer Engineering, University of Massachusetts Lowell
One University Avenue, Lowell, MA 01854, USA

Abstract— A high-quality quantum dot (QD) fabrication technique with precisely-controllable physical (size, position, and density) and electronics properties (energy levels, emission spectrum) is proposed. The fabrication technology employs *in situ* or *ex situ* etching of after-growing multi-quantum-well QCL structures followed by the regrowth of properly p-doped AlGaAs large band-gap materials. The heterostructure formed at the sidewall interface together with build-in electric field allows us to achieve nanometer-scale carrier confinement in the etched QDs and thus reduces the non-radiative recombination rate. In this paper, the fabrication method for high-quality QDs is presented. The photoluminescence (PL) enhancement by the nanometer-scale carrier confinement technology is calculated. Expected performance enhancement is simulated.

1. INTRODUCTION

Room temperature portable THz sources are of great importance in numerous defense, homeland security as well as civilian applications such as screening explosives at airports, imaging biological tissue for medical diagnostics, chemical detections for environment monitoring as well as astronomy imaging [1–4]. The state-of-the-art solid-state THz source is THz quantum cascade lasers (QCL) technique [5–8], which utilizing intersubband transitions in engineered multiple-quantum-well (MQW) active region for laser action [5]. The major limitation of THz QCL lasers lies in its inability to achieve CW operation at room temperature. This is primarily due to the nonradiative relaxation caused by the thermally-activated electron-LO phonon scattering process [8]. The LO-scattering nonradiatively depopulates electrons from the upper states to the lower states at a much faster (2300 times) rate than the radiative emission processes [5, 9]. Such as fast nonradiatively relaxing rate leads to poor emission efficiency and generate device heating problem.

The nonradiative LO-scattering rate can be greatly reduced by converting the QW active region into quantum-dot arrays [9–11]. By quantizing energy levels in an all three dimensions in quantum dots, we can achieve substantially reduced LO-phonon nonradiative relaxation due to the discrete energy levels, which are off-resonance with the LO-phonons with an energy of $\hbar\omega_{LO} \approx 36\,\text{meV}$ in GaAs. This is typically referred to as "phonon bottleneck" effect [12–14]. However, conventional self-assembly QD growth technique via Stanski-Krastanow (SK) growth mode [15] failed to achieve this due to the unavoidable size distribution in QD ensembles as is evidenced by inhomogeneously broadened photoluminescence (PL). The typical full-width-half-magnitude (FWHM) of PL of self-assembly QDs ranges from 30 meV to 50 meV, a value much larger than THz emission ($\sim 10\,\text{meV}$) [16, 17]. Such a large energy band distribution makes it impossible to achieve THz QCL based on self-assembly QDs. In this paper, we propose a high-quality QD fabrication process using an innovative nanometer scale carrier confinement technology. The nanometer scale carrier confinement is achieved by using the build-in electric field and the heterostructure formed at the sidewall interface. Such nanometer-scale carrier confinement is capable of reducing the non-radiative recombination rate at the interface of the etched QDs and thus enhancing the radiative emission efficiency. This paper discusses the fabrication method for high-quality QDs. The photoluminescence (PL) enhancement by the nanometer-scale carrier confinement technology is calculated. Expected performance enhancement is simulated.

2. HIGH-QUALITY QD FABRICATION METHOD

The schematic structure of the QD THz QCL is shown in Fig. 1(a). It is formed by converting the QW active region of a QC laser into the QD array. The simplified QD structure and the schematic for carrier confinement are shown in Fig. 1(b). Fig. 1(c) shows the schematic for carrier confinement profile.

The QD can be fabricated by *in situ* or *ex situ* etching of after-growing multi-quantum-well QCL structures followed by the regrowth of properly p-doped AlGaAs large band-gap materials. The

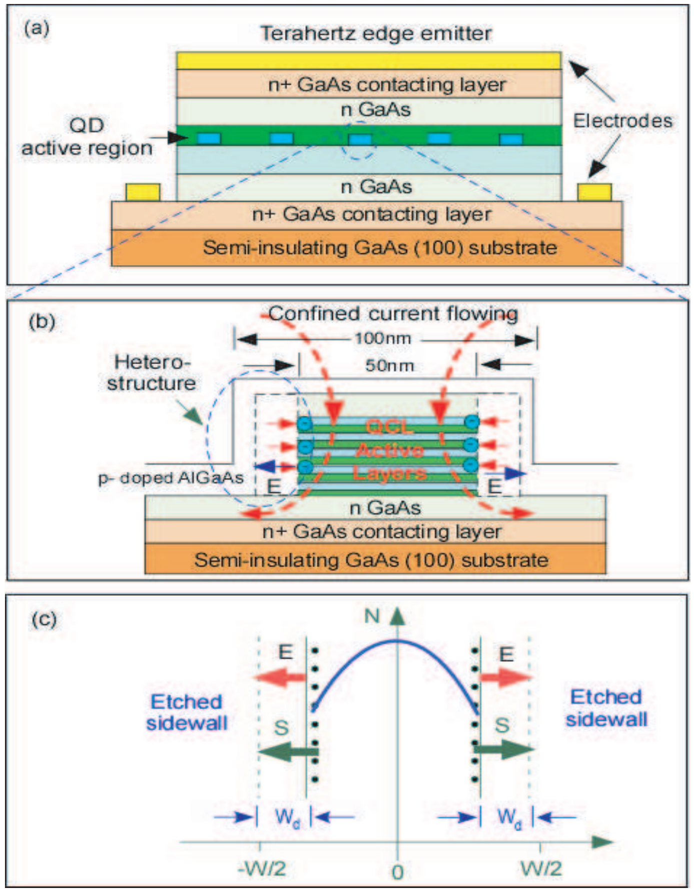


Figure 1: (a) Schematic structure of the QD THz QCL; (b) Simplified structure of the QD; (c) Carrier profile inside the QD.

p-doped AlGaAs large band-gap materials would form a heterostructure at the sidewall interface with a build-in electric field pointing to the edge (Figs. 1(b), (c)). This electric field will drive the negative charged electrons away from the etched edge. The hetero-structure at the sidewall interface helps to confine the carriers in the desired region. Such carrier confinement would push carriers away from the non-radiative recombination center at the sidewall interfaces and thus reduces the non-radiative recombination rate there.

Since the carrier confinement is achieved by the built-in electric field in the doping-level-dependent depletion region, much smaller confinement region could be generated than the actual size of the etched dots. This would enable us to achieve nanometer-scale carrier confinement not limited by the *in situ* or *ex situ* etching resolution capability. In addition, the ultra-small carrier confinement would generate high local current density and reduce the leakage current for highly efficient emission.

3. MODELING OF THE SURFACE RECOMBINATION PROCESS AND CALCULATION OF PHOTOLUMINESCENCE (PL) ENHANCEMENT

Follow [21], the schematic of modeling of the surface recombination process were shown in Fig. 1(c). The carriers in the depletion region experience two opposite processes: (1) diffusion due to the carrier density gradient; (2) drift due to the build-in electric filed in the depletion region. The net carrier flow rate R to the etched surface can be written as:

$$\vec{R} = D\frac{\partial N}{\partial x} - N\mu_{drift}\vec{E}, \tag{1}$$

where, N, and D are the carrier density and carrier diffusion coefficient, respectively. μ_{drift}, and $\vec{E}$ are the carrier drift mobility and electric filed inside the depletion layer, respectively. The normalized surface recombination velocity is [18]:

$$\frac{S_C}{S_{NC}} = \frac{\dfrac{D}{N}\dfrac{\partial N}{\partial x} - \mu_{drift}\vec{E}}{\dfrac{D}{N}\dfrac{\partial N}{\partial x}} = 1 - \frac{\mu_{drift}}{D}\frac{\vec{E}}{\dfrac{1}{N}\dfrac{\partial N}{\partial x}} = 1 - \frac{q}{K_B T}\frac{\vec{E}}{\dfrac{1}{N}\dfrac{\partial N}{\partial x}}, \tag{2}$$

where, S_C and S_{NC} are the surface recombination velocities with and without carrier confinement, respectively. K_B is Boltzmann's constant and $\mu_{drift}/D = q/K_B T$ is the famous Einstein relation in kinetic theory.

The rate equation for photoluminescence (PL) can be written as [18]:

$$\frac{\partial N}{\partial t} = G - \left(\frac{N}{\tau_R} + \frac{N}{\tau_{NR}} + \frac{NS}{(W - 2W_d)}\right), \tag{3}$$

where, G is the electron-hole generation rate per unit volume by the pump beam, and N is the carrier density. The radiative emission rate can thus be obtained by setting $\partial N/\partial t = 0$, i.e., steady-state:

$$R_{sp} = \frac{N}{\tau_R} = \frac{G}{\dfrac{\tau_R}{\tau_R} + \dfrac{\tau_R}{\tau_{NR}} + \dfrac{\tau_R S}{(W - 2W_d)}} = \frac{G}{1 + \dfrac{\tau_R}{\tau_{NR}} + \dfrac{\tau_R S}{(W - 2W_d)}}, \tag{4}$$

$$= \frac{G\tau/\tau_R}{1 + \dfrac{\tau S}{W - 2W_d}}, \tag{5}$$

The radiative emission rate determines the PL intensity. Fig. 2 shows the relative PL intensity as a function of the QD width for different $S\tau_{QW}$ with and without carrier confinement. Fig. 2 clearly shows the PL intensity enhancement by the carrier confinement.

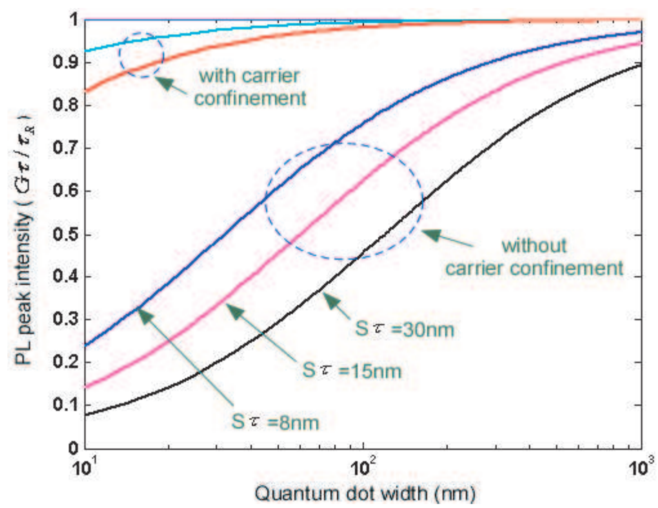


Figure 2: PL intensity as a function of the QD width for different $S\tau_{QW}$ with and without carrier confinement.

4. SIMULATION OF DEVICE PERFORMANCE ENHANCEMENT

The carrier rate equation in the confined region can be expressed as [18]:

$$\frac{\partial N}{\partial t} = -\frac{N}{\tau} + G - \frac{N}{(W - 2W_d)/4S}, \tag{6}$$

where, G is the carrier generation rate due to the current injection.The factor 4 comes from the total 4 edges of the QDs.The carrier lifetime τ in the confined region can thus be expressed as:

$$\frac{1}{\tau} = \frac{1}{\tau_R} + \frac{1}{\tau_{NR}} + \frac{S}{W - 2W_d}, \tag{7}$$

where, τ_R, and τ_{NR} are the radiative lifetime and nonradiative lifetime, respectively.

The spontaneous emission efficiency in etched QDs $\eta_{spon,QD}$ can be written as:

$$\eta_{spon,QD} = \frac{1/\tau_R}{1/\tau} = \frac{1/\tau_R}{1/\tau_R + 1/\tau_{NR} + 4S/(W - 2W_d)}, \tag{8}$$

The corresponding spontaneous emission efficiency in quantum well, $\eta_{spon,QD}$ can be written as:

$$\eta_{spon,QW} = \frac{1/\tau_R}{1/\tau} = \frac{1/\tau_R}{1/\tau_R + 1/\tau_{NR}}, \tag{9}$$

Thus, the normalized spontaneous emission efficiency, i.e., the spontaneous emission efficiency degradation due to the surface recombination can be expressed as:

$$\frac{\eta_{spon,QD}}{\eta_{spon,QW}} = \frac{1/\tau_R + 1/\tau_{NR}}{1/\tau_R + 1/\tau_{NR} + 2S/(W - 2W_d)} = \frac{1}{1 + 4S\tau_{QW}/(W - 2W_d)}, \tag{10}$$

where, τ_{QW} is the carrier lifetime in QWs. Fig. 3 shows the normalized spontaneous emission efficiency as a function of QD width for different $S\tau_{QW}$ with and without carrier confinement.

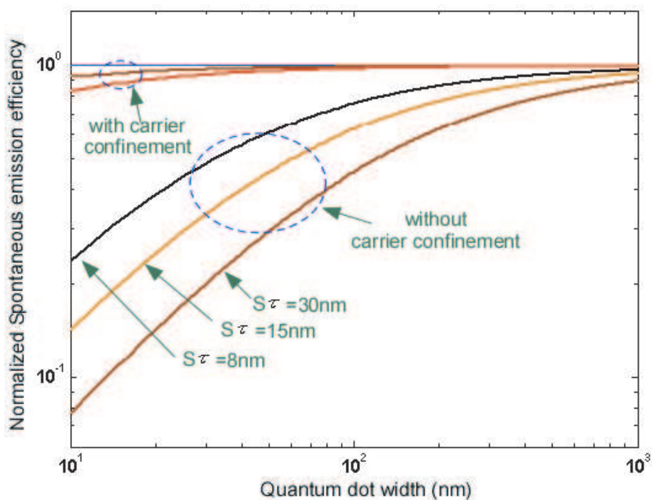


Figure 3: Normalized spontaneous emission efficiency as a function QD width for different $S\tau_{QW}$.

As show in Fig. 3, the surface recombination is severe for small QD dimension. With the cattier confinement technology, the spontaneous emission efficiency penalty can be made minimum. As indicated in Eq. (11), the carrier lifetime τ in etched QDs can be expressed using carrier lifetime in QWsτ_{QW} and recombination velocity S.

$$\frac{1}{\tau} = \frac{1}{\tau_R} + \frac{1}{\tau_{NR}} + \frac{S}{(W - 2W_d)} = \frac{1}{\tau_{QW}} + \frac{S}{(W - 2W_d)}, \tag{11}$$

The threshold current density J_{th} enhancement due to the surface recombination can be written as:

$$J_{th} = \frac{edN_w N_{th}}{\rho\tau} = \frac{edN_w N_{th}}{\rho}\left(\frac{1}{\tau_{QW}} + \frac{S}{(W - 2W_d)}\right) = \frac{J_{th,QW}}{\rho}\left(1 + \frac{S\tau_{QW}}{(W - 2W_d)}\right), \tag{12}$$

where, ρ, and N_{th} are the filling factor and the threshold carrier density, respectively. $\rho = W/\Lambda$, Λ is the period of the QDs, i.e., separation of two adjacent dots. With the carrier confinement, the filling factor ρ is expected to be unity. Fig. 4 shows the threshold current density enhancement factor as a function of QD width for different $S\tau_{QW}$ with and without carrier confinement. Fig. 4 clearly shows that with the current confinement the threshold current density is around one order of magnitude lower than that of without the current confinement case.

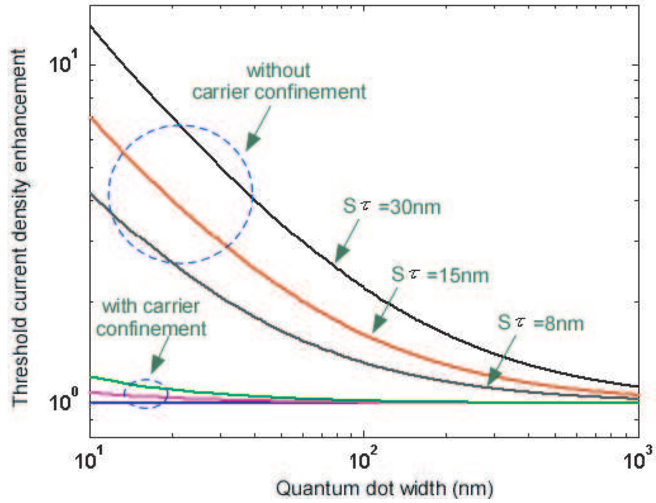


Figure 4: Threshold current density enhancement factor due to the surface recombination as a function QD width for different $S\tau_Q W$.

5. CONCLUSION

We propose a high-quality quantum dot (QD) fabrication technique with precisely-controllable physical (size, position, and density) and electronics properties (energy levels, emission spectrum). The fabrication technology employs nanometer-scale carrier confinement in the etched QDs. The nanometer-scale carrier confinement provided by the build-in electric field and the heterostructure formed at the sidewall interface allows us to reduce the non-radiative recombination rate and thus achieve enhanced radiative emission rate. A factor of 10 threshold current reduction is predicted.

REFERENCES

1. Miles, R. E., P. Harrison, and D. Lippens (eds), "Terahertz sources and systems," *NATO Science Series II*, Vol. 27, Kluwer, Dordrecht, 2001.
2. Han, P. Y., G. C, Cho, and X.-C. Zhang, "Time-domain transillumination of biological tissue with terahertz pulses," *Opt. Lett.*, Vol. 25, 242–244, 2000.
3. Mittleman, D. M., R. H. Jacobsen, and M. C. Nuss, "T-ray imaging," *IEEE J. Sel. Top. Quant. Electron.*, Vol. 2, 679–692, 1996.
4. Mittleman, D. M., S. Hunsche, L. Boivin, and M. C. Nuss, "T-ray tomography," *Opt. Lett.*, Vol. 22, 904–906, 1997.
5. Faist, J., F. Capasso, D. L. Sivco, C. Sirtori, A. L. Hutchinson, and A. Y. Cho, "Quantum cascade laser," *Science*, Vol. 264, 553–556, 1994.
6. Köhler, R., A. Tredicucci, F. Beltram, H. E. Beere, E. H. Linfield, A. G. Davies, D. A. Ritchie, R. C. Iotti, and F. Rossi, *Nature*, Vol. 417, 156, 2002.
7. Rochat, M., L. Ajili, H. Willenberg, J. Faist, H. Beere, G. Davies, E. Linfield, and D. Ritchie, *Appl. Phys. Lett.*, Vol. 81, 1381, 2002.
8. Williams, B. S., H. Callebaut, S. Kumar, Q. Hu, and J. L. Reno, "3.4-THz quantum cascade laser based on longitudinal-optical-phonon scattering for depopulation," *Appl. Phys. Lett.*, Vol. 82, 1015, 2003.
9. Hsu, C.-F., J.-S. O, P. Zory, D. Botez, "Intersubband quantum-box semiconductor lasers," *IEEE J. Select. Top. Quantum Electron*, Vol. 6, 491, 2000.
10. Wingreen, N. S. and C. A. Stafford, "Quantum-dot cascaded laser: Proposal for an ultra-low-threshold semiconductor laser," *IEEE J. Sel. Top. Quant. Electron*, Vol. 33, 1170–1173, 1997.
11. Botez, D., P. Zory, and C.-F. Hsu, "Intersubband quantum box semiconductor laser," U.S. Patent 5 953 356, Sept. 14, 1999.
12. Bockelmann, U. and G. Bastard, "Phonon scattering and energy relaxation in two-, one-, and zero-dimensional electron gases," *Phys. Rev. B*, Vol. 42, 8947–8951, 1990.
13. Benisty, H., C. M. S. Torres, and C. Weisbuch, "Intrinsic mechanism for the poor luminescence properties of quantum-box systems," *Phys. Rev. B*, Vol. 44, 10945, 1991.

14. Inoshita, T. and H. Sakaki, *Phys. Rev. B*, Vol. 46, 7260, 1992.
15. Arakawa, Y. and H. Sakaki, *Appl. Phys. Lett.*, Vol. 40, 939, 1982.
16. Huffaker, D. L., G. Park, Z. Zou, O. B. Shchekin, and D. G. Deppe, "1.3 μm roomtemperature GaAs-based quantum-dot laser," *Appl. Phys. Lett.*, Vol. 73, No. 18, 2564–2566, Nov. 1998.
17. Park, G., D. L. Huffaker, Z. Zou, O. B. Shchekin, and D. G. Deppe, "Temperature dependence of lasing characteristics for long-wavelength (1.3μm) GaAs-based quantum-dot lasers," *IEEE Photon. Technol. Lett.*, Vol. 11, No. 3, 301–303, Mar. 1999.
18. Yagi, H., "GaInAsP/InP multiple-quantum-wire lasers using strain-compensated quantum-well structure," PhD dissertation, Tokyo Institute of Technology, 2004.

Applications of Modular RBF/MLP Neural Networks in the Modeling of Microstrip Photonic Bandgap Structures

M. G. Passos[1], **H. C. C. Fernandes**[1], **and P. H. da F. Silva**[2]

[1]Federal University of Rio Grande do Norte, Brazil
[2]Federal Center of Technological Education, Brazil

Abstract— This paper presents a Radial Basis Function/Multilayer Perceptron (RBF/MLP) modular neural network, training with the Resilient Backpropagation (Rprop) algorithm which has been used for nonlinear device modeling in microwave band. The proposed modular configuration employs three or more neural networks, each one with a hidden layer of neurons, and aim to take advantage of the MLP and RBF networks specific characteristics to improve learning aspects, such as: ability to learn, speed of training and learning with consistency, or generalization. Simulations through the proposed neural network models for microstrip line with anisotropic PBG (Photonic Bandgap) structure and a metallic enclosure microstrip with PBG gave responses in good agreement with accurate results (measured or simulated) available in the literature.

1. INTRODUCTION

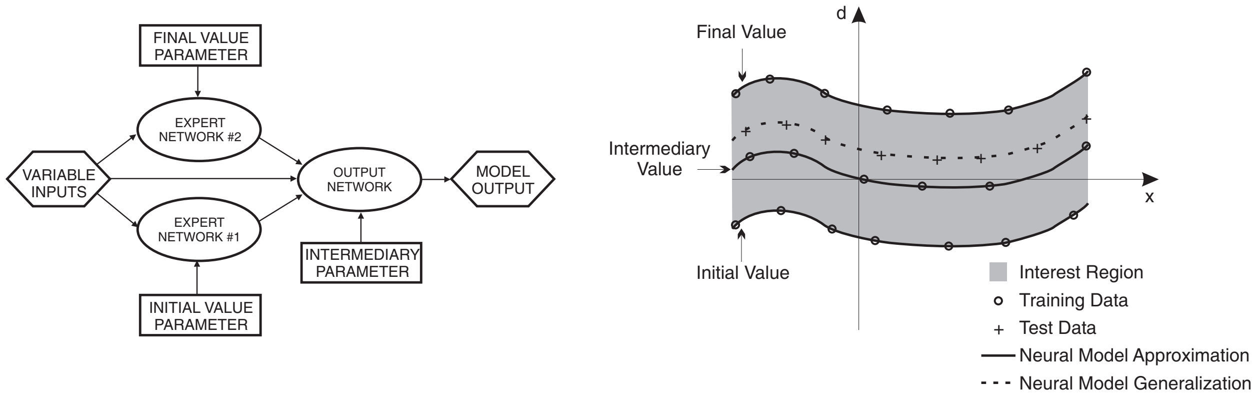
Since the beginning of the 1990s, the artificial neural networks have been used as a flexible numerical tool, which are efficient for the RF/microwave device/circuit modeling. The neural models, which are trained by means of precise data (obtained through measurements or by electromagnetic simulation), are used in the design/optimization phase of devices and circuits, supplying fast and accurate responses. In the CAD (Computer Aided-Design) applications related to microwave engineering and optical systems, the use of ANNs as nonlinear models becomes very common, [1]. Some hybrid modeling techniques have been proposed for the use with empirical models and neural networks, such as: Source Difference Method, [2], PKI (Prior Knowledge Input), [3], KBNN (Knowledge Based Neural Network), [4] and SM-ANN (Space Mapping Artificial Neural Network), [5]. A disadvantage in the hybrid models useage is the need of an empirical model. When this becomes a limitation, for example, when a new component does not have an empirical model or an equivalent circuit, the EM-ANN (Electromagnetic-Artificial Neural Network), [1] conventional technique, is commonly utilized. In this case, a simple neural network, MLP or RBF, is trained directly through electromagnetic data, which represent the behavior of the component under analysis.

In this paper the RBF/MLP modular structure is composed by combination of two expert RBF networks and one output MLP network and that configuration has been used for modeling devices in microwave/optical bands [6, 7]. The development of models through the RBF/MLP modular structure is described in Section 2. The applications of these neural models for microstrip line with anisotropic PBG (Photonic Bandgap) structure and a metallic enclosure microstrip with PBG are presented in Section 3. Section 4 gathers the conclusions of this research.

2. METHODOLOGY THROUGH THE RBF/MLP MODULAR STRUCTURE

The proposed modular structure uses three feed forward neural networks, each one with a hidden neuron layer: two expert RBF networks and an output MLP network. Figure 1(a) illustrates a diagram in RBF/MLP modular structure blocks. This choice was motivated by the individual characteristics of the MLP and RBF networks when used in the function approximation learning tasks: the RBF network performs a local approach, serving as an expert network, since it grasps the models' nonlinearities; the MLP network performs a global approach and acts as an output network, since it favours the generalization capacity of the RBF/MLP modular structure. In this technique, the model input parameters, named by 'initial value' and 'intermediary value', are related to the interested region defined by the training data, Figure 1(b). In order to receive additional information supplied by the pre-trained expert RBF networks, the output MLP network has two extra inputs, as shown in Figure 1(a).

The modeling problem mentioned is established by means of a normalized set of measured/simulated data, cited by $S = [x(n), d(n)]$, where, $1 \leq n \leq N$, and N is the total number of examples in the S training dataset. The $\mathbf{x}$ vector gathers the parameters of the model input (for instance, the



(a) (b)

Figure 1: (a) The proposed modular network configuration, (b) interest region defined by the training dataset.

geometrical parameters of microstrip line filters and frequency). The d desired response describes the device EM/physics behavior under consideration (for instance, the frequency response and scattering parameters). The EM/physics theoretical relation between $\mathbf{x}$ and d is given by,

$$d = f(\mathbf{x}) \tag{1}$$

where, f represents the input-output relation, which can be multidimensional and highly nonlinear. The aim is to develop a fast and accurate neural model for the f relation. The neural model is defined through the relation,

$$y = y(\mathbf{x}, \mathbf{w}) \tag{2}$$

where, $\mathbf{w}$ represents the free parameters (or weights) of the neural network structure.

The use of the RBF/MLP modular structure enables the division of a modeling problem in smaller and easier problems to be solved. To describe this division, the interested region is taken into account defined through the training data for a hypothetical device, Figure 1(b). The data referred to the 'initial value' and the 'initial value' parameters are used in the training of #1 and #2 expert RBF networks, respectively; the training of the MLP output network is done with all the training dataset, including the 'intermediary values' available.

In the MLP and RBF network supervised training with the backpropagation algorithm [8], the adjustment of the weights is carried out through the steepest descent method,

$$\mathbf{w}(t) = \mathbf{w}(t-1) - \eta \nabla E(\mathbf{w}(t-1)) \tag{3}$$

where, ∇ is the gradient operator; η is a training parameter, called learning rate, which controls the adjustments applied to the weights; and E is the square error of instantaneous error between the desired response and the neural network output. The training is carried out until the mean square error $E(t)$ reaches a minimum pre-established value. The $E(t)$ is a parameter that measures the training algorithm performance, and is defined by,

$$E(t) = \frac{1}{N} \sum_{n=1}^{N} E(n) \tag{4}$$

where, t is an index for the number of training epochs. An epoch is counted when all the training examples are presented to the neural network.

Due to the fact that backpropagation learning is too slow for many modeling applications, in this work the use of the Rprop algorithm (using the standard training parameters) is chosen [9]. The Rprop algorithm, proposed by Riedmiller and Braun [9], belongs to the algorithm family derived from backpropagation, which satisfies Jacobs' heuristics for the training acceleration, [10]. In the ANN training using the Rprop, just the gradient signs of the error function, Eq. (3), are taken into account. The negative inuence elimination of the gradient amplitudes in the Eq. (3) is eliminated, as well as the use of adaptive and individual learning rates for each ANN free parameter, awards convergence speed and robustness as regards the choice of the training parameters of Rprop algorithm, [9].

3. NEURAL NETWORK MODELING APPLICATIONS

3.1. Microstrip Line with Anisotropic Planar PBG

The anisotropic uniplanar PBG is a microstrip structure with a ground plane consisting of an array of etched slots of alternating widths, as shown in Figure 2(a). Its pattern is a two-dimensional (2-D) square-lattice periodic structure with a unit cell geometry exhibiting a 180° symmetry. When the line is in the z-direction, the induced current can low freely through the structure and the signal is transmitted, while the signal is rejected when the line is in the y-direction because of the stepped-impedance slots breaking the continuity of metal paths [11].

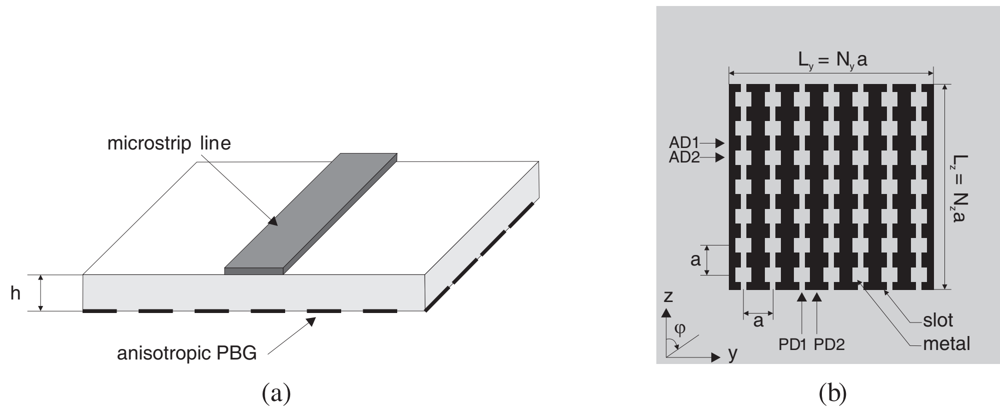


Figure 2: (a) Microstrip line with PBG anisotropic structure, (b) ground plane of the anisotropic PBG structure with parameters.

For the results presented in this paper, the period is $a = 1.524\,\text{mm}$ and the substrate is RT/Duroid with $h = 0.635\,\text{mm}$ and $\epsilon_r = 10.2$. The total dimensions of the PBG are, therefore, $L_y \times L_z = a(N_y \times N_z) = 1.524(N_y \times N_z)\,\text{mm}$, where N_y and N_z are the number of unit cells along y and z, respectively, as shown in Figure 2(b). In this particular case, N_z is fixed in 7, N_y vary 1 to 7 and line is placed in AD direction.

In the neural model training for the anisotropic planar PBG, two scaled input parameters were taken into consideration: the operation frequency, f, and the number of cell on y axis, N_y. The measured values in the scattering parameter S_{21} make up the desired responses for the neural models. The training data were obtained through measurements presented in [11]. The information related to the RBF/MLP modular network training is presented in Table 1.

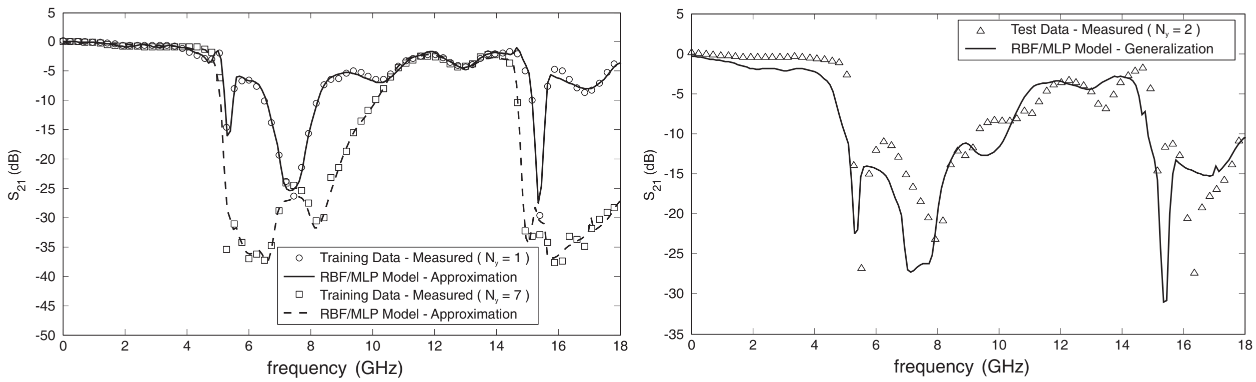
Table 1: Information related to the RBF/MLP modular training for the anisotropic planar PBG.

Neural Network	Expert 1-RBF	Expert 2-RBF	Output-MLP
Input parameter:	$N_y = 1$	$N_y = 7$	$N_y = [1\,7]$
# hidden neurons:	15	25	25
# training examples:	84	84	168
final $E(t)$:	5.86E-005	7.42E-004	1.62E-004
# training epochs:	2000	1500	2000

Figure 3(a) presents the results obtained from simulation of the RBF/MLP modular model developed. A good agreement between this model's responses and the measured data is verified, with excellent interpolation to the training dataset examples. To take account of this highly nonlinear learning task, a reasonable generalization result was obtained around what permits to predict the filter stop-band approaches to microstrip anisotropic PBG, as shown in Figure 3(b).

3.2. Metallic Enclosure Microstrip Line with PBG

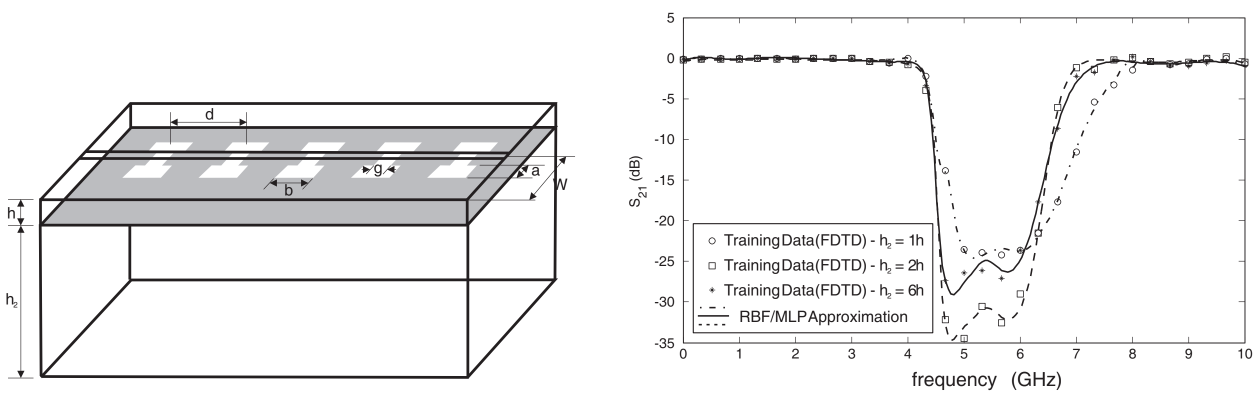
To observe the influence of a metallic enclosure on the S-parameters of a PBG structure with holes etched in the ground plane, the PBG circuit proposed in [12], is analyzed as the basic PBG structure in current section. In this case, a PBG circuit with five unit cells is fabricated on TACONIC CER-10 (dielectric constant ϵ_r of 10) substrate with the thickness $h = 1.5748\,\text{mm}$, as shown in Figure 4.



(a) (b)

Figure 3: (a) RBF/MLP model approximation for $N_y = 1$ and $N_y = 7$, (b) RBF/MLP model generalization for $N_y = 2$.

Other lattice dimensions $a = b = 2.5\,\text{mm}$, $g = 0.2\,\text{mm}$, and a period $d = 5\,\text{mm}$. The gap, whose length is the same as the width of the strip line, is just located under the strip line in the metallic ground plane. A line width of 1.46 mm is used, corresponding to $50\,\Omega$ line for a conventional microstrip line, [13]. The aim of this section is to analyze and to model the influence in terms of the distance between the lower wall of the metallic enclosure and the PBG structure.



(a) (b)

Figure 4: (a) 3-D view of the shielded PBG structure, (b) RBF/MLP model approximation for $h_2 = 1h$, $2h$ and $6h$.

Table 2: Information related to the RBF/MLP modular training for the enclosure microstrip line with PBG.

Neural Network	Expert 1-RBF	Expert 2-RBF	Output-MLP
Input parameter:	$h_2 = 1h$	$h_2 = 6h$	$h_2 =[1h\,2h\,6h]$
# hidden neurons:	10	15	10
# training examples:	31	31	93
final $E(t)$:	1.67E-004	1.64E-004	9.02E-005
# training epochs:	3000	1000	5000

In the neural model training for the enclosure microstrip line, two scaled input parameters were taken into consideration: the operation frequency, f, and the distance between the lower wall and the PBG ground plane, h_2. The simulated values in the scattering parameter S_{21} through finite-difference time-domain (FDTD) method make up the desired responses for the neural models. In the model development for the enclosure microstrip line, the RBFs and MLP modular structure networks were trained separately. The relevant information for the training is in Table 2.

Figure 4(b) presents the results obtained from simulation of the RBF/MLP modular model developed. A good agreement between this model's responses and the simulated data is verified, with excellent interpolation to the training dataset examples. To take account of this difficult nonlinear modeling learning task, a good generalization result is obtained in all considered frequency range, as shown in Figure 5.

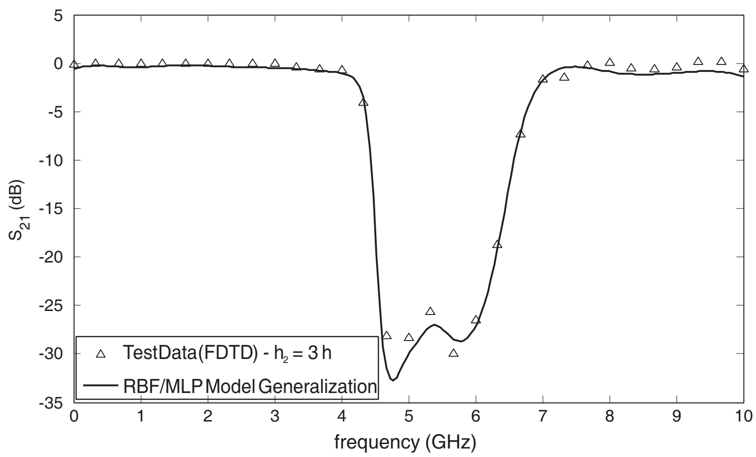


Figure 5: RBF/MLP model generalization for $h_2 = 3h$.

4. CONCLUSIONS

In this paper a modular structure of neural networks, named RBF/MLP trained with the efficient Rprop algorithm, is developed for specific for utilization in stop-band filters modeling. In particular, a microstrip line with anisotropic PBG and enclosure microstrip line with PBG are used in modeling simulations.

The RBF/MLP structure's modules are organized in order to take advantage of the local and global characteristics presented by the RBF and MLP neural networks, respectively, when used in the function approximation learning tasks. This kind of organization, together with the modeling problem division, becomes easier in the training of individuals RBF and MLP networks in the RBF/MLP modular structure. The obtained neural models simulation results indicate a good learning consistency, or generalization, and a major reliability of the models developed through the RBF/MLP modular structure. Besides, the RBF/MLP structure, directly trained by means of measured/simulated data through the EM-ANN technique, becomes very flexible, and it still can be applied as models, mainly when new components/technologies for microwaves circuits arise.

ACKNOWLEDGMENT

This work was supported in part by CNPq and CAPES.

REFERENCES

1. Zhang, Q. J. and K. C. Gupta, *Neural Networks for RF and Microwave Design*, 1st ed., Norwood, MA, Artech House Inc., 2000.
2. Watson, P. M. and K. C. Gupta, "EM-ANN models for microstrip vias and interconnects in multilayer circuits," *IEEE Trans. Microwave Theory and Techniques*, Vol. MTT-44, 2495–2503, 1996.
3. Watson, P. M., K. C. Gupta, and R. L. Mahajan, "Development of knowledge based artificial neural networks models for microwave components," *IEEE MTT-S Int. Microwave Symp. Dig.*, 9–12, 1998.
4. Wang, F. and Q. J. Zang, "Knowledge based neuromodels for microwave design," *IEEE Trans. Microwave Theory and Techniques*, Vol. MTT-45, 2333–2343, 1997.
5. Rayas-Sanchez, J. E., *Neural Space Mapping Methods for Modeling and Design of Microwave Circuits*, Ph.D. thesis, McMaster University, 2001.
6. Passos, G. M., P. H. da F. Silva, and H. C. C. Fernandes, "A RBF/MLP modular neural network for microwave device modeling," *International Journal of Computer Science and Network Security*, Vol. 6, No. 5A, 81–86, 2006.

7. Passos, G. M., P. H. da F. Silva, and H. C. C. Fernandes, "A New RBF/MLP modular neural network for device modeling in microwave/optical bands," *Proceedings of XXVII Iberian Latin-American Congress on Computational Methods in Engineering*, 08–513, 2006.
8. Haykin, S., *Neural Networks: A Comprehensive Foundation*, Macmillan, 1999.
9. Riedmiller, M. and H. Braun, "A direct adaptive method for faster backpropagation learning: The Rprop algorithm," *Proceedings of IEEE Int. Conf. Neural Networks*, Vol. 1, 586–591, 1993.
10. Jacobs, R. A., "Increase rate of convergence through learning rate adaptation," *Neural Networks*, Vol. 1, 295–307, 1998.
11. Caloz, C. and T. Itoh, "Multilayer and anisotropic planar compact PBG structures for microstrip applications," *IEEE Trans. Microwave Theory and Techniques*, Vol. 50, No. 9, 2206–2208, 2002.
12. Kim, C. S., J. S. Park, D. Ahn, and J. B. Lim, "A novel 1-D periodic defected ground structure for planar circuits," *IEEE Microwave Guided Wave Lett.*, Vol. 10, 131–133, 2000.
13. Zhengwei, D., K. Gong, J. S. Fu, B. Gao and Z. Feng, "Influence of a metallic enclosure on the S-parameters of microstrip photonic bandgap structures," *IEEE Trans. Electromagnetic Compatibility*, Vol. 44, No. 2, 324–328, 2002.

Mobile Broadband: The Emergency of IEEE 802.16e

Humberto César Chaves Fernandes and Davi Bibiano Brito

Department of Electrical Engineering, Universidade Federal do Rio Grande do Norte — UFRN
Natal/RN 59.078-970, P. O. Box 1583, Brazil

Abstract— Since the ratification of the IEEE 802.16 standard in 2001 the telecommunication world is witnessing the emergency of wireless broadband services. Now with the publication of IEEE standard 802.16e in 2005, the mobile version of 802.16, new wireless systems will enable better communication, enhanced productivity and better customer service allowing users to access information beyond their desk moving from anywhere. This paper analyses the new possibilities offered by this new standard from various aspects.

1. INTRODUCTION

The IEEE 802.16 and the so called WiMAX "Wireless (Wi) Microwave Access (MA)" technology involves microwaves for the transfer of high speed data wireless at distances up to many kilometers. WiMAX is very similar to Wi-Fi in that it uses the same core technology of wireless modulation developed way back in the 1960's and 1970's the Orthogonal Frequency Division Multiplexing (OFDM).

While clearly based on the same OFDM base technology adopted in 802.16, the 802.16e version is designed to deliver service across many more sub-channels than the OFDM 256-FFT both standards support single carrier, OFDM 256-FFT and at least OFDMA 1K-FFT. But the 802.16e standard adds OFDMA 2K-FFT, 512-FFT and 128-FFT capability. Sub-canalization facilitates access at varying distance by providing operators the capability of dynamically reducing the number of channels while increasing the gain of signal to each channel in order to reach customers to distances farther as 50 Km.

Promoters of 802.16 elected to form an organization to test and certify products for interoperability and standards compliance. That organization is known as the Worldwide Interoperability for Microwave Access (WiMAX), aiming to provide broadband wireless access (BWA) on the scale of the Metropolitan Area Network (MAN). As this technology is a fairly recent phenomenon many aspects must be discussed to determine the viability of a wireless broadband network.

IEEE 802.16e is optimized for dynamic mobile radio channels and supports hard and soft handoffs to provide users with seamless connections as they move across coverage areas of adjacent cells and roaming at vehicular speeds of up to 100 km/h. The deployment of IEEE 802.16e is at frequency range of 2 GHz to 6 GHz, which is low enough to accommodate the non-line-of-sight conditions between the base stations and mobile devices. Mobility support in 802.16e includes power-saving and sleep modes to extend the battery life of mobile devices.

The other improvements for mobile devices include a real-time polling service to provide QoS and the transport protocols supported including TCP/IP. It will offer broadband connectivity similar to Wi-Fi to use a notebook in a hotspot or in a building, the user will be able to move around at pedestrian speeds and maintain your broadband connection. The 802.16e standard will most likely be utilized primarily in licensed spectrum for pure mobile applications promising to deliver high data rates over large areas to a large number of users in the near future.

2. ADAPTIVE ARRAY ANTENNAS

To increase the capacity and coverage in broadband data communication according to the IEEE 802.16e WiMAX standard, there are essentially two types of adaptive array antennas (or smart antennas): the switched-beam antenna and the beam-steering type. The switched-beam antennas can combine only the beams from the different element in a finite number of juxtapositions. The beam can assume a few fixed widths and a few fixed angles and is not infinitely variable. The beam-steering type is, on the other hand, infinitely variable and is far more flexible. It also requires far more processing power to operate effectively.

The greater part of the transmitted signal power is radiated into directions other than toward the specific user. This causes interference, reducing efficiency and the range of coverage. Especially in new wireless broadband services as 802.16e, where the user front-end is very simple, it becomes necessary to provide every user with a specific beam offering enough gain to increase the range so the smart antennas can provide this in a number of non-line-of-sight (NLOS) applications.

The antenna array has the ability to focus a beam very tightly toward each subscriber unit on a packet-by-packet basis. The RF energy in that beam is not dispersed through the atmosphere as in a normal broad beam transmission and instead is delivered almost in its entirety to the subscriber site, where it can blast through considerable obstructions and still provide a usable signal. By concentrating energy in such tight patterns on a channel-by-channel or even packetby-packet basis, the adaptive array can also increase frequency reuse, theoretically up to several times within a single cell, while still adhering to regulatory limits on transmitter power. This in turn reduces the need for sectorization.

All adaptive array antennas provide signal diversity; that is, the mere fact that several spaced antenna elements are exposed simultaneously to the transmission practically guarantees that signal quality will vary from element to element. The system can then select the best signal, and in an NLOS situation, there is a far higher likelihood that an array will find a usable signal than a single element. In some systems the adaptive-antenna array can even take multipath reflections impinging on the various elements and phase-align them so as to construct a single, coherent, high-strength signal.

Adaptive array antennas are most beneficial when used in both the base station and the subscriber terminal. Such double-ended systems are known as multiple-in, multiple-out (MIMO) links (the multiples refer to the antenna arrays). Other applications include the use of the smart antennas to provide not only signal to 802.16e, but also to Wireless Local Area Networks (WLAN) that are compatible with the IEEE 802.11 family. Fig. 1 shows the use of an adaptive base station to cover an area with different types of subscribers.

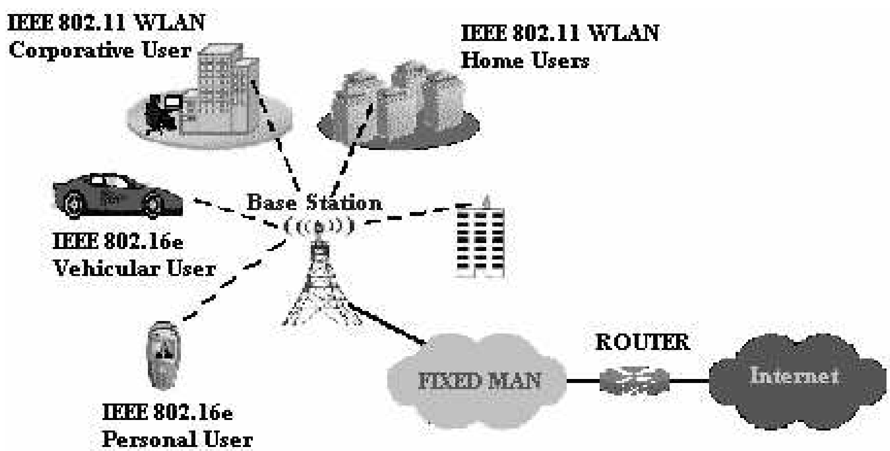


Figure 1: Smart base station providing signal to different subscribers.

3. APPLICATIONS

The 802.16e standard will provide solutions across multiple broadband segments. This technology was developed to become a last mile access solution to compete with technologies like the third cellular generation 3G and a new mobile standard known as High-Speed Downlink Packet Access (HSDPA).

It is a rapidly growing technology that is most viable for backhauling the rapidly increasing volumes of traffic being generated by Wi-Fi hotspots. WiMAX is a MAN technology that fits between wireless LANs, such as 802.11, and wireless wide-area networks (WANs), such as the cellular networks. 802.16e was developed to provide low-cost, high-quality, flexible, mobility BWA using certified, compatible and interoperable equipments.

3.1. Mobile Voice

While VoIP (voice over IP) has been around for years, it has not been a viable alternative for most wireless applications. Recent technology advancements have tremendously improved quality and now VoIP service providers are positioned to offer an affordable alternative to traditional circuit-switched voice services.

Wireless mobile telephony is a simple and cost-effective service which allows a subscriber to use VoIP services while on the move. This is possible because of WiMAX which can provide carrier-grade connectivity while being wireless. Such services could enhance the attractiveness of the broadband wireless network to some customers especially if a single smart phone, could be made to serve as both a phone and a high-speed access device.

The 802.16e will also provide much higher throughputs for data, purportedly in the megabits per second, at least in burst mode, and the smart phones will allow the user not only to talk but to have a high speed access anywhere in the area covered by the network, and ultimately fourth-generation cellular and broadband wireless may not fill different niches either.

Mobile voice is a desirable service offering for broadband wireless in that it may be something to entice users onto the network, it may also be a loss leader. Mobile voice is already a price-eroded business, and the arrival of new service providers and one-number portability cannot improve that situation.

3.2. Video Applications

Video and audio conferencing are well within the capabilities of 802.16e systems. Videoconferencing, while never experiencing the explosive growth predicted for it in the past, must be considered a successful application and is increasingly utilized by enterprises for training purposes and for controlling travel expenses and to smart phones with high definition video cameras. Furthermore, it has always been a high-value application for which businesses have been willing to pay premium prices.

Video-on-demand is another application that is one of the most hyped technologies which never took off. Now with WiMAX has been found a technology which can make the base wider and price points better suited. WiMAX can reach the masses at low cost, and hence more people who need services. Another interesting feature is that alternative videos and contents related to learning, training, can become a revenue-generating mechanism due to obvious financial value. The mobile user will be able to watch high definition videos and make video conferences on the same device with flexibility and mobility offered by the wireless connection.

4. CONCLUSIONS

The ability to offer distinct services with mobility is the only sure way to differentiate the IEEE 802.16e standard from other technologies. In the future, new offerings will emerge such as location-based services, highly personalized content and reporting services, advanced filtering, and others that have scarcely even been envisioned. Exploring the complete range of services that broadband will support in the midterm and determining their impact is quite impossible in a book of this sort, and I can say only that tracking developments in this area is essential to the long-term health of the broadband wireless service provider. Above all, network operators must come to view themselves as marketing agents and not as public utilities. Public utilities are products of the era of regulated monopoly. In the new era of competition, they are anachronisms.

REFERENCES

1. Pareek, D., *The Business of WiMAX*, John Wiley & Sons Ltd., The Atrium, Southern Gate, Chichester, West Susexx, POO19 8SQ, England, 2006.
2. Wang, F., A. Ghosht, R. Love, K. Stewart, R. Ratasuk, R. Bachu, Q. Zhao, and Y. Sun, *Journal Title Personal, Indoor and Mobile Radio Communications, 2005. PIMRC 2005. IEEE 16th International Symposium on Abbreviation*, Vol. 2, 900–904, 2005.
3. Mahler, W. and F. M. Landstorfer, "Design and optimisation of an antenna array for WiMAX base stations," *IEEE/ACES International Conference on Wireless Communications and Applied Computational Electromagnetics*, 1006–1009, Honolulu, USA, April 2005.
4. Abichar, Z., Y. Peng, and J. M. Chang, "WiMax: the emergence of wireless broadband," *IT Professional*, Vol. 8, No. 4, 40–48, 2006.
5. Ma, L. and D. Jia, "The competition and cooperation of WiMAX, WLAN and 3G," *2005 2nd International Conference on Mobile Technology, Applications and Systems*, 1–5, Munich, Germany, November 2005.

Rectangular Slot Resonator with Four Dielectric Layers

Humberto César Chaves Fernandes, **Humberto Dionísio de Andrade**
Davi Bibiano Brito, **and Manoel Bonfim Lins de Aquino**
Department of Electrical Engineering, Technological Center, Federal University of Rio Grande do Norte
UFRN, P. O. Box: 1583, 59, 078-970 Natal, RN, Brazil

Abstract— This work shows the development and mathematics analysis to the rectangular slot line resonator with four dielectric overlapped layers, using the full wave Transversal Transmission Line LTT-method. Deriving from the Maxwell's equations, we can get a set of equations that represent the electromagnetic fields. After several numerical manipulations, we can get the complex resonant frequency through double spectral variables. The numerical results are shown graphically.

1. INTRODUCTION

The rectangular slot line resonator with four layers, consists of one rectangular slot line resonator, where there are two layers under and two layers over it. This structure is shown in the Fig. 1, with width w and length l. The general equations for the electromagnetic fields are obtained by the method TTL, with adequate basis function and Garlekin's procedure. The complex resonant frequency is calculated using double spectral variables, being the same, used in the elaboration of the efficiency and bandwidth's parameters.

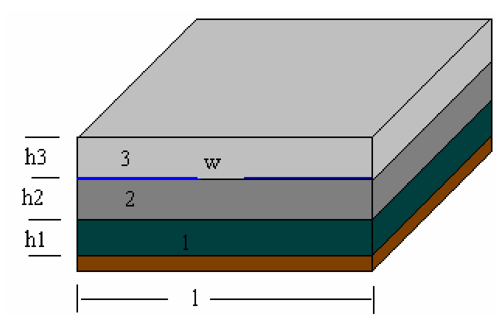


Figure 1: Spatial view of the four layers slot line resonator.

With the aid of the Cartesian coordinates system, the dimensional nomenclatures and electromagnetic as presented in Fig. 1(a) (spatial view) and Fig. 1(b) (traverse section of the structure), all are obtained referred them equations of fields, being considered despicable the thickness of the slot line.

2. FIELDS IN STRUCTURE

Due to the limitation of the length, the equations should be used for the analysis in the spectral domain in "x" and "z" directions as function. Therefore the field equations are applied for double Fourier transformed and defined as:

$$\widetilde{f}(\alpha_n, y, \beta_k) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y, z) \cdot e^{j\alpha_n x} \cdot e^{j\beta_k z} dx dz \tag{1}$$

where α_n is the spectral variable in the "x" direction and β spectral variable in the "z" direction.

After using the Maxwell's equations in the spectral domain, the general equations of the electric and magnetic fields to the method TTL, are obtained:

$$\widetilde{E}_{xi} = \frac{1}{\gamma_i^2+k_i^2}\left[-j\alpha_n\frac{\partial}{\partial y}\widetilde{E}_{yi}+\omega\mu\beta_k\widetilde{H}_{yi}\right] \quad (2.1)$$

$$\widetilde{E}_{zi} = \frac{1}{\gamma_i^2+k_i^2}\left[-j\beta_k\frac{\partial}{\partial y}\widetilde{E}_{yi}-\omega\mu\alpha_n\widetilde{H}_{yi}\right] \quad (2.2)$$

$$\widetilde{H}_{xi} = \frac{1}{\gamma_i^2+k_i^2}\left[-j\alpha_n\frac{\partial}{\partial y}\widetilde{H}_{yi}-\omega\varepsilon\beta_k\widetilde{E}_{yi}\right] \quad (2.3)$$

$$\widetilde{H}_{zi} = \frac{1}{\gamma_i^2+k_i^2}\left[-j\beta_k\frac{\partial}{\partial y}\widetilde{H}_{yi}+\omega\varepsilon\alpha_n\widetilde{E}_{yi}\right] \quad (2.4)$$

where: $i = 1, 2, 3, 4$ represent four dielectrics regions of the structure;

$$\gamma_i^2 = \alpha_n^2+\beta_k^2-k_i^2 \quad (2.5)$$

is the constant of the propagation in y direction; α_n is the spectral variable in "x" direction and β_k the spectral variable in "z" direction.

$k_i^2 = \omega^2\mu\varepsilon = k_0^2\varepsilon_{ri}^*$ is the number of wave of ith term of Dielectric region;

$\varepsilon_{ri}^* = \varepsilon_{ri} - j\frac{\sigma_i}{\omega\varepsilon_0}$ is the dielectric constant relative of the material with losses;

$\omega = \omega_r + j\omega_i$ is the complex angular frequency;

$\varepsilon_i = \varepsilon_{ri}^* \cdot \varepsilon_0$ is the dielectric constant of the material;

The equations above are applied to the resonator, the fields E_y and H_y are gotten from the solution of the Helmoltz's wave equations in the spectral domain [2–4]:

$$\left(\frac{\partial^2}{\partial y^2}-\gamma^2\right)\widetilde{E}_y = 0 \quad (3.1)$$

$$\left(\frac{\partial^2}{\partial y^2}-\gamma^2\right)\widetilde{H}_y = 0 \quad (3.2)$$

The solutions of Helmoltz's equations for the four regions of the structure are given, for examples:

Region 2:

$$\widetilde{E}_{y2} = A_{2e}\cdot\sinh\gamma_2 y + B_{2e}\cdot\cosh\gamma_2 y \quad (4.1)$$

$$\widetilde{H}_{y2} = A_{2h}\cdot\sinh\gamma_2 y + B_{2h}\cdot\cosh\gamma_2 y \quad (4.2)$$

Region 4:

$$\widetilde{E}_{y3} = A_{3e}\cdot e^{-\gamma_3 y} \quad (4.3)$$

$$\widetilde{E}_{y3} = A_{3h}\cdot e^{-\gamma_3 y} \quad (4.4)$$

Substituting these solutions in the equations of the fields (2.1) to (2.4), in function of the unknown constants A_{21}, A_{22}, B_{21} and B_{22} are obtained, for example, for the region 2:

$$\widetilde{E}_{x2} = \frac{-j}{K_2^2+\gamma_2^2}\left[(j\omega\mu_0\beta_k B_{21}+\alpha_n\gamma_2 A_{22})\cosh(\gamma_2 y)) + (j\omega\mu_0\beta_k B_{22}+\alpha_n\gamma_2 A_{21})\sinh(\gamma_2 y))\right] \quad (4.5)$$

$$\widetilde{H}_{x2} = \frac{-j}{K_2^2+\gamma_2^2}\left[(j\omega\varepsilon_2\beta_k A_{21}+\alpha_n\gamma_2 B_{22})\cosh(\gamma_2 y)) + (j\omega\varepsilon_2\beta_k A_{22}+\alpha_n\gamma_2 B_{21})\sinh(\gamma_2 y))\right] \quad (4.6)$$

For the determination of the unknown constants,the conditions of close contour to the regions 1, 2 and 3 are applied as follows,

Regions 1 e 2: $y = h_1$

$$\widetilde{E}_{x1} = \widetilde{E}_{x2} \quad (5.1)$$

$$\widetilde{E}_{z1} = \widetilde{E}_{z2} \quad (5.2)$$

$$\widetilde{H}_{x1} = \widetilde{H}_{x2} \quad (5.3)$$

$$\widetilde{H}_{z1} = \widetilde{H}_{z2} \quad (5.4)$$

Regions 2 e 3: $y = d$-; $(g = h_1 + h_2)$

$$\widetilde{E}_{x2} = \widetilde{E}_{x3} = \widetilde{E}_{xg} \tag{5.5}$$
$$\widetilde{E}_{z2} = \widetilde{E}_{z3} = \widetilde{E}_{zg} \tag{5.6}$$

After several calculations, for region two, we obtain:

$$A_{21} = \frac{\varepsilon_1 \cosh(\gamma_2 y)}{\varepsilon_2 \gamma_1 \sinh(\gamma_1 g_1)\cosh(\gamma_2 g_2) + \gamma_2 \frac{\varepsilon_1}{\varepsilon_2}\cosh(\gamma_1 g_1)\sinh(\gamma_2 g_2)} * \left[j(\alpha_n \tilde{E}_{xg} + \beta_k \tilde{E}_{zg}\right] \tag{6.1}$$
$$A_{22} = \frac{\gamma_1 \sinh(\gamma_2 y)}{\gamma_2 \gamma_1 \sinh(\gamma_1 g_1)\cosh(\gamma_2 g_2) + \gamma_2 \frac{\varepsilon_1}{\varepsilon_2}\cosh(\gamma_1 g_1)\sinh(\gamma_2 g_2)} * \left[j(\alpha_n \tilde{E}_{xg} + \beta_k \tilde{E}_{zg}\right] \tag{6.2}$$
$$B_{21} = -\frac{\sinh(\gamma_1 g_1)}{\omega\mu_0 \sinh(\gamma_1 g_1)\cosh(\gamma_2 g_2) + \frac{\gamma_1}{\gamma_2}\cosh(\gamma_1 g_1)\sinh(\gamma_2 g_2)} * \left[-\beta_k \tilde{E}_{xg} + \alpha_n \tilde{E}_{zg}\right] \tag{6.3}$$
$$B_{22} = -\frac{\gamma_1}{\gamma_2}\frac{\cosh(\gamma_1 g_1)}{\omega\mu_0 \sinh(\gamma_1 g_1)\cosh(\gamma_2 g_2) + \frac{\gamma_1}{\gamma_2}\cosh(\gamma_1 g_1)\sinh(\gamma_2 g_2)} * \left[-\beta_k \tilde{E}_{xg} + \alpha_n \tilde{E}_{zg}\right] \tag{6.4}$$

The general equations of the electromagnetic fields are obtained as function of tangential electric fields on the antenna resonator.

3. CALCULATION OF THE ADMITANCE MATRIX

The following Equations (7.1) and (7.2) show the relations of the current densities in the sheets ($\tilde{J}_{xt}$ and $\tilde{J}_{zt}$) and the magnetic fields in the interface $y = h_1 + h_2$:

$$\tilde{H}_{x2} - \tilde{H}_{x3} = \tilde{J}_{zt} \tag{7.1}$$
$$\tilde{H}_{z2} - \tilde{H}_{z3} = -\tilde{J}_{xt} \tag{7.2}$$

It has been done for the substitutions of the equations of the magnetic fields, so after some calculations we obtain,

$$Y_{xx}\widetilde{E}_{xg} + Y_{xz}\widetilde{E}_{zg} = \widetilde{J}_{zg} \tag{8.1}$$
$$Y_{zx}\widetilde{E}_{xg} + Y_{zz}\widetilde{E}_{zg} = \widetilde{J}_{xg} \tag{8.2}$$

This equation can be represented in the matrix form as follows,

$$\begin{bmatrix} Y_{xx} & Y_{xz} \\ Y_{zx} & Y_{zz} \end{bmatrix} \begin{bmatrix} \widetilde{E}_{xg} \\ \widetilde{E}_{zg} \end{bmatrix} = \begin{bmatrix} \widetilde{J}_{zg} \\ \widetilde{J}_{xg} \end{bmatrix} \tag{9}$$

The "Y" admittance terms are Green's functions dyadic to the antenna, where,

$$Y_{xx} = -\frac{j}{\varpi\mu_0(\gamma_2^2+k_2^2)}\left[-\beta_k^2\gamma_2 E + k_2^2\alpha_n^2 F\right] + \frac{j}{\varpi\mu_0\gamma_3}\left[\alpha_n^2 k_3^2 E - \beta_k^2\gamma_3^2 D\right] \tag{9.1}$$

$$Y_{xz} = \frac{-j\alpha_n\beta_k}{\varpi\mu_0(\gamma_2^2+k_2^2)}\left[A + k_2^2(B)\right] - \frac{\alpha_n\beta_k}{\varpi\mu_0\gamma_3(k_3^2+\gamma_3^2)}\left[k_3^2 C + \gamma_3^2 D\right] \tag{9.2}$$

$$Y_{zx} = \frac{-j\alpha_n\beta_k}{\varpi\mu_0(\gamma_2^2+k_2^2)}\left[A + k_2^2(B)\right] - \frac{\alpha_n\beta_k}{\varpi\mu_0\gamma_3(k_3^2+\gamma_3^2)}\left[k_3^2 C + \gamma_3^2 D\right] \tag{9.3}$$

$$Y_{zz} = \frac{j}{\varpi\mu_0(\gamma_2^2+k_2^2)}\left[\alpha_n^2 A - \beta_k^2 k_2^2 B\right] - \frac{j}{\varpi\mu_0\gamma_3(k_3^2+\gamma_3^2)}\left[\alpha_n^2\gamma_3^2 C - \beta_k^2\gamma_3^2 D\right] \tag{9.4}$$

$$A = \frac{\gamma_1\cdot\gamma_2}{\gamma_2\tanh(\gamma_1\cdot h_1)+\gamma_1\tanh(\gamma_2 h_2)} + \frac{\gamma_2^2}{\frac{\gamma_2}{\tanh(\gamma_2\cdot h_2)}+\frac{\gamma_1}{\tanh(\gamma_1\cdot h_1)}} \tag{9.5}$$

$$B = \left(\frac{\varepsilon_1}{\gamma_1\varepsilon_2\tanh(\gamma_1 h_1)+\gamma_2\varepsilon_1\tanh(\gamma_2 h_2)} + \frac{\gamma_1\varepsilon_2}{\frac{\gamma_1\gamma_2\varepsilon_2}{\tanh(\gamma_2 h_2)}+\frac{\gamma_2^2\varepsilon_1}{\tanh(\gamma_1\cdot h_1)}}\right) \tag{9.6}$$

$$C = \left(\frac{\frac{\gamma_3\varepsilon_4}{\gamma_4\varepsilon_3}+\tanh(\gamma_3 h_3)}{1+\frac{\gamma_3\varepsilon_4\tanh(\gamma_3 h_3)}{\gamma_4\varepsilon_3}}\right) \tag{9.7}$$

$$D = \left(\frac{\frac{\gamma_4}{\gamma_3}+\tanh(\gamma_3 h_3)}{1+\frac{\gamma_4\tanh(\gamma_3 h_3)}{\gamma_3}}\right) \tag{9.8}$$

$$E = \left(\frac{\gamma_1+\gamma_2\tanh(\gamma_1 h_1)\tanh(\gamma_2 h_2)}{\gamma_2\tanh(\gamma_1 h_1)+\gamma_2\tanh(\gamma_2 h_2)}\right) \tag{9.9}$$

The electric tangents fields in the interface are expanded by using base functions [3, 5] as follows:

$$\widetilde{E}_{xg} = \sum_{i=1}^{n} a_{xi}\cdot\widetilde{f}_{xi}(\alpha_n,\beta_k) \tag{10.1}$$

$$\widetilde{E}_{zg} = \sum_{j=1}^{m} a_{zj}\cdot\widetilde{f}_{zj}(\alpha_n,\beta_k) \tag{10.2}$$

where a_{xi} and a_{zj} are unknown constants and the terms n and m are positive integers that can be equal to 1, as in the Equations (12.1) and (12.2):

$$\widetilde{E}_{xg} = a_x\cdot\widetilde{f}_x(\alpha_n,\beta_k) \tag{11.1}$$

$$\widetilde{E}_{zg} = a_z\cdot\widetilde{f}_z(\alpha_n,\beta_k) \tag{11.2}$$

The base functions of the Fourier transformation are chosen as [6]:

$$\widetilde{f}_x(\alpha_n) = \pi\cdot J_o\left(\alpha_n\frac{w}{2}\right) \tag{12.1}$$

$$\widetilde{f}_x(\beta_k) = \frac{2\pi l\cdot\cos\left(\frac{\beta_k l}{2}\right)}{\pi^2-(\beta_k l)^2} \tag{12.2}$$

$$\widetilde{f}_x(\alpha_n,\beta_k) = \frac{2\pi^2 l\cdot\cos\left(\frac{\beta_k l}{2}\right)}{\pi^2-(\beta_k l)^2}\cdot J_0\left(\alpha_n\frac{w}{2}\right) \tag{12.3}$$

where J_0 is zero order Bessel's function of the first kind.

The Garlekin's method is applied to (9), to eliminate curent densities and new equation in matrix form is obtained [5, 7].

$$\begin{bmatrix} K_{xx} & K_{xz} \\ K_{zx} & K_{zz} \end{bmatrix}\cdot\begin{bmatrix} a_x \\ a_z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{13}$$

where,

$$K_{xx} = \sum_{-\infty}^{\infty} \widetilde{f}_x \cdot Y_{xx} \cdot \widetilde{f}_x^* \tag{14}$$

The solution to the characteristic equation of the determinant (14) supplies the resonant frequency.

4. RESULTS

According to the figure, the resonant frequency decreases when the width or length of the slot is decreased. Fig. 2 shows the curves of resonant frequency as function of length slot for different thicknesses of substrate at the first range. Fig. 3 shows the curves of resonant frequency in function of width slot for different thicknesses of substrate at the third range.

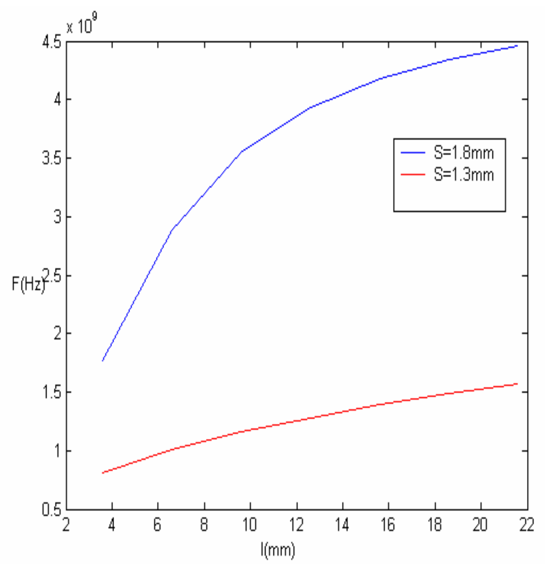


Figure 2: Frequency (GHz) as function of the length (mm).

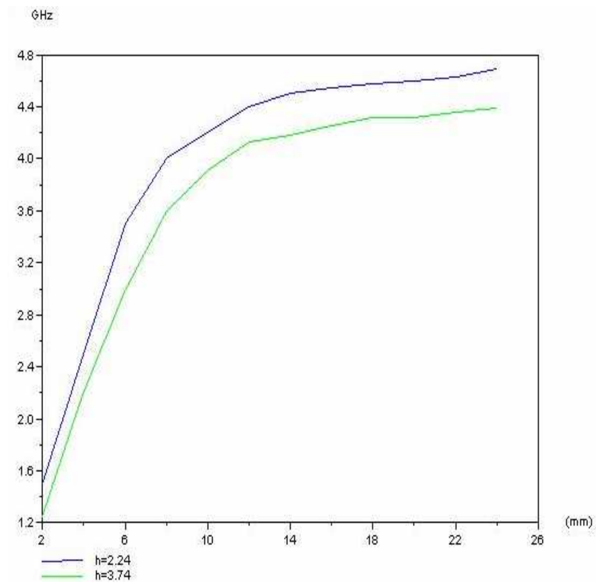


Figure 3: Frequency (GHz) as function of the width (mm).

5. CONCLUSION

The Transverse Transmission Line — TTL method was used in the analysis of obtaining the numerical results of the four layers slot line resonator. Using the concise and effective procedures the calculus of the complex resonant frequency was obtained with accuracy. The possibility of the alternate various materials is the greatest advantage of multiple layers slot line resonator that can be used as antenna.

REFERENCES

1. Agrawal, A. K. and B. Bhat, "Resonant characteristics and end effects of a slot resonator in unilateral fin line," *Proc. IEEE*, Vol. 72, 1416–1418, Oct. 1984.
2. Fernandes, H. C. C., S. A. P. Silva, and J. P. Silva, "Coupling analisys at the coupler and unilateral edge-coupled fin line," *International Conference on Millimeter and Submillimeter Waves and Applications II, SPIE's 1998 International Symposium on Optical Science, Engineering and Instrumentation, Conf. Proc.*, 53–54, San Diego, Califórnia, USA. July 1998.
3. Fernandes, H. C. C. and S. A. P. Silva, "Asymmetric and unilateral thick edge-coupled fin line and coupling analysis," guest, *PIERS 1999 — The Progress in Electromagnetics Research Symposium, Conf. Proc.*, Vol. 90, Taipei, Taiwan, China, Març 1999.
4. Fernandes, H. C. C., S. A. P. Silva, Costa, and O. S. Delfino, "3D complex propagation of coupled unilateral and antipodal arbitrary finlines," *Brazilian CBMAG'96-Congress of Electromagnetism*, 159–162, Ouro Preto-MG, Nov. 1996.

5. Baht, B. and S. K. Koul, *Analysis, Design and Applications of Finlines*, Artech House, 1987.
6. Silva, S. A. P. and H. C. C. Fernandes,"Functions of basis in analisys of the acoupled unilateral fin line coupler," *IV SPET-symposium of Research and Extension in Technology*, 79–81, Christmas-RN, Annals, Nov. 1998.
7. Fernandes, H. C. C., "Planar structures general in guides of waves Millimeters: finlines," Thesis of Doctorate, 189, FEC, UNICAMP, Campinas, SP, July 1984.

Multilayer Planar Resonators with Superconductive Patch on PBG Substrate

Humberto César Chaves Fernandes and George Dennes Fernandes Alves
Federal University of Rio Grande do Norte, DEE, P.O. Box 1583, Lagoa Nova, 59078-970-Natal, RN, Brazil

Abstract— The rectangular discretized multilayer microstrip resonators, analyzed using the full wave TTL — Transverse Transmission Line method in the spectral domain, jointly with the moment method are presented. The homogenization theory is used to calculate the equivalent permittivity for the s and p polarizations of the structure composed of PBG (*Photonic Band Gap*) material. Numerical results of the resonance frequency for the rectangular multilayer, on isotropic and anisotropic PBG material substrate resonators are presented.

1. INTRODUCTION

The microstrip antennas have received more attention due to the characteristics and advantages comparison with the conventional antennas. These antennas present small volume, reduced weight, planar configuration, compatibility with integrated hybrid circuits and the possibility of acting with dual frequency [1]. They can be used in several systems, such as: radars, wireless, mobile telephony and communication by satellites.

However, they present disadvantages with relationship to the narrow band, losses due to the low gain and losses by irradiation due to the surface waves. These disadvantages can be eliminated, by using a screening to avoid losses by irradiation [2] and antennas with PBG substrate (Photonic Bandgap) [3]. Some techniques to increase the band width are used among the: thicker substrate [4], multilayer antennas or stacked patches [5, 6].

Antennas with dielectric multilayer have advantages, such as, the flexibility in the operation frequency band, a smaller physical size, etc.

The PBG materials are a new type of periodic substrates structure in which the propagation in certain bands of frequencies is prohibited [7]. In this work the rectangular multilayer resonators on PBG substrate are analyzed, as illustrated in the Fig. 1.

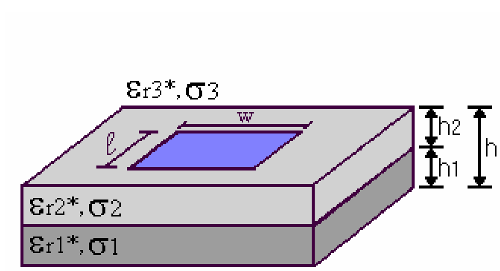


Figure 1: Rectangular multilayer microstrip resonator.

The TTL — Transverse Transmission Line method [8–10] is used in the determination of the electromagnetic fields components in the Fourier transform domain (FTD), for the three regions of the structures. The moment method is applied and adequate basis functions are used to expand the current densities in the metallic strip.

2. THEORY

The TTL method in the Fourier transform domain, uses a component of propagation in the y direction, treating the general equations of electric and magnetic field as functions of $\vec{E}_y$ and $\vec{H}_y$.

Starting from the Maxwell equations, and after several algebraic manipulations, the equations that represent the electromagnetic fields in the x and z directions are obtained as a function of the electromagnetic fields the in the y direction,

$$\left\{ \begin{array}{c} \vec{E}_{Ti} \\ \vec{H}_{Ti} \end{array} \right\} = \frac{1}{k_i^2 + \gamma_i^2} \left[j\omega \nabla_T \times \left\{ \begin{array}{c} -\mu \vec{H}_{yi} \\ \varepsilon_1 \vec{E}_{yi} \end{array} \right\} + \frac{\partial}{\partial y} \nabla_T \left\{ \begin{array}{c} \vec{E}_{yi} \\ \vec{H}_{yi} \end{array} \right\} \right] \quad (1)$$

After then, the two dimensional Fourier transforms are applied. The electromagnetic fields for i-th dielectric region are then obtained:

$$\widetilde{E}_{xi} = \frac{1}{\gamma_i^2 + k_i^2}\left[-j\alpha_n \frac{\partial}{\partial y}\widetilde{E}_{yi} + \omega\mu\beta_k \widetilde{H}_{yi}\right] \tag{2}$$

$$\widetilde{E}_{zi} = \frac{1}{\gamma_i^2 + k_i^2}\left[-j\beta_k \frac{\partial}{\partial y}\widetilde{E}_{yi} - \omega\mu\alpha_n \widetilde{H}_{yi}\right] \tag{3}$$

$$\widetilde{H}_{xi} = \frac{1}{\gamma_i^2 + k_i^2}\left[-j\alpha_n \frac{\partial}{\partial y}\widetilde{H}_{yi} - \omega\varepsilon\beta_k \widetilde{E}_{yi}\right] \tag{4}$$

$$\widetilde{H}_{zi} = \frac{1}{\gamma_i^2 + k_i^2}\left[-j\beta_k \frac{\partial}{\partial y}\widetilde{H}_{yi} - \omega\varepsilon\alpha_n \widetilde{E}_{yi}\right] \tag{5}$$

where $i = 1, 2, 3\ldots$ are the dielectric regions, $\gamma_i^2 = \alpha_n^2 + \beta_k^2 - k_i^2$ where γ_i is the propagation constant in $\boldsymbol{y}$ direction, α_n is the spectral variable in $\boldsymbol{x}$ direction, β_k is the spectral variable in $\boldsymbol{z}$ direction, $k_i^2 = \omega^2\mu\varepsilon = k_0^2\varepsilon_{ri}^*$ and $\varepsilon_{ri}^* = \varepsilon_{ri} - j\dfrac{\sigma_i}{\omega\varepsilon_0}$ is the relative dielectric permittivity of the material, k is the wave number, ω is the angular frequency, ε_0 is the dielectric permittivity in the free space, and σ_i is the conductivity of the i-th layer. The analytical development, using the LTT method, for the multilayer resonators studied in this work is given in [11].

The electromagnetic field and the currents are related with the dyadic Green matrix [Z], [11, 13],

$$\begin{bmatrix} Z_{xx} & Z_{xz} \\ Z_{zx} & Z_{zz} \end{bmatrix} \begin{bmatrix} \widetilde{J}_z \\ \widetilde{J}_x \end{bmatrix} = \begin{bmatrix} \widetilde{E}_x \\ \widetilde{E}_z \end{bmatrix} \tag{6}$$

Appling the Galerkin technique, the electric fields out of the metallic strip are eliminated. The current densities are expanded in appropriate basis functions and (6) becomes a homogeneous complex matrix as shown in (7).

$$\begin{bmatrix} K_{xx} & K_{xz} \\ K_{zx} & K_{zz} \end{bmatrix} \cdot \begin{bmatrix} a_x \\ a_z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{7}$$

Each element of the [K] characteristic matrix is shown in (8)–(9) for examples:

$$K_{xx} = \sum_{-\infty}^{\infty} \widetilde{f}_x(x,z) Z_{xx} \widetilde{f}_x^*(x,z) \tag{8}$$

$$K_{xz} = \sum_{-\infty}^{\infty} \widetilde{f}_z(x,z) Z_{xz} \widetilde{f}_x^*(x,z) \tag{9}$$

The no-trivial solution generates the characteristic equation, in which the roots allow the obtaining of the resonance frequency of the structure.

1) Rectangular Resonator

After obtaining (7), appropriate basis functions are used to approximate the values of the current densities to the form the real function, as presented in (10) and (11):

$$\widetilde{J}_{xh}(x,z) = \sum_{i=1}^{Ni} a_{xi} f_{xi}(x,z) \tag{10}$$

$$\widetilde{J}_{zh}(x,z) = \sum_{i=1}^{Ni} a_{zi} f_{zi}(x,z) \tag{11}$$

The basis functions are used in the characteristic matrix [K] for the expansion of the current densities [15].

3. NUMERICAL RESULTS

As a comparison, Fig. 2 presents the results obtained for the multilayer microstrip rectangular resonator with superconductive patch, on isotropic and PBG material, and the second comparison results of the microstrip rectangular resonator of one layer, without superconductive patch. Another comparison using the Cavity model [19] is also made.

For the simulation of the three-layer resonator, the two-layer under patch were considered as a single-layer, whose height of each layer was considered as 0.635 mm (see Fig. 1). In this way, the height of the added two-layers ($h_1 + h_2 = h$) will result in a height $h = 1.27$ mm. The substrate was RT Duroid 5880 of relative electric permittivities $\varepsilon_{r1} = 2.2$ and $\varepsilon_{r2} = 2.2$, the three layer was considered as being the air $\varepsilon_{r3} = 1.0$, the width of the patch is of $w = 15.0$ mm.

The resonator of one layer was simulated for $h = 1.27$ mm, with relative electric permittivities $\varepsilon_{r1} = 2.2$ and $\varepsilon_{r2} = 1.0$, width of the patch $w = 15.0$ mm. The Fig. 3 has shown the result of the variation of the resonance frequency in function of several length of the patch resonator.

Observing the Fig. 2, one notices that the dielectric three-layer resonator with superconductive patch works correctly when it simulates a resonator of one layer without superconductive patch.

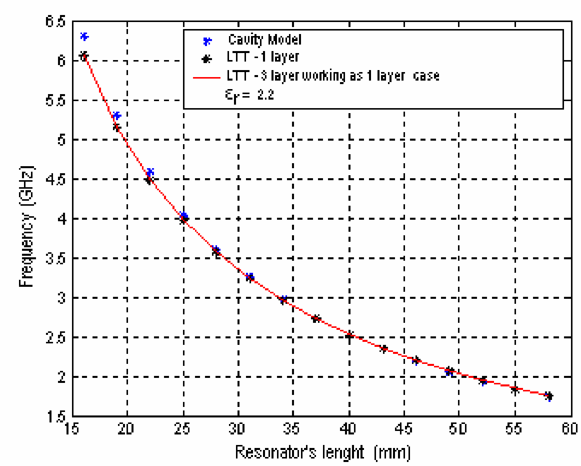


Figure 2: Resonance frequency as function of the length of the patch for $\varepsilon_r = 2.2$.

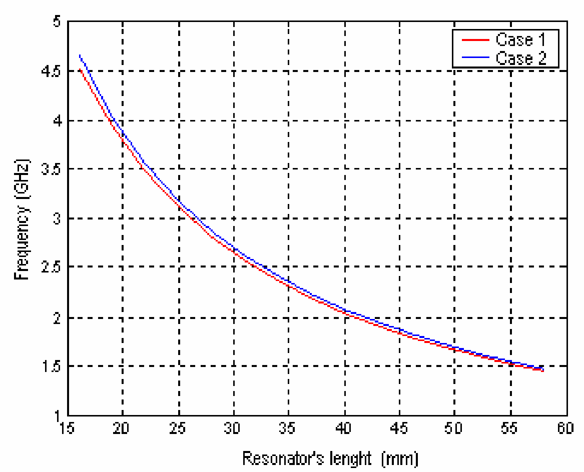


Figure 3: Resonance frequency as function of the length of the patch for the cases 1 and 2.

After the verification that the multilayer resonator is working correctly, simulations were accomplished with PBG 2D material substrate, where the effective permittivity of the dielectric depends on the polarization of the electric field $\boldsymbol{s}$ (parallel to the axis z) and $\boldsymbol{p}$ (perpendicular to the axis z). Being the equivalent permittivity of the structure in the substrate of Silicon (Si) for the $\boldsymbol{p}$ polarization $\varepsilon_r = 8.7209$ and for the $\boldsymbol{s}$ polarization $\varepsilon_r = 10.233$, the width of the patch is $w = 15.0$ mm.

The Fig. 3 has shown the curves for the simulated cases: Case 1: Dielectric layer 1 composed of RT Duroid 5880 with $\varepsilon_{r1} = 2.2$, dielectric layer 2 composed of PBG 2D material considering the incident wave to the $\boldsymbol{s}$ polarization and the dielectric layer 3 is the air with $\varepsilon_{r3} = 1.0$; Case 2: Dielectric layer 1 composed of RT Duroid 5880 with $\varepsilon_{r2} = 2.2$, dielectric layer 2 composed of PBG 2D material considering the incident wave to the $\boldsymbol{p}$ polarization and the dielectric layer 3 is the air with $\varepsilon_{r3} = 1.0$.

4. CONCLUSION

A theoretical study about the rectangular multilayer dielectric microstrip resonators was presented using the Transverse Transmission Line method in the spectral domain. The theory of the Homogenization was applied for the determination of the equivalent permittivity of the structure composed of PBG.

The rectangular geometry is made in agreement with the desired application, because the triangular geometry tends to occupy a smaller physical space than the rectangular, being able to be used in applications where the miniaturization of the antenna or resonator is necessary. The authors thank the Brazilian agencies CNPQ, CAPES and UFRN for supporting this work.

REFERENCES

1. Bahl, I. J. and P. Barthia, *Microstrip Antennas*, Artech House, 1982.
2. Polichronakis, I. and S. S. Kouris, "Computation of the dispersion characteristics of a shielded suspended substrate microstrip lines," *IEEE Trans. on Microwave Theory Tech.*, Vol. 40, No. 3, 581–584, Mar. 1992.
3. Almeida, J. F. and C. L. S. Sobrinho, "Position influence of one PBG structure on the microstrip antenna bandwidth," (in portuguese), *Dig. MOMAG 2004*, 5, CD, Ago. 2004.
4. Chen, W., K. F. Lee, and R. Q. Lee, "Input impedance of coaxially fed rectangular microstrip antenna on electrically thick substrate," *Microwave Opt. Tech. Lett.*, Vol. 5, No. 6, 387–390, May 1993.
5. Damiano, J. P., J. Bennegueouche, and A. Papiernik, "Study of multilayer microstrip antennas with radiating elements of various geometry," *IEE Proceedings*, Vol. 137, No. 3, Pt. H, Jun. 1990.
6. Hassani, H. R. and D. Mirshekar-Syahkal, "Study of electromagnetically coupled stacked rectangular patch antennas," *IEE Proc. Microw. Propag.*, Vol. 142, No. 1, Feb. 1995.
7. Oliveira, L. C. M., "Applications of PBG structures in planar microwave devices — lines and antennas — in dielectric and semiconductors substrates: Technology development and characterization," Master Thesis, UNICAMP, Campinas-SP, Brazil.
8. Fernandes, H. C. C. and I. S. Queiroz, "Overlay shielded superconducting microstrip lines on double semiconducting regions," *IEEE / Twentieth International Conference on Infrared and Millimeter Waves , Conf. Dig.*, 379–380, Orlando-Florida-USA, Dec. 1995.
9. Fernandes, H. C. C., E. A. M. Souza, and I. S. Queiroz, "Conductor thickness in unilateral finlines on semiconductor substrate," *International Journal of Infrared and Millimeter Waves*, Vol. 16, No. 1, USA, Jun. 1994.
10. Fernandes, H. C. C., R. M. R. P. Spinelli, A. J. Giarola, L. M. Mendonç and J. E. F. Madrigales, "A simplified method applied to various finlines structures including dispersion, asymmetry, loss, finite strip thickness, and discontinuity," *12th International Conference on Infrared and Millimeter Waves*, 161–162, Orlando-Florida-USA, Dec. 1987.
11. Santos, K. C., "Applicaton of the TTL method to the rectangular and triangular structures in multilayers PBG substrates for mobile communications," (in Portuguese), Master Thesis, PPGEE — Federal University of rio Grande do Norte, Prof. H. C. C. Fernandes, Natal-RN, Brazil, Dec. 2005.
12. Santos, K. C. and H. C. C. Fernandes, "Study of resonance frequency of the microstrip antenna with photonic material," *IV Meeting of the SBPMat*, Recife-PE, Brazil, Oct. 2005.
13. Queiroz Jr., I. S., "Superconducting microstrip analysis on multilayers wiht losses and finite thickness," (in Portuguese), Master Thesis, PPGEE — Federal University of rio Grande do Norte, Prof. H. C. C. Fernandes, Natal-RN, Brazil, Feb. 1996.
14. Bahl and S. K. Koul, *Analysis, Designs and Applications of Finlines*, Artech House, 1987.
15. Silva, S. A. P., "Ressoadores de linha de lamina de microfita retangular com passos e triangular discretizado," (in Portuguese), Master Thesis, PPGEE — Federal University of rio Grande do Norte, Prof. H. C. C. Fernandes, Natal-RN, Brazil, Dec. 1999.
16. Fernandes, H. C. C., "General planar structures in millimeter waveguides:finlines," Doctoral Thesis, FEC, UNICAMP, 189, Campinas-SP, Brazil, Jul. 1984.
17. Felbacq and G. Bouchitté, "Homogenization of a set parallel fibers," *Waves Random Media*, Vol. 7, 245–255, 1997.
18. Centeno, E. and D. Felbacq, "Rigorous vector diffraction of electromagnetic waves by bidimensional photonic crystals," *J. Opt. Soc. Am. A*, Vol. 17, No. 2, 320–327, February 2000.
19. Balanis, C. A., *Antenna Theory: Analysis and Design*, John Wiley & Sons, 1997.

Photonic Crystal at Millimeter Waves Applications

Humberto César Chaves Fernandes, **Joêmia Leilane G. Medeiros**
Isnaldo M. A. Júnior, **and Davi B. Brito**

Department of Electrical Engeneering — Federal University of Rio Grande do Norte
Natal/RN 59.078-970, P. O. Box 1583, Brasil

Abstract— The photonic band gap (PBG) crystals have been used as a perfectly reflecting substrate for many millimeter wave applications. In his work the fin line directional coupler with PBG substrate was analyzed using the TTL — Transverse Transmission Line — method. Comparing with the other full wave methods the TTL is efficient, making possible a significant algebraic simplification of the equations involved in the process. In order to analyze the structure the coupling were determined. Numerical results obtained for this finline coupler are presented.

1. INTRODUCTION

Photonic band gap crystals have emerged as a new class of periodic dielectric structures where propagation of electromagnetic waves is forbidden for all frequencies in the photonic band gap [1]. This material has a periodic arrangement of cylinders immersed in air with diameters and spacing of less than a wave length [2–4]. This substrate can improve the band width and eliminate the propagation of undesirable modes. Many integrated circuits for millimeter wave applications can be made using fin line techniques. This includes, beyond the fin line circuit, circuits inserted in metal and other standards circuits, mounted in the wave guides E-plane [5].

This letter demonstrates an application of the 2D layer-by-layer PBG crystal; an efficient unilateral fin line directional coupler. This type of coupler can be realized by the use of the natural coupling between the 2 slots symmetrically localized. The analysis is made using the TTL method and the coupling definitions. The Fig. 1 shows a project of the device.

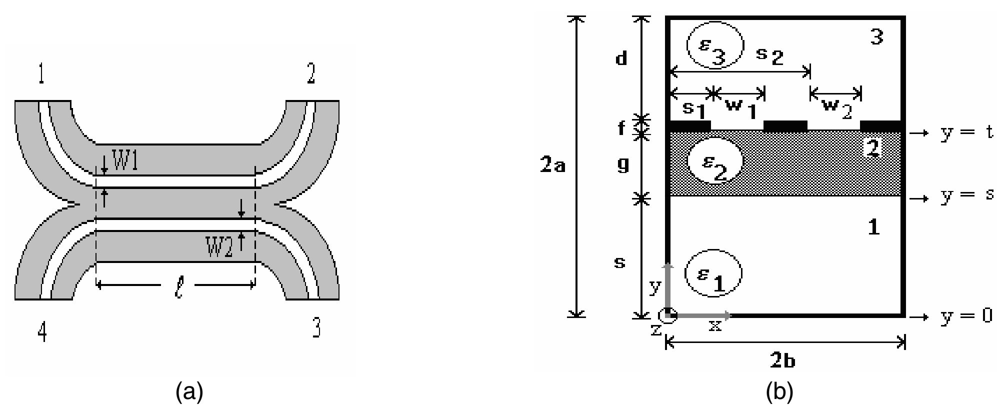


Figure 1: (a) Fin line coupler superior and intern view, and (b) transversal section of unilateral fin line coupler.

The coupled unilateral fin lines consist of a rectangular wave guide with three dielectric regions inside being the second region a substrate placed in the center of the wave guide and having three conductors fins on top of the substrate and the other two regions being air.

2. THEORY

Starting for the rotational Maxwell equations the electromagnetic fields are developed. The "x" and "z" components the final fields equations in the Fourier Transform Domain for the structures

ith regions are obtained:

$$\widetilde{E}_{xi} = \frac{1}{\gamma_i^2 + k_i^2}\left[-j\alpha_n \frac{\partial}{\partial y}\widetilde{E}_{yi} - j\omega\mu\Gamma\widetilde{H}_{yi}\right] \tag{1}$$

$$\widetilde{E}_{zi} = \frac{1}{\gamma_i^2 + k_i^2}\left[-\Gamma \frac{\partial}{\partial y}\widetilde{E}_{yi} - \omega\mu\alpha_n\widetilde{H}_{yi}\right] \tag{2}$$

$$\widetilde{H}_{xi} = \frac{1}{\gamma_i^2 + k_i^2}\left[-j\alpha_n \frac{\partial}{\partial y}\widetilde{H}_{yi} + j\omega\varepsilon\Gamma\widetilde{E}_{yi}\right] \tag{3}$$

$$\widetilde{H}_{zi} = \frac{1}{\gamma_i^2 + k_i^2}\left[-\Gamma \frac{\partial}{\partial y}\widetilde{H}_{yi} + \omega\varepsilon\alpha_n\widetilde{E}_{yi}\right] \tag{4}$$

were: , is the propagation constant in "y" direction; α_n is the spectral variable in "x" direction. $k_i^2 = \omega^2\mu\varepsilon = k_0^2\varepsilon_{ri}^*$, is the wave number of ith term of dielectric region;
$\varepsilon_n^* = \varepsilon_{ri} - j\frac{\sigma_i}{\omega\varepsilon_0}$, is the relative dielectric constant of the material with losses;
$\varepsilon_i = \varepsilon_{ri}^* \cdot \varepsilon_0$, is the dielectric constant of the ith region;
$\Gamma = \alpha + j\beta$, is the complex propagation constant; and,
$\omega = \omega + j\omega_i$, is the complex angular frequency.

The solutions of the fields equations for the three regions in study are given by example: For region 2:

$$\widetilde{E}_{y2} = A_{2e} \cdot \text{senh}\gamma 2y + B_{2e} \cdot \cosh\gamma 2y \tag{5}$$

$$\widetilde{H}_{y2} = A_{2h} \cdot \text{senh}\gamma 2y + B_{2h} \cdot \cosh\gamma 2y \tag{6}$$

For the determination of the unknown constants described above the boundary conditions are applied.

For the propagation constant determination, the $\widetilde{E}_{xt}$ and $\widetilde{E}_{zt}$ (still unknown) components must be isolated in the magnetic field equations in the slots regions. Theses equations are shows in sequence:

$$\widetilde{H}_{x2} - \widetilde{H}_{x3} = \widetilde{J}_{zt} \tag{7}$$

$$\widetilde{H}_{z2} - \widetilde{H}_{z3} = -\widetilde{J}_{xt} \tag{8}$$

were $\widetilde{J}_{xt}$ and $\widetilde{J}_{zt}$ are the electric current densities in the fins.

Substituting the above equations in the magnetic fields equations (presented in the last section) and isolating the electric fields terms the admittance functions are obtained.

$$Y_{xx}\widetilde{E}_{xt} + Y_{xz}\widetilde{E}_{zt} = \widetilde{J}_{xt} \tag{9}$$

$$Y_{zx}\widetilde{E}_{xt} + Y_{zz}\widetilde{E}_{zt} = \widetilde{J}_{zt} \tag{10}$$

The $\widetilde{E}_{xt}$ and $\widetilde{E}_{zt}$ fields are expanded in terms of base functions and have a contribution of the two slots,

$$\widetilde{E}_{xt} = \sum_{i=1}^{n} a_{xi} \cdot \widetilde{f}_{xi} \tag{11}$$

$$\widetilde{E}_{zt} = \sum_{j=1}^{m} a_{zj} \cdot \widetilde{f}_{zj} \tag{12}$$

The base functions utilized in the Fourier Transform Domain are obtained; for the odd mode:

$$\widetilde{f}_{ixm}(\alpha_n) = \text{Re}\left\{\frac{\pi wi}{4} e^{j\alpha_n(xi+wi/2)}\left[e^{jm\pi/2} J_0\left[\frac{1}{2}(\alpha_n wi + m\pi)\right] + e^{-jm\pi/2} J_0\left[\frac{1}{2}(\alpha_n wi - m\pi)\right]\right]\right\} \tag{13}$$

$$\widetilde{f}_{izm}(\alpha_n) = \text{Im}\left\{-\frac{j\pi wi}{4} e^{j\alpha_n(xi+wi/2)}\left[e^{jm\pi/2} J_0\left[\frac{1}{2}(\alpha_n wi + m\pi)\right] - e^{-jm\pi/2} J_0\left[\frac{1}{2}(\alpha_n wi - m\pi)\right]\right]\right\} \tag{14}$$

And for the even mode:

$$\widetilde{f}_{ixm}(\alpha_n)=\mathrm{Im}\left\{\frac{\pi wi}{4}e^{j\alpha_n(xi+wi/2)}\left[e^{jm\pi/2}J_0\left[\frac{1}{2}(\alpha_n wi+m\pi)\right]+e^{-jm\pi/2}J_0\left[\frac{1}{2}(\alpha_n wi-m\pi)\right]\right]\right\} \quad (15)$$

$$\widetilde{f}_{izm}(\alpha_n)=\mathrm{Rm}\left\{-\frac{j\pi wi}{4}e^{j\alpha_n(xi+wi/2)}\left[e^{jm\pi/2}J_0\left[\frac{1}{2}(\alpha_n wi+m\pi)\right]-e^{-jm\pi/2}J_0\left[\frac{1}{2}(\alpha_n wi-m\pi)\right]\right]\right\} \quad (16)$$

were J_0 is the first species and zero order Bessel function; $x_i\, e\, w_i$ are the dimensional terms presented in Fig. 1(b) with $I = 1, 2$ for first and second slots.

In the sequence the Gallerkin method, particular case of Moment method, is used and, a new matrix homogeneous matrix with two variables is obtained.

$$\begin{bmatrix} K_{xx}^{11} & K_{xx}^{12} & K_{xz}^{11} & K_{xz}^{12} \\ K_{xx}^{21} & K_{xx}^{22} & K_{xz}^{21} & K_{xz}^{22} \\ K_{zx}^{11} & K_{zx}^{12} & K_{zz}^{11} & K_{zz}^{12} \\ K_{zx}^{21} & K_{zx}^{22} & K_{zz}^{21} & K_{zz}^{22} \end{bmatrix} \begin{bmatrix} a_{1x} \\ a_{2x} \\ a_{1z} \\ a_{2z} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad (17)$$

were for example:

$$K_{xx}^{1j} = \sum_{-\infty}^{\infty} Y_{xx}\widetilde{f}_{jx} \cdot \widetilde{f}_{1x}^{*} \quad (18)$$

$$K_{xz}^{1j} = \sum_{-\infty}^{\infty} Y_{xz}\widetilde{f}_{jz} \cdot \widetilde{f}_{1x}^{*} \quad (19)$$

The determinant of (17) is represented by a transcendent equation witch the roots are the attenuation constant "α" and the phase constant "β", and the complex propagation constant $\Gamma = \alpha + j\beta$, is obtained.

Finally effective dielectric constant is determined:

$$\varepsilon_{ef} = \left(\frac{\beta}{k_0}\right)^2 \quad (20)$$

3. DIRECTIONAL COUPLER

In the directional coupler when the signal arrives in port 1, port 3 will be coupled and port 4 isolated. The even and odd modes propagates with different velocities and coupling is periodical along the fin length.

The length required for total power transference from port 1 to port 3 is [6].

$$L = \pi/(\beta_{\text{even}} - \beta_{\text{odd}}) \quad (21)$$

were β_{even} and β_{odd} are the phase constants for the even and odd modes respectively, and are calculated with the TTL method. The coupling amplitude coefficient between ports 1 and 3 is,

$$|S_{13}| = \mathrm{sen}(\pi/2 \cdot l/L) \quad (22)$$

and the amplitude coefficient between ports 1 and 2 is,

$$|S_{12}| = \cos(\pi/2 \cdot l/L) \quad (23)$$

The length l is calculated with Eq. (48). The expression for the coupling is defined as [6]:

$$C_3 = 20.\log\left(\frac{1}{|s_{13}|}\right) \quad (24)$$

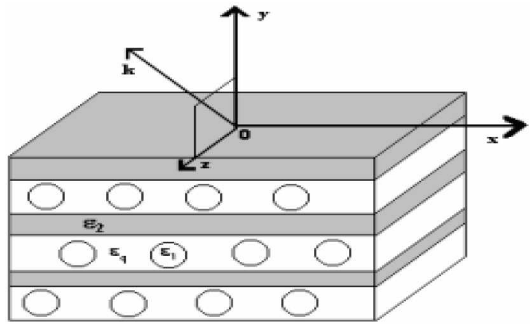


Figure 2: Homogenized bidimensional crystal.

4. PBG STRUCTURE

For a non-homogeneous structure submitted, the incident sign goes at the process of multiple spread. A solution can be obtained through a numerical process called homogenization [5]. The process is based in the theory related to the diffraction of an incident electromagnetic plane wave imposed by the presence of a air immerged cylinders in a homogeneous material.

In the Cartesian coordinates system of axes (O, x, y, z), are shown in the Fig. 2. A cylinder is considered with relative permittivity ε_1, with a traverse section in the plane xy, embedded in a medium of permittivity ε_2. For this process the two-dimensional structure is sliced in layers whose thickness is equal at the cylinder diameter. In each slice is realized the homogenization process.

According to homogenization theory the effective permittivity depends on the polarization. For the s and p polarization, respectively, we have:

$$\varepsilon_{eq} = \beta(\varepsilon_1 - \varepsilon_2) + \varepsilon_2 \tag{25}$$

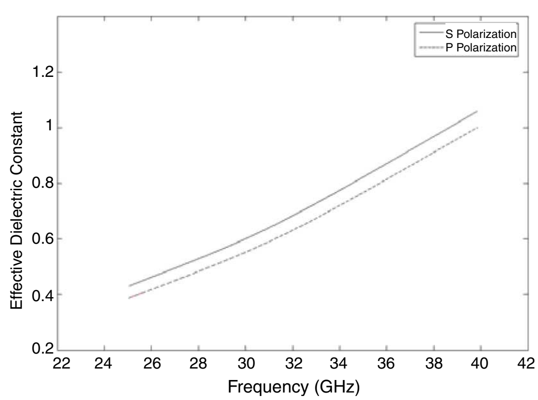
$$\frac{1}{\varepsilon_{eq}} = \frac{1}{\varepsilon_1}\left\{1 - \frac{3\beta}{A_1 + \beta - A_2\beta^{10/3} + O(\beta^{14/3})}\right\}, \tag{26}$$

where:

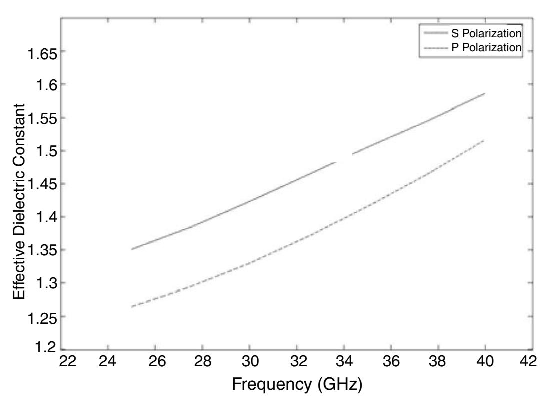
$$A_1 = \frac{2/\varepsilon_1 + 1/\varepsilon_2}{1/\varepsilon_1 - 1/\varepsilon_2} \tag{27}$$

$$A_2 = \frac{\alpha(1/\varepsilon_1 - 1/\varepsilon_2)}{4/3\varepsilon_1 + 1\varepsilon_2} \tag{28}$$

And β is defined as the ratio between the area of the cylinders and the area of the cells, α is an independent parameter whose value s equal to 0.523. The A1 and A2 variables in (27) and (28) were included only for simplify (26) equation.



(a)



(b)

Figure 3: Effective dielectric constant as a function of the frequency for a unilateral directional coupler with 2D PBG substrate, in a WR-28millimeter wave guide, (a) for the even mode (b) for the odd mode.

5. NUMERICAL RESULTS

For the numerical results determination a computational program in Fortran Power Station was developed according to the theory, using a Pentium IV, 3.4 GHz. The recourses present by the program include the determination of the attenuation, phase and effective dielectric constants for the even and odd modes.

The results were obtained for a unilateral directional coupler with 2D PBG substrate in WR-28 millimeter wave guide with dimensions $g = 0.254$ mm, $s = 3.302$ mm (region 1 thickness1), $s_1 = 1.078$ mm, $s_2 = 2.278$ mm, $w_1 = w_2 = 0.2$ mm, $\varepsilon_{r2} = 8.7209$ for p polarization and $\varepsilon_{r2} = 10.233$ for s polarization, $s' = 0.5$ mm (half the distance between the slots), $2a = 7.112$ mm and $2b = 3.556$ mm.

Figure 3 shows the effective dielectric constant as a function of the frequency Fig. 3(a) for the even mode and Fig. 3(b) for the odd mode.

The Fig. 4 shows the attenuation constant as a function of the frequency (a) for the even mode and (b) for the odd mode.

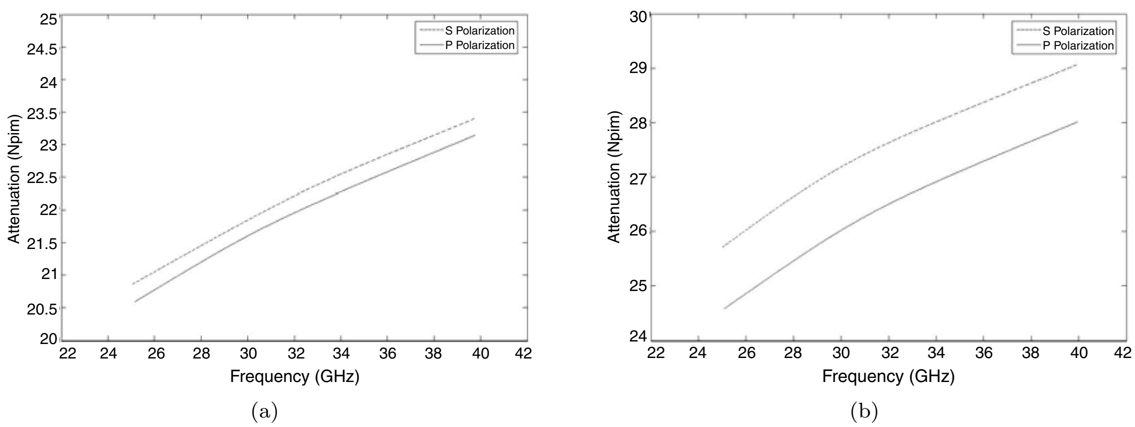


Figure 4: Attenuation constant as a function of the frequency for a unilateral fin line coupler with PBG substrate in a WR-28 millimeter wave guide; (a) for the even mode; (b) for the odd mode.

6. CONCLUSIONS

The full wave transverse transmission line (TTL) method was used to the characterization of the unilateral fin line directional coupler, considering a 2D Photonic Band Gap (PBG) substrate in the millimeter wave bands. The full wave TTL method was used to the electromagnetic fields determination. Numerical results for the attenuation, effective dielectric constant and coupling were presented.

ACKNOWLEDGMENT

This work was partially supported by CNPQ.

REFERENCES

1. Brat, B. and S. K. Koul, *Analysis, Design and Applications of Fin Lines*, Artech House, Inc., 383–471, Norwood, MA, USA, 1987.
2. Wiersma, D. S., P. Bartolini, A. Lagendijk, and R. Righini, "Localization of light in a disordered medium," *Letters to Nature*, Vol. 390, 18/25, Dec. 1997.
3. Grning, U., V. Lehmann, S. Ottow, and K. Busch, "Macroporous silicon with a complete two dimensional photonic band gap centered at 5 μm," *Appl. Phys. Lett.*, Vol. 68, Feb. 1996.
4. Radisic, V., Y. Qian, R. Coccioli, and T. Itoh, "Novel 2-D photonic bandgap structure for mictostrip lines," *IEEE Microwave and Guided Wave Letters*, Vol. 8, No. 2, Feb. 1998.
5. Centeno E. and D. Felbacq, "Rigorous vector diffraction of electromagnetic waves by bidimensional photonic crystals," *J. Optical Soc. American A*, Vol. 17, No. 2, 320–327, February 2000.
6. Silva, S. A. P., J. P. Silva, L. C. Freitas Jr, and H. C. C. Fernandes, "Coupling analysis at the coupler and edge-coupled unilateral fin line," *SPIE'98*, San Diego, USA, 1998.
7. Wiersma, D. S., P. Bartolini, A. Lagendijk, and R. Righini, "Localization of light in a dimensional medium," *Letters to Nature*, Vol. 390, 18/25, Dec. 1997.

8. Guillouard, K., M. F. Wong, V. Fouad Hanna, and J. Citerne, "Diakoptics using finite element analysis," *MWSYM 96*, Vol. 1, 363–366.
9. Tan, J. and G. Pan, "A general functional analysis to dispersive structures," *MWSYM 96*, Vol. 2, 1027–1030.
10. Rocha, A. R. B. and H. C. C. Fernandes, "Analysis of antennas with PBG substrate," *International Journal of Infrared and Millimeter Waves*, Vol. 24, 1171–1176, USA, Jul. 2003.

Novel Neural Network Models of Q-Type Integrals and Their Use for Circular-loop Antenna Analysis

Humberto César Chaves Fernandes[1], **M. G. Passos**[1], **and P. H. da F. Silva**[2]

[1]UFRN/TECFOTON — Campus Universitário, CEP: 59072-970 — Natal, RN, Brazil
[2]CEFET-PB/GTEMA — Av. 1° de Maio, 720 Jaguaribe, CEP: 58015-430 — João Pessoa, PB, Brazil

Abstract— In this paper, the neural network methodology was applied to develop analytical models, defined by means of weighted sums of basis functions, to approximate/interpolate the class of Q-type power-radiation integrals arising in antenna theory. The associated problems in this proposed approach, i.e., the choice of the type and number of basis functions, and the model's parameters optimization, were replaced by the problem of training weights in neural networks. The better neural model resultant of this research (obtained through of the *damped-sinusoid* basis functions set) was used for efficient evaluations of the circular-loop antenna radiation resistance and directivity. The neural model accuracy and computational efficiency were compared with recent approaches published in literature.

1. INTRODUCTION

In the recent years, the analysis of circular-loop antennas has been revisited. The analytical approximations for radiation resistance and directivity have been presented by means of series of Bessel functions [1–2] and approximate expressions [3]. A class of power-radiation integrals named Q-type integrals that arise in antenna theory were considered. Such integrals often appear when cylindrical coordinates are used in the analysis of the radiated power (particularly, in the case of a circular-loop, a circular microstrip antenna, and a circular aperture). The Q-type integrals representation was given in [1], using the usual Bessel function notation, as follows:

$$Q_{mn}^{(p)}(\alpha) = \int_{0}^{\pi/2} J_m[\alpha \sin(\theta)] J_n[\alpha \sin(\theta)] \sin^{(p)}(\theta) \mathrm{d}\theta \tag{1}$$

where $\alpha = ka$, $k = 2\pi/\lambda$; a is the loop radius; and $J_m(\cdot)$ is the Bessel function of the first kind, m-th order.

A general closed-form solution of Q-type integrals is not available in standard handbooks on integrals and special functions [5–6]. On the other hand, its evaluation is a difficult numerical task in the case of large antenna dimensions, due to the rapid oscillation of the integrand.

The aim of this work was to develop analytical models of Q-type integrals, defined by weighted sums of basis functions. The problems regarding the choice of the basis functions type/number and model's parameters optimization, which may result in better analytical models, were replaced by the problem of supervised learning in one hidden layer feed forward neural networks. Thus, the learning task corresponds to the function approximation of Q-type integrals, and the problem to obtain the analytical models is directly linked to that one corresponding to the training weights in neural networks. The Q-type integrals solutions were used in the analysis of circular-loop antennas, taking into account two current distributions: constant and co-sinusoidal. The new neural network model and the previously established approaches were compared with respect to the accuracy and computational efficiency performance criterions.

2. PROBLEM DEFINITION AND NEURAL NETWORK APPROACH

The evaluation of Q-type integrals can be a tedious and inefficient numerical exercise. However, for $n = m$ and $p = 1$, it is possible to circumvent this by rewriting (1) through of a simplified integral [4]:

$$Q_{nm}^{(1)}(\alpha) = \frac{1}{2\alpha} \int_{0}^{2\alpha} J_{2n}(t) \mathrm{d}t \tag{2}$$

Further, an analytical approach through of the series of Bessel functions was presented as follows [4, 7]:

$$Q_{nm}^{(1)}(\alpha) = \frac{1}{\alpha}\sum_{k=n}^{\infty} J_{2k+1}(2\alpha) \tag{3}$$

Exact, integral expressions for radiation resistance $R_r(\alpha)$ and directivity $D(\alpha)$, for all loop radius a, may be found in [8]. For convenience, the formulas are reproduced here, in the form:

$$R_r(\alpha) = 60(\pi\alpha)^2 T(\alpha) \tag{4}$$

$$D(\alpha) = F_m(\alpha)/T(\alpha) \tag{5}$$

For constant or co-sinusoidal loop currents, the auxiliary functions $F_m(\alpha)$ and $T(\alpha)$ are given by, [7–8]:

$$F_m(\alpha) = \begin{cases} \begin{cases} 2J_1^2(1.840) = 0.677, & \alpha > 1.840 \\ 2J_1^2(\alpha) & \alpha < 1.840' \end{cases} & I_\varphi = I_0 \\ 1, & I_\varphi(\varphi) = I_0\cos(\varphi) \end{cases} \tag{6}$$

$$T(\alpha) = \begin{cases} Q_{11}^{(1)}(\alpha), & I_\varphi = I_0 \\ \frac{1}{2}\left[Q_{00}^{(1)}(\alpha) - 2Q_{11}^{(1)}(\alpha)/a^2 + Q_{22}^{(1)}(\alpha)\right], & I_\varphi(\varphi) = I_0\cos(\varphi) \end{cases} \tag{7}$$

The neural network configuration developed to learn the general relationship defined by (2) is illustrated in Fig. 1, and it has been defined with one hidden layer of damped-sinusoid basis functions, in the form:

$$q_j(\mathbf{w}_j, \overline{\alpha}) = \frac{k_j}{\sqrt{|1-\xi_j^2|}} \exp(-\xi_j \tau_j \overline{\alpha}) \cdot \sin(\omega_j \overline{\alpha} + \phi_j) \tag{8}$$

where $\mathbf{w}_j = [k_j \quad \xi_j \quad \tau_j \quad \omega_j \quad \phi_j]^T$ designate the weight vector of the j-th node in the hidden layer; and $0 \leq \overline{\alpha} \leq 1$ is the scaled input of neural network model (after linear scaling).

The *damped-sinusoid* function, based on step response of a second order linear system, with five free parameters, is a very flexible function and can be assume various shapes depending of the weight vector values. It was chosen from various tested basis functions, including sigmoidal and Gaussian common functions, rational-sinusoid and sigmoidal-sinusoid proposal functions. The proposed model, resultant of this research, and that provides the better results, named DSFNN. — *Damped-Sinusoid Functions Neural Network*, was able to learn efficient approximations for Q-type integrals. In accordance with the DSFNN configuration (Fig. 1), with one linear output neuron, the overall DSFNN forward computation can be expressed, in the form:

$$q_{nn}(\mathbf{W}, \alpha) = \left[1 - \sum_{j=0}^{N_h-1} q_j(\mathbf{w}_j, \overline{\alpha})\right](q_{max} - q_{min}) + q_{min} \tag{9}$$

where $\mathbf{W} = [\mathbf{w}_0 \ldots \mathbf{w}_j \ldots \mathbf{w}_{Nh-1}]$ symbolize the neural network weight matrix; N_h is the number of hidden neurons (or basis function); and $q_{min,\,max} = \min, \max\left(\alpha Q_{nn}^{(1)}(\alpha)\right)$ correspond to the de-scaling parameters. Finally, the approximations for Q-type integrals are obtained through of the DSFNN model as follows:

$$Q_{nn}^{(1)}(\alpha) \approx \frac{1}{\alpha} q_{nn}(\mathbf{W}, \alpha) \tag{10}$$

By means of a trial and error process and training restarts, we choose the minimal DSFNN configuration (with 10 basis functions) needed to learn fast and accurate approximations for Q-type integrals.

The training datasets were obtained through of the precise computation of (2) in the interval $0.1 \leq \alpha \leq 25$. The involved evaluation/integration of Bessel functions were carried out using the Matlab specialized functions *besselj()* and *quadl()*, respectively. The *quadl()* Matlab function implements the Lobatto quadrature method and was applied using a tolerance parameter: $TOL = 1E-12$.

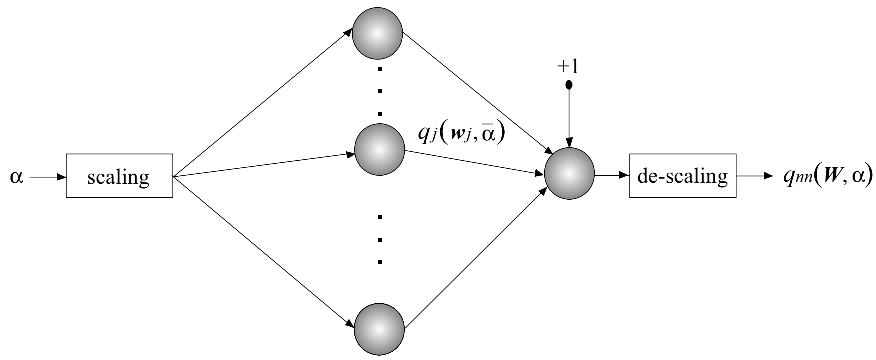


Figure 1: The DSFNN configuration.

In the DSFNN model optimization, the weight matrix adjustments were carried out by the use of the Rprop — *Resilient Backpropagation* algorithm, [9]. Again, by means of a trial and error process and training restarts, and from various trained DSFNNs, the best one was chosen to provide accurate approximations for Q-type integrals.

3. NUMERICAL RESULTS

The circular-loop antenna radiation resistance and directivity were calculated using four approaches: numerical integration, approximate expressions, series of Bessel functions (with 30 terms), and DSFNN approach (with 10 basis functions). The Lobatto quadrature numerical integration method was utilized to provide reference values (considered exacts) and to verify the accuracy from the others approaches.

When the DSFNN training is finished, the weight matrix $\boldsymbol{W}$ no more suffers adjustments and their weights values correspond to the model parameters through of the weighted sum of *damped-sinusoid* functions. Table (1) presents the DSFNN weights for approximation/interpolation of integral $Q_{11}^{(1)}(\alpha)$ using (8), (9) and (10).

Table 1: The DSFNN trained weights to approximate the $Q_{11}^{(1)}(\alpha)$ integral using (8), (9), and (10).

j	k_j	ξ_j	τ_j	ω_j	ϕ_j
0	−0.0709568	0.8130902	0.5086460	49.7032121	2.7010168
1	0.0311538	0.9052742	0.3059687	4.1402434	1.6282065
2	−0.0941926	0.8970338	12.70495940	9.6053935	0.3553070
3	0.1071911	0.9489009	0.70639380	−0.2338617	2.2606259
5	−0.0476289	0.9378222	4.3380320	48.5350052	3.0261462
6	0.0535575	0.9596534	13.1990856	42.8734507	0.7351803
7	0.1643971	0.8974420	23.5600778	25.2240012	1.9201970
8	0.1244032	0.9432581	1.1375807	0.2377621	0.9807844
9	−0.1069789	0.9107270	2.5510573	4.4421668	−0.1574657

Figure 2(a) shows the excellent agreement between DSFNN approximations and exact Q-integrals training data values (regarding , $n = 0$, 1, 2, and 1000 test samples over limited range, $0.1 \leq \alpha \leq 16$). Fig. 2(b) shows the obtained results for circular-loop directivity (regarding the constant and co-sinusoidal current distributions, and $0.1 \leq \alpha \leq 10$).

In Fig. 3(a)–(b) are presented the comparisons (using Matlab) with respect to the efficiency of the approaches for the directivity and radiation resistance evaluations, regarding the constant current distribution, and $0.1 \leq \alpha \leq 25$. Fig. 3(a) shows the errors on directivity calculations, and may be observed that the DSFNN approach presents an intermediary accuracy. While approximate expressions presents a maximum error of 0.2 dB, [3], the discrepancy between the DSFNN approximations and exact directivity values is, at worst, no more than about 0.00006 dB. The computational cost of each approach (used for the radiation resistance evaluations) is presented in Fig. 3(b). The CPU time spent by DSFNN approach was very close to the approximate expressions ones, and both approaches speedup the simulation (around 10x) with regard to the series of Bessel functions approach given in [7].

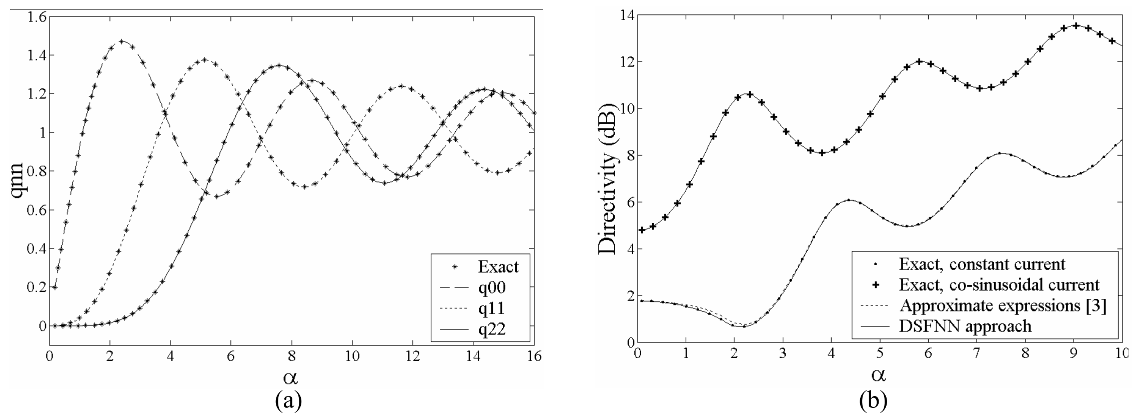


Figure 2: (a) The DSFNN approximations of Q-type integrals using (10). (b) The circular-loop antenna directivity evaluations through of the Lobatto quadrature(exact), approximate expressions and DSFNN approach.

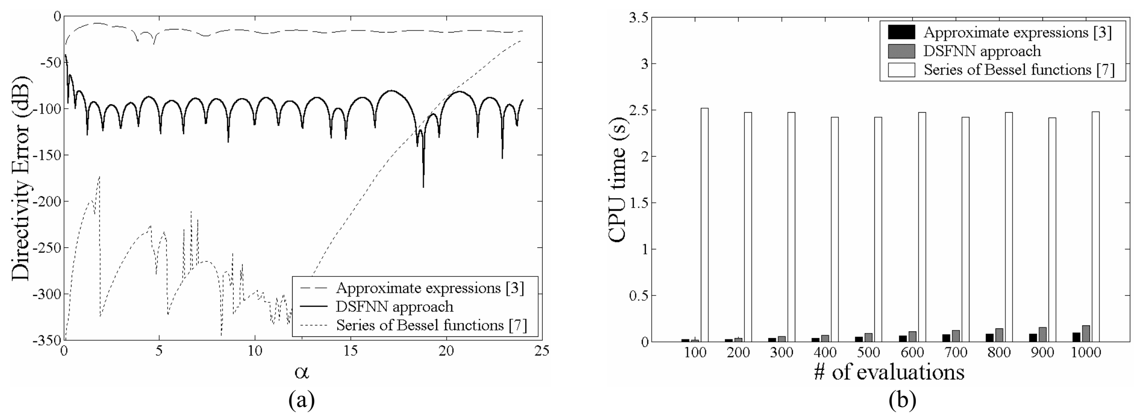


Figure 3: The comparisons among implemented approaches for circular-loop antenna analysis: (a) directivity errors; (b) computational costs.

4. CONCLUSIONS

In this paper, the neural network methodology was exploited to develop fast and accurate analytical models of Q-type integrals through of weighted sums of basis functions. Among tested basis functions, best results were obtained with the use of *damped-sinusoid* basis functions, regarding the resultant model accuracy and minimal number of basis functions necessary. The new DSFNN model, resultant of this research, was applied for efficient circular-loop antenna analysis. The computational cost of the DSFNN model was comparable with approximate expressions and speedup (around 10x) the series of Bessel functions approach. In addition, the DSFNN model is very accurate with maximum error for exact directivity values, at worst, no more than about 0.00006 dB.

This work was supported in part by CNPq and CAPES.

REFERENCES

1. Savov, S. V., "An efficient solution of a class of integrals arising in antenna theory," *IEEE Antennas and Propagation Magazine*, Vol. 44, No. 5, 98–101, 2002.
2. Mahony, J. D., "Circular microstrip-patch directivity revisited: an easily computable exact expression," *IEEE Antennas and Propagation Magazine*, Vol. 45, No. 1, 120–122, 2003.
3. Mahony, J. D., "Approximations to the radiation resistance and directivity of circular-loop antennas," *IEEE Antennas and Propagation Magazine*, Vol. 36, No. 4, 52–55, 1994.
4. Mahony, J. D., "A comment on Q-type integrals and their use in expressions for radiated power," *IEEE Antennas and Propagation Magazine*, Vol. 45, No. 3, 127–128, 2003.
5. Abramowitz, M. and I. A. Stegun, *Handbook of Mathematical Functions*, Dover, New York, 1970.

6. Gradshteyn, I. and I. Ryzhik, *Tables of Integrals, Series and Products*, Academic Press, New York, 1965.
7. Savov, S. V., "A comment on the radiation resistance,"*IEEE Antennas and Propagation Magazine*, Vol. 45, No. 3, 129, 2003.
8. Balanis, C. A., *Antenna Theory Analysis and Design*, Wiley-Interscience, New Jersey, 2005.
9. Riedmiller, M. and H. Braun, "A direct adaptive method for faster backpropagation learning: The rprop algorithm," *Proceedings of the IEEE International Conference on Neural Networks*, 586–591, San Francisco, EUA, 1993.

Coupled-mode Theory for Spun 4-lobe Stress Region Fibers

Jingren Qian and Xuxu Wang

Department of Electronic Engineering and Information Science
University of Science and Technology of China, Hefei 230026, Anhui, China

Abstract— Based on coupled-mode theory for LP modes, spun 4-lobe stress region fibers are theoretically analyzed. It is indicated for the first time that there are no any birefringence in spun 4-lobe stress region fibers if four lobes are symmetrically placed around the fiber axis. The same result is valid for spun fibers with 8-lobe and 16-lobe stress region, both have zero birefringence.

1. INTRODUCTION

The effort to create circular high-birefringence (hi-bi) fibers was never discontinued. Early in 1986, a spun 4-lobe core fiber was proposed by Fujii and Hussey as one of candidates of circular hi-bi fibers [1]. Since the suggested spun fiber is drawn in a viscous state at a high fiber-formation temperature and is not able to yield shear stress, no torsional circular birefringence occurs. They only intended to utilize index distribution to achieve circular high birefringence in spun fibers. Bassett pointed out that high birefringence could not acquire in such kinds of fibers as they expected, he suggested spin 4-lobe stress region to achieve circular hi-bi fibers [2]. However, Bassett has only given a range (from zero to 4π/spinning period) of circular birefringence for this spun fiber, which might mislead people to believe this spun fiber to be a nice candidate. According to authors' knowledge, so for the unclear conclusion for spun 4-lobe stress region fibers has not been changed.

In this paper, based on coupled-mode theory [3, 4], a rotated 4-lobe stress region fiber is analyzed and a conclusion of not any birefringence is obtained.

2. THEORETICAL ANALYSIS

The cross-section of a 4-lobe stress region fiber is shown in Fig. 1(a), there are four stress regions placed symmetrically around the fiber axis. Similar to a 2-lobe stress region fiber (see Fig. 1(b)) known as linear birefringence fiber, the four stress regions in the fiber cladding result in an anisotropic index change in the in the core region due to the elasto-optic effect. The argument

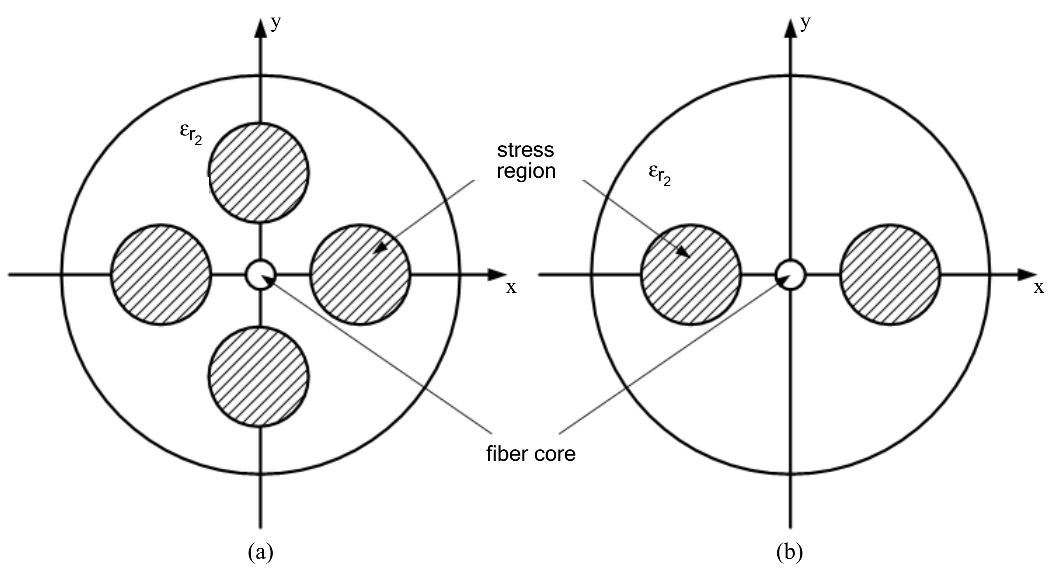


Figure 1: Cross sections of a 4-lobe stress region fiber and a dual -lobe stress region fiber.

makes use of two approximations. One is the weak guidance approximation, which is accurate for single-mode fibers and can be expressed as:

$$\frac{\varepsilon_{r_1} - \varepsilon_{r_2}}{\varepsilon_{r_2}} \ll 1 \tag{1}$$

Where ε_{r_1} and ε_{r_2} are relative dialectical constant of fiber core and cladding, respectively.

The other is the approximation of small changes of dielectic constants of the fiber core, which are induced by the stress, thus we have

$$\frac{\Delta\varepsilon_x}{\varepsilon_{r_2}} \ll 1, \quad \frac{\Delta\varepsilon_y}{\varepsilon_{r_2}} \ll 1 \tag{2}$$

The conditions (1) and (2) are valid for conventional PM fibers, i.e., dual-lobe stress region fibers.

Before we use coupled-mode theory, we have to express the anisotropic dielectric tensor of the additional stress-induced dielectric constant at the core region of an unspun 4-lobe stress region fiber. If the lobes are identical and are placed symmetrically to x- and y- axes as shown in Fig. 1, the directions of maximum local polarizability and so of a principal axis of a local dielectric tensor coincide in x- and y- axes.

Since the x- or y-polarized LP mode, LP_{01}^x or LP_{01}^y, has its electric vector in x- or y-direction only, respectively, it is adequate to describe the medium in the core region by two corresponding diagonal elements of the dielectric tensor as follows:

$$\Delta\overleftrightarrow{\varepsilon}(r,\varphi) = \begin{bmatrix} \Delta\varepsilon_x(r,\varphi) & 0 \\ 0 & \Delta\varepsilon_y(r,\varphi) \end{bmatrix} \tag{3}$$

where $\Delta\overleftrightarrow{\varepsilon}$ is the change of the relative dielectric constant induced by stress, $\Delta\varepsilon_x(r,\varphi)$ and $\Delta\varepsilon_y(r,\varphi)$ are the corresponding changes seen by x- and y-polarized light, respectively.

When the principal axes of the 4-lobe stress region fiber do not coincide in x- and y- axes and there exists a rotation angle θ between them as shown in Fig. 2, then the dielectric tensor will be as follows:

$$\Delta\overleftrightarrow{\varepsilon}_\theta(r,\varphi) = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \Delta\overleftrightarrow{\varepsilon} \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

or

$$\Delta\overleftrightarrow{\varepsilon}_\theta(r,\varphi) = \begin{bmatrix} \Delta\varepsilon_x\cos^2\theta + \Delta\varepsilon_y\sin^2\theta & -\frac{1}{2}(\Delta\varepsilon_x - \Delta\varepsilon_y)\sin 2\theta \\ -\frac{1}{2}(\Delta\varepsilon_x - \Delta\varepsilon_y)\sin 2\theta & \Delta\varepsilon_x\sin^2\theta + \Delta\varepsilon_y\cos^2\theta \end{bmatrix} \tag{4}$$

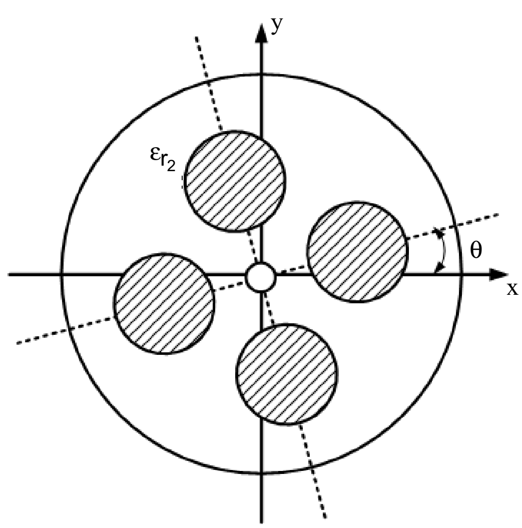


Figure 2: The counter-clockwise rotating angle θ between one of the fiber principal axes and x-axis.

Based on coupled-mode theory for LP modes, coupled-mode equations for the ideal LP_{01}^x and LP_{01}^y propagating in a 4-lobe stress region fiber with rotation angle θ (as shown in Fig. 2) can be obtained by inserting (4) in to (1)–(5) of *Reference* [3]:

$$\begin{aligned} \frac{\mathrm{d}A_x}{\mathrm{d}z} + j\beta A_x &= -j(\Delta\beta + C\cos 2\theta)A_x - jC\sin 2\theta A_y \\ \frac{\mathrm{d}A_y}{\mathrm{d}z} + j\beta A_y &= -jC\sin 2\theta A_x - j(\Delta\beta - C\cos 2\theta)A_y \end{aligned} \tag{5}$$

where A_x and A_y are the amplitudes of the LP_{01}^x and LP_{01}^y modes, respectively; β denotes the phase constant of the LP_{01} mode;

$$C = \frac{\omega\varepsilon_0}{2}\int_0^{\infty}\int_0^{2\pi}[\Delta\varepsilon_x(r,\varphi) - \Delta\varepsilon_y(r,\varphi)]e_{01}^2 r\mathrm{d}r\mathrm{d}\varphi \tag{6}$$

$$\Delta\beta = \omega\varepsilon_0\int_0^{\infty}\int_0^{2\pi}[\Delta\varepsilon_x(r,\varphi) + \Delta\varepsilon_y(r,\varphi)]e_{01}^2 r\mathrm{d}r\mathrm{d}\varphi \tag{7}$$

$e_{01}$is the normalized scalar mode function of LP_{01} mode.

It is easy to found from (5) that the coupling coefficient between the ideal LP_{01}^x and LP_{01}^y modes is equal to $C\sin 2\theta$. Then, let us evaluate C from (6) for rotated 4-lobe stress region fibers. Due to symmetry of this fiber, the distribution function of the dielectric constant is converted into itself by rotation through 90°, thus,

$$\Delta\varepsilon_x(r,\varphi \pm \pi/2) = \Delta\varepsilon_y(r,\varphi) \tag{8}$$

Using (8) in (6) we readily have

$$C = 0 \tag{9}$$

It implies that the ideal LP_{01}^x and LP_{01}^y modes propagating in 4-lobe stress region fibers have no coupling between them except an additional phase constant B, no matter how the rotation angle changes. Therefore the LP_{01}^x and LP_{01}^y modes propagate in a spun 4-lobe stress region fiber as a pair of eigen-modes in an ideal isotropic fiber. Consequently, there is not any birefringence in spun 4-lobe stress region fibers.

Doing it in more rigorous way, we utilize local modes or eigen-modes in unspun 4-lobe stress region fibers instead of ideal LP_{01}^x and LP_{01}^y modes in ideal single mode fibers. Due to condition (2), the electric fields of these two modes are modified from $\mathbf{i}_x e_{01}$(for LP_{01}^x mode) or $\mathbf{i}_y e_{01}$ (for LP_{01}^y mode):

$$\begin{aligned}\mathbf{e}_{01}^x(r,\varphi) &= \mathbf{i}_x e_{01}(r) + \Delta\mathbf{e}_1(r,\varphi)\\ \mathbf{e}_{01}^y(r,\varphi) &= \mathbf{i}_y e_{01}(r) + \Delta\mathbf{e}_2(r,\varphi)\end{aligned} \tag{10}$$

where $\Delta\mathbf{e}_1$ and $\Delta\mathbf{e}_2$ are the modifying fields resulted from appearance of $\Delta\varepsilon_x(r,\varphi)$ and $\Delta\varepsilon_y(r,\varphi)$, respectively. From (8) we have following relations between $\Delta\mathbf{e}_1$ and $\Delta\mathbf{e}_2$:

$$\begin{aligned}\mathbf{i_z} \times \Delta\mathbf{e_1}(r,\varphi + \frac{\pi}{2}) &= \Delta\mathbf{e_2}(r,\varphi)\\ \mathbf{i_z} \times \Delta\mathbf{e_2}(r,\varphi + \frac{\pi}{2}) &= -\Delta\mathbf{e_1}(r,\varphi)\end{aligned} \tag{11}$$

Starting from *References* [3, 4] , we also have coupled-mode equations (5) for the eigen-modes of the unspun fiber (Fig. 1) propagating in the same fiber except a rotation angle θ (see Fig. 2). Using (14) of *Reference* [4], we have the coupling coefficient C' equal to

$$\omega\varepsilon_0\int_0^{\infty}\int_0^{2\pi}[\mathbf{e_{01}^x}(r,\varphi)\cdot\Delta\vec{\vec{\varepsilon}}_\theta(r,\varphi)\cdot\mathbf{e_{01}^y}(r,\varphi)]r\mathrm{d}r\mathrm{d}\varphi \tag{12}$$

Inserting (10) along with (8) and (11) into (12), we also have $C' = 0$ or no coupling. Thus, we have reconfirmed that under condition of (1) and (2) spun 4-lobe stress region fibers have not any birefringence.

The 2-lobe stress region fibers (PM fibers) perform entirely differently [5] when rotating, the coupling between two eigen-modes occurs as well as linear birefringence, however, in spun 4-lobe stress region fibers, coupling caused by one pair of lobes is compensated by the other pair of lobes, consequently, no coupling can be found.

3. CONCLUSION

Coupled-mode equations for a pair of ideal modes propagating in a 4-lobe stress region fiber with rotation angle are formulated, no coupling between the ideal modes means no birefringence when spinning the fiber. The above conclusion is also valid for spun 8-lobe and 16-lobe stress region fibers.

ACKNOWLEDGMENT

Supported by Anhui Key Laboratory Optoelectronic Science and Technology

REFERENCES

1. Fujii, Y. and C. D. Hussey, "Design considerations for circularly form-birefringent optical fibers," *IEE proc. J Optoelectron.*, Part J, Vol. 133, No. 4, 249–255, 1986.
2. Bassett, L. M., "Design principle for a circularly birefringent optical fiber," *Optics Letters*, Vol. 13, No. 10, 844–846, 1988.
3. Qian, J. R. and W. P. Huang, "Couple-mode theory for *LP* modes," *J. of Lightwave Technology*, Vol. 14, No. 6, 619–626, 1986.
4. Qian, J. R. and W. P. Huang, "*LP*-modes and idea-modes on optical fibers," *J. of Lightwave Technology*, Vol. 14, No. 6, 626–630, 1986.
5. Qian, J. R. and L. Li, "Spun highly linearly birefringent fibers for current sensors," *Sci. China (Ser.A)*, Vol. 33, No. 1, 99–107, 1990.

The Effect of Radiation Coupling in Higher Order Fiber Bragg Gratings

Li Yang[1], **Wei-Ping Huang**[2], **and Xi-Jia Gu**[3]

[1]Department EEIS, University of Science and Technology of China, China
[2]Department ECE, McMaster University, Canada
[3]Department ECE, Ryerson University, Canada

Abstract— The effect of radiation coupling in second and higher order (marked as N-th order) fiber Bragg gratings is investigated by using coupled mode theory, where the radiation field is solved and expressed in terms of guided waves by using Green's function. Different from the direct index coupling between two contra-propagating guided modes due to the N-th order diffraction, the radiation coupling lead to the indirect self or mutual complex coupling between the two guided waves via the diffractions lower than N-th order. It will become dominant and result in some interesting properties potentially useful in the DFB fiber laser at some special cases such as a rectangular index modification with the duty cycle equals 0.5. For this reason, the dependences of coupling coefficients, spectrum characteristics and the bandwidth of Bragg reflection peak on the parameters of gratings with rectangular profiles, such as duty cycle, index change and grating length, are simulated and discussed in this paper.

1. INTRODUCTION

Over the past decades, much theoretical and experimental work has been devoted to the first order fiber Bragg Grating (FBG), while there are fewer reports on the higher order FBGs possibly because of the limitation of the fabrication techniques. However, the distinct properties of higher order FBGs indicate potential applications in laser, sensor and etc, which draw more and more interests [1–5].

One notable property in higher order FBGs is the complex radiation coupling via the diffraction lower than N-th order which is the order of FBG. It means, in the total coupling coefficient, in addition to index coupling coefficient due to the direct coupling between two contra-propagating guided waves similar to the first order gratings, there is radiation coupling coefficient indicating the indirect coupling between the two guided waves via the excitation and resonant of the radiation wave. Since the index coupling coefficient is proportional to the index change if it does not equal zero, while radiation coupling coefficient is proportional to the square of the index change, the effect of radiation coupling will be prominent only when the former one is around zero at some special cases such as a rectangular index modification with the duty cycle equaling 0.5 [3]. At such cases, the complex radiation coupling coefficient provides possibility to control the total coupling coefficient and then the spectrum characteristics in a wider range, which has been proved to be useful in DFB laser based on multilayer planar waveguides [6, 7].

On the other hand, in experiments, usually, the perturbed index profile in FBG is almost sinusoidal; and it will be more similar to rising-cosine after the saturation of the index modification [5]. During the growth of the index change, the higher order gratings attributed to the effect of the higher order Fourier expansion coefficients show up; meanwhile, the radiation coupling enhances gradually, which offers the possibility and necessity to study the effect of radiation coupling in higher order FBGs.

In this paper, the spectrum characteristics of higher order FBGs with rectangular index modification in a theoretical model are investigated by using coupled mode theory. Hereinto, the radiation fields are determined in terms of guided waves by using Green's function method. The influence of grating parameters such as duty cycle, index change and grating length on the coupling coefficients and the bandwidth of Bragg reflection peak are simulated and discussed.

2. THEORETICAL MODEL AND BASIC FORMULATIONS

A schematic diagram of a cylindrical symmetric single mode FBG in the theoretical model is shown in Fig. 1. In cylindrical coordinates based on the $\exp(j\omega t)$ time variation, the refractive index in the perturbed optical fiber can be expressed as,

$$\tilde{n}^2(r,z) = \begin{cases} \tilde{n}_{co}^2 \cong n_{co}^2 + \sum\limits_{m=-\infty}^{\infty} \eta_{\mathrm{m}} \exp\left(-\mathrm{j}\frac{2\mathrm{m}\pi}{\Lambda}z\right) = n_0^2 + \sum\limits_{\substack{m=-\infty \\ m\neq 0}}^{\infty} \eta_{\mathrm{m}} \exp\left(-j\frac{2\mathrm{m}\pi}{\Lambda}z\right), & 0 < r \leq r_0 \\ n_{cl}^2, & r > r_0 \end{cases} \tag{1}$$

where r_0, n_{co} and n_{cl} are respective the core radius, the refractive index of the core and the cladding in the unperturbed fiber, while n_0 is the refractive index of the core in the reference fiber. The Fourier expansion coefficients of the periodic index modification is

$$\eta_{\mathrm{m}} = \frac{2n_{co}\delta n}{\Lambda} \int_{-\Lambda/2}^{\Lambda/2} f(z) \exp\left(\mathrm{j}\,\frac{2\mathrm{m}\pi}{\Lambda}z\right) dz \tag{2}$$

$f(z)$ is the profile of the periodic modification.

Following the previous analysis for a N-th order grating, the electrical field is expressed in terms of two contra-propagation fundamental guided waves (named as LP_{01} mode in fiber) and radiation waves due to the resonant via the diffraction lower than N-th [3, 6, 7]. For a fiber with a small index difference, the radiation waves E_n^{rad} can be solved and expressed in terms of the two guided waves by Green's function in a uniform region [3]. Then from the partial differential equation the electrical field of LP_{01} mode satisfies, the following coupled mode equations are obtained as follows,

$$\begin{aligned} \frac{\mathrm{d}a(z)}{\mathrm{d}z} &= -\mathrm{j}\left(\delta + \kappa_{\mathrm{sa}}^{\mathrm{rad}}\right) a(z) - \mathrm{j}\left(\kappa_{\mathrm{N}} + \kappa_{\mathrm{m(ab)}}^{\mathrm{rad}}\right) b(z) \\ \frac{\mathrm{d}b(z)}{\mathrm{d}z} &= \mathrm{j}\left(\delta + \kappa_{\mathrm{sb}}^{\mathrm{rad}}\right) b(z) + \mathrm{j}\left(\kappa_{-\mathrm{N}} + \kappa_{\mathrm{m(ba)}}^{\mathrm{rad}}\right) a(z) \end{aligned} \tag{3}$$

In (3), δ is the detune factor,

$$\delta = \beta - \beta_0 = \beta - \frac{\mathrm{N}\pi}{\Lambda} \tag{4}$$

$\kappa_{\pm\mathrm{N}}$is the index coupling coefficient analytically expressed as

$$\kappa_{\pm\mathrm{N}} = \frac{\pi Y A_0^2 r_0^2 k_0^2 \eta_{\pm\mathrm{N}}}{4\beta_0 \mathrm{J}_0^2(k_{1t}r_0)} \left[\mathrm{J}_0^2(k_{1t}r_0) + \mathrm{J}_1^2(k_{1t}r_0)\right] = \frac{k_0}{2}\eta_{\pm\mathrm{N}}\Gamma \tag{5}$$

where Y and A_0 are respectively the mode admittance and normalized amplitude of LP_{01} mode in the reference fiber. $\mathrm{J}_0(k_{1t}r_0)$ and $\mathrm{J}_1(k_{1t}r_0)$ are Bessel Functions, where k_{1t} is the transverse propagation constant of LP_{01} mode in the core. More briefly, $\kappa_{\pm\mathrm{N}}$ is expressed as the product of the half of the propagation constant in free space, k_0, the $\pm N$-th order Fourier expansion coefficient, $\eta_{\pm\mathrm{N}}$, and the confinement factor of LP_{01} mode, Γ.

When $N > 1$, the analytical expression of self radiation coupling coefficients $\kappa_{\mathrm{sa}}^{\mathrm{rad}}$ and $\kappa_{\mathrm{sb}}^{\mathrm{rad}}$, mutual radiation coupling coefficients $\kappa_{\mathrm{m(ab)}}^{\mathrm{rad}}$ and $\kappa_{\mathrm{m(ba)}}^{\mathrm{rad}}$ are obtained by using Green's function method and analytically expressed as follows,

$$\kappa_{\mathrm{sa}}^{\mathrm{rad}} = \sum_{n=1}^{\mathrm{N}-1} \kappa_{n,-n}^{\mathrm{rad}}, \quad \kappa_{\mathrm{sb}}^{\mathrm{rad}} = \sum_{n=1}^{\mathrm{N}-1} \kappa_{n-\mathrm{N},\mathrm{N}-n}^{\mathrm{rad}}, \quad \kappa_{\mathrm{m(ab)}}^{\mathrm{rad}} = \sum_{n=1}^{\mathrm{N}-1} \kappa_{n,\mathrm{N}-n}^{\mathrm{rad}}, \quad \kappa_{\mathrm{m(ba)}}^{\mathrm{rad}} = \sum_{n=1}^{\mathrm{N}-1} \kappa_{n-\mathrm{N},-n}^{\mathrm{rad}} \tag{6}$$

$$\kappa_{i,j}^{\mathrm{rad}} = -\frac{\mathrm{j}\pi^2 Y k_0^4 A_0^2 \eta_i \eta_j}{4\beta_0 \mathrm{J}_0^2(k_{1t}r_0)\left(k^2 - k_{1t}^2\right)} \int_0^{r_0} C(r)\mathrm{J}_0(k_{1t}r) r\mathrm{d}r, \quad n = 1, 2 \tag{7}$$

$$\begin{aligned} C(r) &= kr\mathrm{J}_0(k_{1t}r)\left[\mathrm{H}_0^{(2)}(kr)\mathrm{J}_1(kr) - \mathrm{H}_1^{(2)}(kr)\mathrm{J}_0(kr)\right] \\ &\quad + r_0\left[k\mathrm{H}_1^{(2)}(kr_0)\mathrm{J}_0(k_{1t}r_0) - k_{1t}\mathrm{H}_0^{(2)}(kr_0)\mathrm{J}_1(k_{1t}r_0)\right]\mathrm{J}_0(kr) \end{aligned} \tag{8}$$

where k is propagation constant in uniform region. For symmetrical structure, there are $\kappa_{\mathrm{N}} = \kappa_{-\mathrm{N}}$ and $\kappa_{\mathrm{rad}}^{=}\kappa_{\mathrm{sa}}^{\mathrm{rad}} = \kappa_{\mathrm{sb}}^{\mathrm{rad}} = \kappa_{\mathrm{m(ab)}}^{\mathrm{rad}} = \kappa_{\mathrm{m(ba)}}^{\mathrm{rad}}$. The total coupling coefficient is defined as $\kappa_{\mathrm{total}} = \kappa_{\mathrm{N}} + \kappa_{\mathrm{rad}}$.

The reflection, transmission spectrums are obtained by solving the coupled-mode equations via transfer matrix method. The radiation power is obtained by subtracting the reflection and transmission power from the total power.

3. NUMERICAL RESULTS

The effect of radiation coupling in second order FBG with theoretically rectangular index modification is concentrated on. The designed Bragg wavelength is set at 1.55 μm.

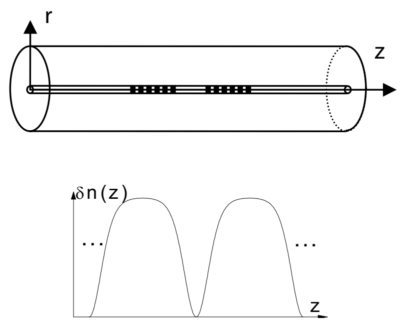


Figure 1: Schematic diagram of FBG with an arbitrary profile.

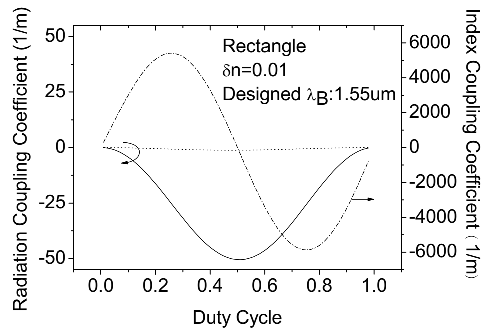


Figure 2: Dependences of coupling coefficients on duty cycle.

The effect of duty cycle is firstly studied. Seen from Fig. 2, the index coupling coefficient sinusoidally varies, obtaining the maximum values when duty cycle equals 0.25 and 0.75, equaling zero and changing sign when duty cycle equals 0.5. The radiation coupling coefficient varies with the square of sinusoidal function, obtaining the maximum value when duty cycle equals 0.5. These are all determined by the corresponding analytical expressions of coupling coefficients and the Fourier expression properties of the periodic rectangular function. It is thus not difficult to understand the radiation wave becomes prominent with the decrease of the reflection power only when the duty cycle is about 0.5, seen from Fig. 3. The prominent occurrence of the radiation wave corresponds to the narrowest 3 dB bandwidth of the Bragg reflection peak seen from Fig. 4. It also leads to a blue shift of the Bragg wavelength, seen from Fig. 5. Conclusively, at the special case of duty cycle around 0.5, the dominant complex radiation coupling results in a decrease of the total coupling coefficient and offer a possibility to control the total coupling coefficient in a wider range, which thus leads to some properties different from the first order FBG, such as a lower reflection power, a narrower and blue-shifted Bragg reflection peak.

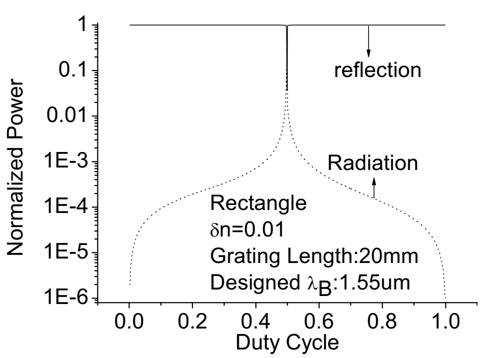


Figure 3: Dependence of reflection and radiation power on duty cycle.

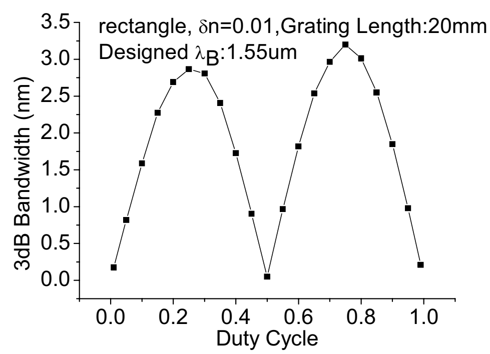


Figure 4: Dependence of 3 dB bandwidth on duty cycle.

Moreover, taking the advantage of other grating parameters, the effect of radiation coupling mentioned above can be further controlled. Seen from Figs. 6–8, due the enhancement of radiation coupling with the increase of the index modification, the reflection power will increase and the Bragg reflection peak will be broader. On the other hand, the increase of the grating length will result in the increase of the reflection power and the decrease of the Bragg reflection bandwidth, as shown in Figs. 9–10.

4. CONCLUSION

By using Coupled mode theory and Green's function method, the effect of radiation coupling in higher order FBG is modeled and simulated. For a second order FBG with rectangular index

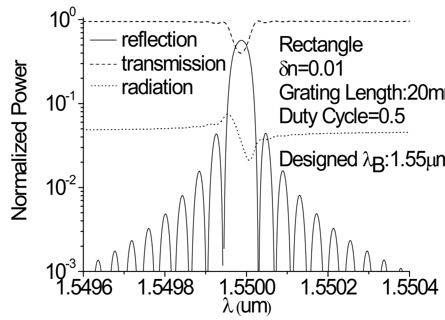


Figure 5: Spectrums at duty cycle equaling 0.5.

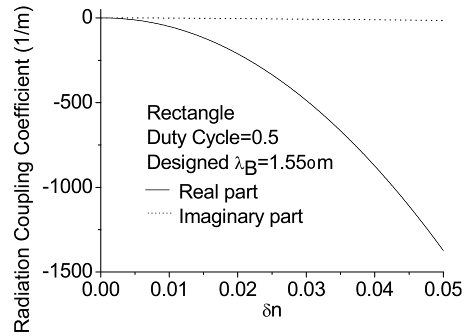


Figure 6: Dependences of coupling coefficients on index change.

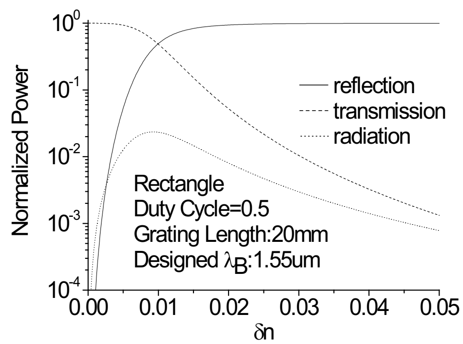


Figure 7: Dependence of powers on index change.

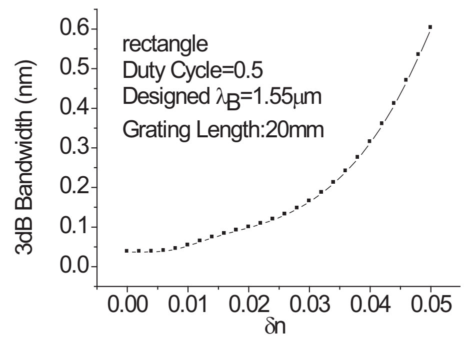


Figure 8: Dependence of 3dB bandwidth on index change.

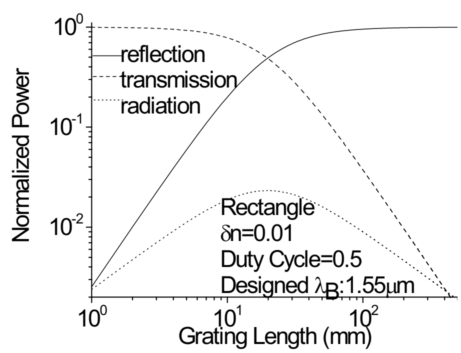


Figure 9: Dependence of powers on grating length.

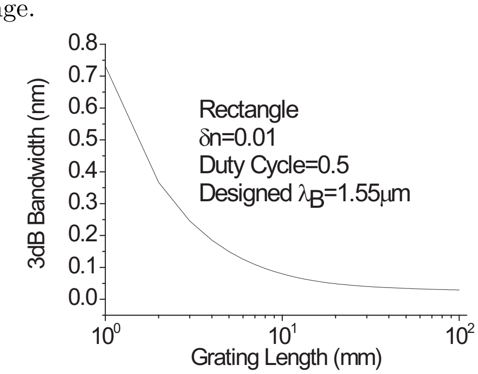


Figure 10: Dependence of 3 dB bandwidth on index change.

modification, when the duty cycle is around 0.5, the radiation wave becomes prominent. Meanwhile, with the elimination of the usually dominant index coupling, the complex radiation coupling is dominant in the total coupling coefficient. It results in the decrease and a wider controllable range for the total coupling coefficient, and then the decrease of reflection power and Bragg reflection bandwidth. These properties are simultaneously influenced by the index change and grating length. The study will be useful for the applications of high order FBGs, such as the design of DFB fiber laser.

ACKNOWLEDGMENT

Thank the Chinese Academy of Sciences for providing Li YANG with the fellowship and the Ontario Photonics Consortium for financial support.

REFERENCES

1. Xie, W. X., et al., "Second order diffraction efficiency of Bragg gratings written within germanosilicate fibres," *Opt. Commun.*, Vol. 101, No. 1–2, 85–91, 1993.

2. Jia, H. Z. and Y. L. Li, "First- and second-order diffraction characteristics of fiber Bragg gratings," *Opt. Commun.*, Vol. 178, 339–343, 2000.
3. Soltani, T., "Second order Bragg grating in slab waveguide and optical fiber," Master thesis, McMaster University, 2000.
4. Echevarria, J., et al., "Uniform fiber Bragg grating first-and second-order diffraction wavelength experimental characterization for strain-temperature discrimination," *IEEE, Phot. Tech. Lett.*, Vol. 13, No. 7, 696–698, 2001.
5. Carballar, A., et al., "Growth modeling of fiber gratings: a numerical investigation," *Fiber and Integrated Optics*, Vol. 21, 451–461, 2004.
6. Hardy, A., et al., "Analysis of second-order gratings," *IEEE J. QE*, Vol. 25, No. 10, 2096–2105, 1989.
7. Shams-Zadeh-Amiri, A. M., et al., "Second- and higher order resonant gratings with gain or loss-Part I: Greens function analysis," *IEEE J. QE*, Vol. 36, No. 12, 1421–1430, 2000.

Applications of Cladding Stress Induced Effects for Advanced Polarization Control in Silicon Photonics

D.-X. Xu, **P. Cheben**, **A. Delâge**, **S. Janz**, **B. Lamontagne**, **M.-J. Picard**
E. Post, **P. Waldron**, **and W. N. Ye**

Institute for Microstructural Sciences, National Research Council Canada (NRC)
Ottawa, Ontario, Canada

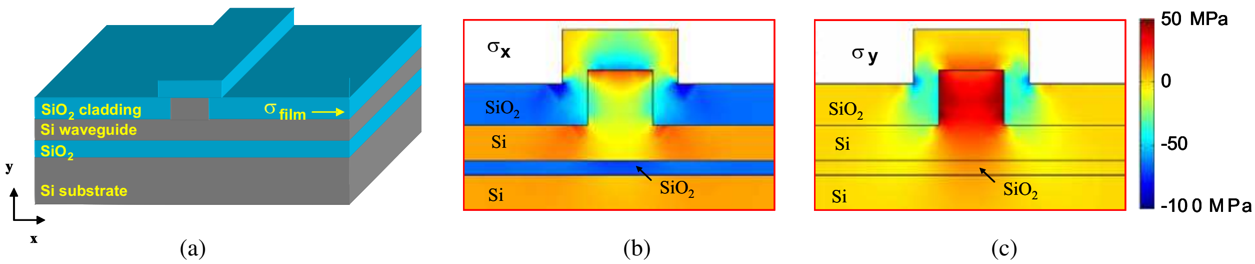
Abstract— The applications of cladding stress in SOI waveguide components to enhance device functionality and improve fabrication tolerance are reviewed. Assisted by FEM design tools, characteristics of stress-induced effects are studied in depth. Design strategies are developed for using stress engineering to achieve a variety of functions. Polarization insensitive arrayed waveguide gratings (AWGs) and ring resonators, and polarization splitters and filters are demonstrated using these design principles.

Silicon photonics is a rapidly growing research field with many advances in recent years. Motivated by the potential of high integration density, compatibility with mature CMOS technologies, silicon-on-insulator (SOI) has been the main material platform for silicon photonic waveguide components [1]. Along with these benefits, control and utilization of polarization dependent properties arise as new challenges, as well as opening new possibilities of novel designs and functionalities.

Polarization dependent properties have long been an important issue in integrated optical systems. In applications such as wavelength demultiplexing and high resolution spectroscopy, the shift in channel wavelength with the polarization state of the incoming optical signal often limits the device spectral resolution. One approach to handle this issue is to produce components or systems with polarization insensitive performance. In some cases this approach may not be practical, and then polarization diversity may be adopted where the signal is divided into orthogonal polarization states and processed separately. In this contribution, we review the solutions we have developed to manage the polarization, namely using cladding stress-induced effects to control the waveguide polarization properties in SOI waveguides. Design methodologies for using this technique to achieve polarization-insensitive arrayed waveguide gratings (AWGs) and ring resonators, Mach-Zehnder interferometer (MZI) based polarization splitters and filters are presented.

In planar waveguides, the modal birefringence results from a combination of geometrical, compositional and stress-induced effects. In low index contrast glass waveguides, the geometrical effect is minimal and the birefringence is primarily controlled by the material residual stress. A large body of research has been devoted to this subject, where the main goal is to reduce the residual stress in the waveguide core region by adjusting the thermal expansion coefficients of the cladding and core layers. In high index contrast systems such as SOI where the light confinement is strong, electromagnetic boundary conditions dictated by the cross-section geometry of the waveguides have the largest impact on the waveguide effective index and also the birefringence. These geometrical dependencies have been used to obtain single mode waveguides with large cross-sections, photonic wires with submicron dimensions which can afford bending radius on the order of microns, and to adjust the waveguide (geometrical) birefringence Δn_{geo}. Although it is possible to minimize the birefringence by tailoring the waveguide aspect ratio [2–4], this method only works well for waveguides with large cross-sections. As the core size is reduced, typically to 2 μm, it becomes increasingly difficult to maintain the designed birefringence values. Since ridge dimensions also determine the number of waveguide modes, the minimum usable bend radius, the mode size, and the coupling between adjacent waveguides, it is often impossible to simultaneously meet these different design objectives.

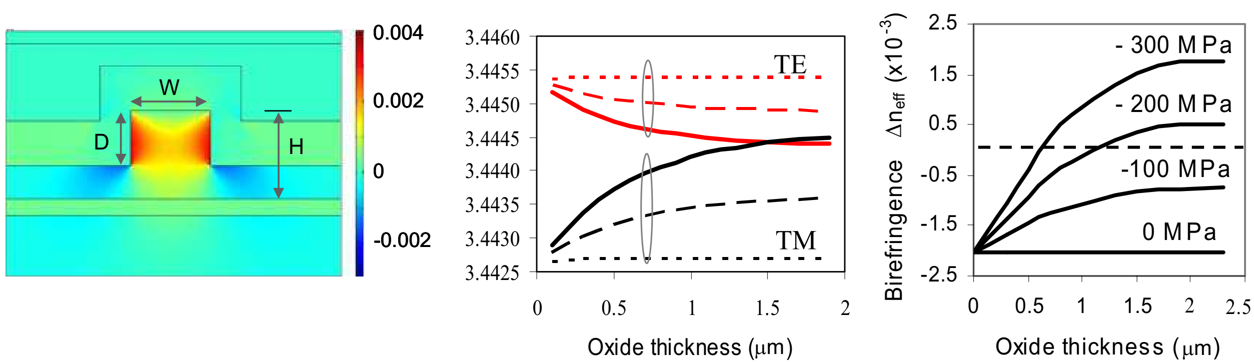
When a stress with axial anisotropy is imposed on the originally isotropic material, a stress contribution Δn_s to the modal birefringence is induced. The total modal birefringence can then be expressed as $\Delta n_{\text{eff}} = n^{\text{TM}} - n^{\text{TE}} = \Delta n_{\text{geo}} + \Delta n_s$, provided the stress induced index change is much smaller than the core-cladding index step, which is the case for high index contrast SOI waveguides. Cladding layers such as silicon dioxide deposited or grown over the silicon core generally retain a stress, mainly due to the thermal mismatch between the materials, and in turn create an anisotropic stress field inside the Si waveguide. Due to the photoelastic effect, the material birefringence is given by [5–7]: $(n_y - n_x) = (C_2 - C_1)(\sigma_y - \sigma_x)$. Here $\sigma_i (i = x, y)$ are the stress



(a) (b) (c)

Figure 1: (a) Cross-section of a SOI ridge waveguide; Stress distributions in the (b) x-direction, and (c) y-direction. The ridge height is 2.2 µm, the width is 1.8 µm and the etch depth is 1.47 µm. The cladding oxide is 1 µm thick and the stress is $\sigma_{\text{film}} = -320$ MPa.

tensor components, n_i is the material refractive index, n_0 is the refractive index without stress, and C_1 and C_2 are the stress-optic constants. We have evaluated the stress-induced effects using a finite-element differential equation solver (FEMLAB) and include stress equations that modify the permittivity tensor via the photoelastic effect. The stress distributions in a SOI ridge waveguide (Fig. 1(a)) are shown in Figs. 1(b) and (c). The corresponding local material birefringence ($n_y - n_x$) is shown in Fig. 2(a), with values as large as 4×10^{-3}. These modifications in the material cause $n_{\text{eff}}^{\text{TM}}$ to increase with the oxide thickness and the stress magnitude, and $n_{\text{eff}}^{\text{TE}}$ to decrease (Figs. 2 (b) and (c)). Two parameters control the stress-induced birefringence: the oxide thickness t_c and the film stress level σ_{film}. The stress-induced index variations resulted from commonly used cladding films are of comparable magnitude to the geometrical birefringence found in typical SOI ridge waveguides. Depending on the value of the geometrical birefringence Δn_{geo}, the total modal birefringence may be designed to be zero or other desired values. Fig. 3 gives the calculated results of maximum birefringence that can be induced by a cladding of -300 MPa in waveguides with 2 µm ridge height and different aspect ratios. These characteristics are the bases for stress engineering in SOI waveguides, and they can be advantageously employed in many applications.



(a) (b) (c)

Figure 2: (a) Stress-induced material birefringence ($n_y - n_x$) corresponding to the stress distributions shown in Figs. 1(b), (c); (b) Effective index vs. the oxide upper cladding thickness, for different cladding stress levels σ_{film} of 0 MPa (dotted lines), -200 MPa (dashed lines), and -400 MPa (solid lines), respectively, (c) Δn_{eff} as a function of t_c and σ_{film}.

In the following, we present several examples of using stress engineering for the control of polarization properties. First we discuss the case of polarization insensitive AWG demultiplexers. In AWGs (Fig. 4(a)) the polarization dependent wavelength shift $\Delta\lambda = \lambda_{\text{TM}} - \lambda_{\text{TE}}$ is mainly determined by the birefringence of the arrayed waveguides as $\Delta\lambda = \lambda \Delta n_{\text{eff}}/n_g$, where n_g is the waveguide group index and λ is the free-space operating wavelength. Fig. 4(b) shows the TE and TM spectra of an AWG with a waveguide ridge height of 2.2 µm. The geometrical birefringence of the as fabricated arrayed waveguides was $\Delta n_{\text{geo}} \sim -1.3 \times 10^{-3}$, giving a polarization dependent wavelength shift of $\Delta\lambda \sim 0.6$ nm. Here the waveguides were mainly designed to support only a single mode, while achieving reasonable coupling and desired bending radius. A PECVD oxide upper cladding film was then deposited, with the stress level and thickness properly chosen (-320 MPa

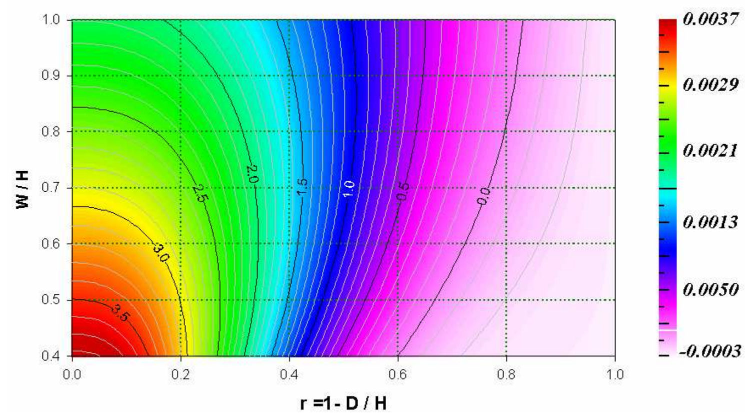
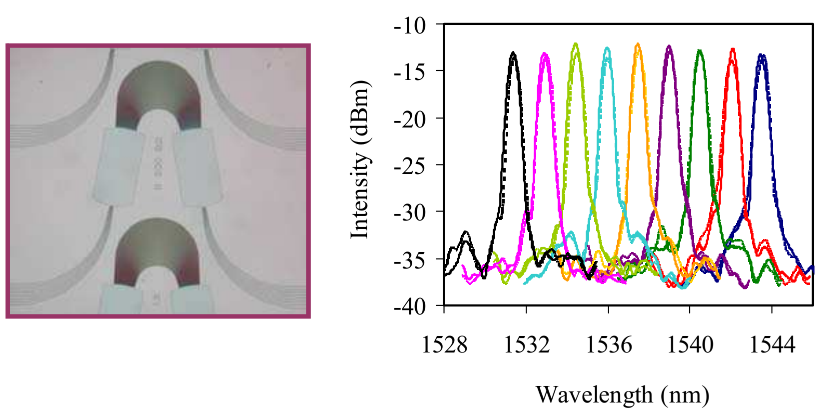


Figure 3: Stress-induced birefringence Δn_s for ridges with different aspect ratios, with $H = 2\,\mu\text{m}$, $t_c = 2\,\mu\text{m}$ and $\sigma_{\text{film}} = -300$ MPa.

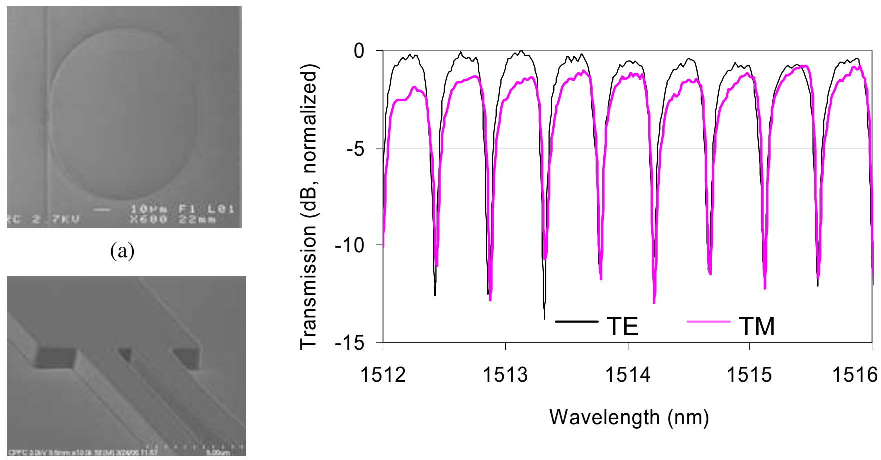
and $0.6\,\mu\text{m}$) using calculations similar to that presented in Fig. 2. The $\Delta\lambda$ was then reduced to below 0.04 nm (corresponding to $\Delta n_{\text{eff}} < 1 \times 10^{-4}$) [8]. Polarization dependent loss was also negligible in these devices. The cladding stress level can also be adjusted by post-processing such as thermal annealing, providing a practical method of post-process tuning [9]. This technique for polarization management is also applicable to high resolution microspectrometers developed in our laboratory. Compared to demultiplexers, spectrometers require a better resolution and a larger free spectral range, while crosstalk and band flatness are relatively minor factors in the specifications. A silica-on-silicon demultiplexer used as a high resolution and high bandwidth microspectrometer would lead to prohibitively large footprint; while the use of SOI can considerably reduce its size. Our present spectrometer prototype has 100 output channels designed with $1.5 \times 2\,\mu\text{m}^2$ arrayed waveguide cross-sections and occupies $8 \times 8\,\text{mm}^2$, two orders of magnitude less than a similar device based on glass waveguides [10]. Polarization dependence is again an important issue here, which can be managed with the assistance of stress engineering.



(a) (b)

Figure 4: (a) Optical image of an fabricated AWG, (b) Measured spectra for an SOI AWG with $\Delta\lambda$ compensated using $0.6\,\mu\text{m}$ thick SiO_2 cladding with $\sigma_{\text{film}} = -320\,\text{MPa}$. TM (solid) and TE (dashed).

We also applied the approach of stress engineering to design polarization-insensitive ring resonators (RR) [11]. Ring resonators can be used as the building block in applications including add-drop, switching, modulation, and sensing. However, polarization sensitivity is often an obstacle that limits their use. Polarization-independent RR designs using directional couplers were investigated [4], but showed very stringent fabrication requirements. We proposed the use of multi-mode-interference (MMI) couplers for light transfer between the bus and ring waveguides (Figs. 5(a) and (b)). Compared to the commonly used directional couplers, MMI couplers provide broadband and low polarization sensitivity in coupling ratios. The polarization dependent shift in resonator phase accumulation caused by waveguide geometrical birefringence Δn_{geo} in the different sections can be compensated by applying the appropriate cladding stress-induced birefringence Δn_s. We have fabricated MMI-based ring resonators in SOI wafers with 1.5 μm thick silicon at the Canadian Photonics Fabrication Center (CPFC). A standard in-house PECVD process produces oxide films with −250 MPa stress. Simulations similar to that in Fig. 2 show that an oxide thickness of 0.8 μm will compensate the overall birefringence. Preliminary measurement results are in shown in Fig. 5(c). Although there is some polarization dependent loss, the birefringence is well compensated in these devices.



(a) (b) (c)

Figure 5: (a) A ring resonator using an MMI coupler, (b) a close-up of the MMI section, (c) TE and TM transmission spectra of the ring resonator, with an upper SiO_2 cladding of 0.8 μm thick and film stress of $\sigma_{\text{film}} = -250$ MPa.

Stress-induced effects can also be used to produce high level of birefringence in selected areas. Polarization splitters and filters can be made by applying a cladding patch on one arm of a Mach-Zehnder interferometer (MZI). The cladding stress induces a phase difference $\Delta\varphi$. between the two interference paths. When the phase relations between the two arms are such that $\Delta\varphi^{\text{TE}} = 2M\pi$ and $\Delta\varphi^{\text{TM}} = (2N+1)\pi$ (M and N are integers), the device can operate as either a polarization splitter or filter, depending on the number of output ports of the combining coupler. An extinction ratio of −35 dB is predicted for a polarization filter [12].

As we have demonstrated, stress-induced modifications of the effective indices for the TE and TM modes and the associated birefringence are important for a wide range of commonly employed SOI waveguide geometries. The significance of these effects is starting to be recognized in the research community. If not taken into consideration, these effects can lead to large deviations in device characteristics from the designed specifications. Fortunately, the use of cladding stress to correct the birefringence allows a considerable degree of freedom in designing SOI waveguides to meet other performance criteria such as relaxed dimensional tolerance, reduced loss at waveguide bends, and overall improved device performance. The stress-induced modifications to the effective indices in SOI waveguides are mainly controlled by the stress level and the thickness of the upper oxide cladding. As we have demonstrated, polarization insensitivity in AWGs, spectrometers and ring resonators can be achieved using this technique, leaving the freedom of optimizing the waveguide geometry for considerations other than the birefringence. Since the effect of a SiO_2 cladding on mode shape is negligible, there is little mode mismatch loss or polarization dependent loss at the junctions between waveguide sections with and without the claddings. Therefore tailored cladding patches can be applied at discrete locations in a planar waveguide circuit with negligible insertion loss penalty, which we demonstrated in the design of broadband polarization splitters and filters. The applications of cladding stress induced effects can be envisioned in a variety of situations to enhance device functionality, simplify fabrication and improve operation tolerance.

ACKNOWLEDGMENT

The authors gratefully acknowledge the contribution of CPFC — Canadian Photonics Fabrication Center (http://cpfc-ccfdp.nrc-cnrc.gc.ca) in fabricating the MMI-based polarization insensitive ring resonators.

REFERENCES

1. Pavesi, L. and D. Lockwood, Eds. *Silicon Photonics*, Springer Verlag, Berlin, 2004; Reed, G. T. and A. P. Knights, *Silicon Photonics, An Introduction*, John Wiley & Sons, England, 2004.
2. Vivien, L., S. Laval, B. Dumont, S. Lardenois, A. Koster, and E. Cassan, "Polarization-independent single-mode rib waveguides on silicon on insulator for telecommunications wavelengths," *Opt. Commun.*, Vol. 210, 43–49, 2002.
3. Dai, D. and S. He, "Analysis of the birefringence of a silicon-on-insulator rib waveguide," *Applied Optics*, Vol. 43, No. 5, 1156–1161, 2004.
4. Reed, G. T., S. P. Chan, W. Headley, V. M. N. Passaro, A. Liu, and M. Paniccia, "Polarisation independent devices in small SOI waveguides," *Proc. IEEE/LEOS Group VI Photonics Conference*, Paper FB5, Hong Kong, CD, 2004.
5. Xu, D.-X., P. Cheben, D. Dalacu, A. Delage, S. Janz, B. Lamontagne, M. Picard, and W. N. Ye, "Eliminating the birefringence in silicon on insulator ridge waveguides using the cladding stress," *Opt. Lett.*, Vol. 29, 2384, 2004.
6. Ye, W. N., D.-X. Xu,, S. Janz, P. Cheben, M.-J. Picard, B. Lamontagne, and N. G. Tarr, "Birefringence control using stress engineering in silicon-on-insulator (SOI) waveguides," *J. Lightwave Technol.*, Vol. 23, No. 3, 1308, 2005.
7. Huang, M., "Stress effects on the performance of optical waveguides," *Inter. J. Solids and Struct.*, Vol. 40, 1615, 2003.
8. Xu, D.-X., W. Ye, A. Bogdanov, P. Cheben, D. Dalacu, S. Janz, B. Lamontagne, M.-J. Picard, and N. G. Tarr, "Stress engineering for the control of birefringence in SOI waveguide components," Photonics West 2005 (invited), *SPIE Proc.*, Vol. 5730, 158–172, 2005.
9. Xu, D.-X., P. Cheben, D. Dalacu, S. Janz, M.-J. Picard, N. G. Tarr, and W. N. Ye, "Control and compensation of birefringence in SOI waveguides," *Proc. 16th LEOS*, WM4, 590, Tucson, Arizona, October 26–30, 2003.
10. Cheben, P., A . Bogdanov, A . Delage, S. Janz, B. Lamontagne, M.-J. Picard, E. Post, and D.-X. Xu, "A 100-channel near-infrared SOI waveguide microspectrometer: Design and fabrication challenges," *SPIE*, Vol. 5644, 103–109, 2004.
11. Xu, D.-X., S. Janz, and P. Cheben, "Design of polarization-insensitive ring resonators in silicon-on-insulator using MMI couplers and cladding stress engineering," *Photon. Technol. Lett.*, Vol. 18, No. 2, 343, 2006.
12. Ye, W. N., D.-X. Xu, S. Janz, P. Waldron, and N. G. Tarr, "Stress-induced SOI polarization splitter based on Mach-Zehnder interferometers (MZI)," *Proc. 3rd IEEE/LEOS Group VI Photonics Conference*, Paper FC4, Ottawa, Canada, CD, 2006.

Analysis of Crosstalk between Single-ended and Differential Lines

F. Xiao[1], **R. Hashimoto**[1], **K. Murano**[2], **and Y. Kami**[1]

[1]University of Electro-Communications, Japan
[2]Tokai University, Japan

Abstract— The crosstalk between single-ended and differential lines is investigated in this paper. First, the telegraphers equations for multiconductor line are applied. Then the differential pair is terminated respectively with T and Π termination networks, and the crosstalk from the single-ended trace to the differential pair is evaluated at the grounded resistors in the T or Π termination networks. The analysis of the crosstalk is obtained by incorporating the termination network with the solution of the telegraphers equations that are solved by using a mode decomposition technique.

1. INTRODUCTION

As digital interfaces are changing from massive parallel buses to gigabit serial data links, signal integrity issues, such as crosstalk between signals, have become more of a problem in the interconnection of single-ended signals. The analysis of crosstalk between two and multi single-ended transmission lines, both in time-domain and frequency-domain, has been reported extensively in the literature [1–5]. Comparing with the single-ended circuits, crosstalk between tightly coupled differential circuits are small [6]. The direct measurements of crosstalk between differential circuits was discussed in [7], and in [8] the analysis of crosstalk between differential delay lines was reported.

In this paper, the crosstalk between a single-ended trace on the printed circuit board (PCB) to an adjacent differential pair is discussed. In our analysis, the single-ended trace with the differential pair are considered as three parallel lines thus the telegrapher's equations for multiconductor line can be applied. Then the differential pair is terminated respectively with T and Π termination networks, and the crosstalk from the single-ended trace to the differential pair is evaluated at the grounded resistors in the T or Π termination networks. The analysis of this time-domain crosstalk is obtained by incorporating the termination network with the solution of the telegrapher's equations that are solved by using a mode decomposition technique and using the fast inverse Laplace transform. The measurement of this crosstalk can be conducted easily by using a two-port network analyzer or a digital oscilloscope since the crosstalk is evaluated on the resistors that are connected to ground. Both the calculated and measured time domain crosstalk are presented in this paper. The characteristics of the near-end crosstalk (NEXT) and far-end crosstalk (FEXT) are also examined and discussed.

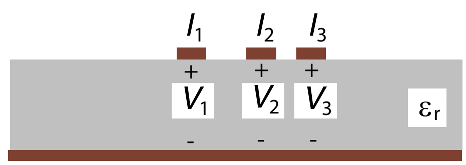


Figure 1: Schematic of three conductor lines above ground plane and definition of line voltages and currents.

2. NETWORK EXPRESSION FOR (3+1)-CONDUCTOR LINES

A single-ended trace runs parallel with a differential pair traces on a PCB is shown in Fig. 1. The line voltages V_i and line currents $I_i\,(i = 1, 2, 3)$ for each trace are also shown in the figure. At the first step, we just consider the traces as three conductor transmission lines above a reference conductor, i. e., (3+1)-conductor lines. If the traces are x-oriented, then the telegrapher's equations in the Laplace domain are

$$\frac{\partial \boldsymbol{V}}{\partial x} = -s\boldsymbol{L}\boldsymbol{I}, \quad \frac{\partial \boldsymbol{I}}{\partial x} = -sc\,\boldsymbol{V} \tag{1}$$

where $\boldsymbol{V} = [V_1, V_2, V_3]^T$ and $\boldsymbol{I} = [I_1, I_2, I_3]^T$ are respectively the line voltage and current vectors with the superscript T denoting the transpose of the vector, $\boldsymbol{L}$ and c are the per-unit-length inductance matrix and capacitance matrix, respectively. Both $\boldsymbol{L}$ and c are symmetric matrices. The equations in (1) are a set of six, coupled, first-order partial differential equations. They can be solved using the mode decomposition technique [2]. This technique decomposes the line voltages and currents into three independent modes, then the six coupled telegrapher's equation (1) are transformed to six uncoupled equations. The primary solution process of the mode decomposition technique is to find transformation matrices $\boldsymbol{T}_v$ and $\boldsymbol{T}_i$, which change the actual line voltages and currents, $\boldsymbol{V}$ and $\boldsymbol{I}$, to the mode voltages and currents $\boldsymbol{V}_m$ and $\boldsymbol{I}_m$, i. e.,

$$\boldsymbol{V} = \boldsymbol{T}_v \boldsymbol{V}_m, \quad \boldsymbol{T} = \boldsymbol{T}_i \boldsymbol{I}_m \tag{2}$$

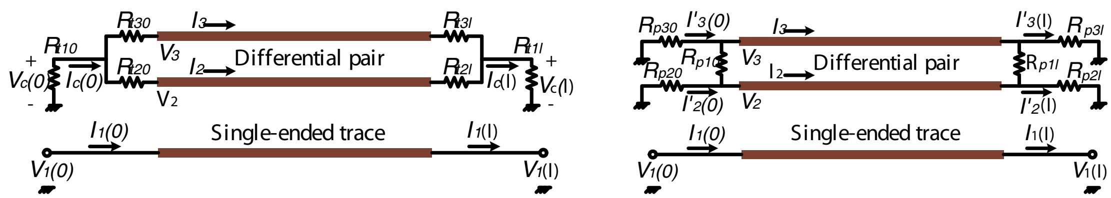
The solution of the six uncoupled equations can be easily obtained in the same fashion of solving a bifilar transmission line. For example, without loss of generality, for lines of length ℓ, we can express the network expression between the mode voltages and currents in chain matrix as

$$\begin{bmatrix} \boldsymbol{V}_m(0) \\ \boldsymbol{I}_m(0) \end{bmatrix} = \begin{bmatrix} \boldsymbol{A}_m & \boldsymbol{B}_m \\ \boldsymbol{C}_m & \boldsymbol{D}_m \end{bmatrix} \begin{bmatrix} \boldsymbol{V}_m(\ell) \\ \boldsymbol{I}_m(\ell) \end{bmatrix} \tag{3}$$

where the matrices $\boldsymbol{A}_m$, $\boldsymbol{B}_m$, $\boldsymbol{C}_m$, and $\boldsymbol{D}_m$ can be easily calculated from the mode impedances, phase constants, and the line length. Thus for the actual line voltages and currents we have

$$\begin{bmatrix} \boldsymbol{V}(0) \\ \boldsymbol{I}(0) \end{bmatrix} = \begin{bmatrix} \boldsymbol{T}_v & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{T}_i \end{bmatrix} \begin{bmatrix} \boldsymbol{A}_m & \boldsymbol{B}_m \\ \boldsymbol{C}_m & \boldsymbol{D}_m \end{bmatrix} \begin{bmatrix} \boldsymbol{T}_v^{-1} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{T}_i^{-1} \end{bmatrix} \begin{bmatrix} \boldsymbol{V}(\ell) \\ \boldsymbol{I}(\ell) \end{bmatrix} \equiv \begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{bmatrix} \begin{bmatrix} \boldsymbol{V}(\ell) \\ \boldsymbol{I}(\ell) \end{bmatrix} \tag{4}$$

This is the obtained chain matrix expression for the $(3+1)$-conductor transmission line system.



(a) (b)

Figure 2: Differential pair terminated by (a) T-type and (b) Π-type networks.

3. INCORPORATING THE T AND Π TERMINAL CONDITIONS

So far we have not care whether the traces are single-ended or differential. Now we will back to the crosstalk problem between a singled-ended and a differential lines. Without loss of generality, we assume trace 1 is single-ended, and traces 2 and 3 composing a differential pair. We terminate the differential pair respectively with T and Π termination networks. Our purpose to use these termination schemes are twofold. One is to extract the common- and differential-mode components of the crosstalk, and the other is to measure the crosstalk using the common single-ended network analyzer or oscilloscope. Considering the crosstalk voltage at the same cross section on the differential pair, V_2 and V_3, we may decompose each of them as a linear combination of two components. The differential-mode voltages $V_{DM}/2$ are equal in magnitude and opposite in polarity, and the common-mode voltages, V_{CM} are equal in magnitude and with same polarity, i. e.,

$$V_2 = V_{CM} + V_{DM}/2, \quad V_3 = V_{CM} - V_{DM}/2 \tag{5}$$

The differential pairs terminated by the T and Π networks are shown in Fig. 2. In the T termination configuration, if we choose the resistors as $R_{t2k} = R_{t3k}$, with $k(=0, \ell)$ denoting the input and output ends, then the voltages on resistors R_{t10} and $R_{t1\ell}$, denoted as $V_C(0)$ and $V_C(\ell)$, will be

$$V_C(k) = \frac{R_{t1k}}{2R_{t1k} + R_{t2k}}[V_2(k) + V_3(k)] = \frac{R_{t1k}}{2R_{t1k} + R_{t2k}} 2V_{CM}(k), \quad k = 0, \ell. \tag{6}$$

This means the voltage on the R_{t10} and R_{t10} can be used to evaluated the common-mode components of the near- and far-end crosstalk in the T termination case, with a scale factor related to the termination resistors. As will be explained later, in the measurements, we have chosen

$R_{t2k} = R_{t3k} = 0$, thus from (6), one can see $V_C^{(k)}$ are be simply equal to $V_{CM}^{(k)}$. In the Π termination configuration, we choose the resistors $R_{p2k} = R_{p3k}$ and evaluate the voltages across them. Since all of these resistor are grounded, thus the conventional single-ended equipment can be used for the measurement. Then from (5), we can see the differences between the voltages on the two resistors R_{p2k} and R_{p3k} are simply equal to $V_{DM}^{(k)}$, thus they can be used to evaluated the differential-mode components of the NEXT and FEXT in the Π termination case. From the above discussion, one can see only the voltages on the grounded resistors in the termination networks will be directly used, thus now we will incorporate the non-grounded resistors into the network expression obtained in Section 2, and derived a new network expression for characterizing the crosstalk.

3.1. T Termination Network

In the T termination case, we express the chain matrix expression (4) as the admittance matrix

$$\begin{bmatrix} \boldsymbol{I}(0) \\ -\boldsymbol{I}(\ell) \end{bmatrix} = \boldsymbol{Y} \begin{bmatrix} \boldsymbol{V}(0) \\ \boldsymbol{V}(\ell) \end{bmatrix} = \begin{bmatrix} \boldsymbol{Y}_1 & \boldsymbol{Y}_2 \\ \boldsymbol{Y}_3 & \boldsymbol{Y}_4 \end{bmatrix} \begin{bmatrix} \boldsymbol{V}(0) \\ \boldsymbol{V}(\ell) \end{bmatrix} \tag{7}$$

The admittance matrix can be obtained from the chain matrix. Considering the relationship between the line voltages and currents with the voltages and currents at the T termination resistor, the T termination condition can be expressed as

$$\begin{bmatrix} \boldsymbol{V}(0) \\ \boldsymbol{V}(\ell) \end{bmatrix} = - \begin{bmatrix} \boldsymbol{R}_{t0} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{R}_{t\ell} \end{bmatrix} \begin{bmatrix} \boldsymbol{I}(0) \\ -\boldsymbol{I}(\ell) \end{bmatrix} + [V_1(0),\, V_C(0),\, V_C(0),\, V_1(\ell),\, V_C(\ell),\, V_C(\ell)]^T \tag{8}$$

where $\boldsymbol{R}_{t0}$ and $\boldsymbol{R}_{t\ell}$ are the termination resistor matrices at the input and output. Let $\boldsymbol{V}'(k) = [V_1(k),\, V_c(k),\, V_c(k)]^T$, we have

$$\begin{bmatrix} \boldsymbol{I}(0) \\ -\boldsymbol{I}(\ell) \end{bmatrix} = \begin{bmatrix} \boldsymbol{U} + \boldsymbol{Y}_1\boldsymbol{R}_{t0} & \boldsymbol{Y}_2\boldsymbol{R}_{t\ell} \\ \boldsymbol{Y}_3\boldsymbol{R}_{t0} & \boldsymbol{U} + \boldsymbol{Y}_4\boldsymbol{R}_{t\ell} \end{bmatrix}^{-1} \boldsymbol{Y} \begin{bmatrix} \boldsymbol{V}'(0) \\ \boldsymbol{V}'(\ell) \end{bmatrix} \tag{9}$$

Considering $I_2(k) + I_3(k) = I_c(k)$, and denoting $\boldsymbol{I}_T(k) = [I_1(k),\, I_c(k)]^T$, $\boldsymbol{V}_T(k) = [V_1(k),\, V_c(k)]^T$, we can get

$$\begin{bmatrix} \boldsymbol{I}_T(0) \\ -\boldsymbol{I}_T(\ell) \end{bmatrix} = \boldsymbol{Y}_T \begin{bmatrix} \boldsymbol{V}_t(0) \\ \boldsymbol{V}_t(\ell) \end{bmatrix} \tag{10}$$

where $\boldsymbol{Y}_T$ is the obtained new admittance matrix by incorporating the T termination condition. Considering the measurement system are of characteristics impedance of 50 Ohm, the scattering matrix can be obtained from this admittance matrix, with which the crosstalk can be analyzed.

3.2. Π Termination Network

Next, we consider the Π termination configuration as shown in Fig. 2(b). The network expression for the Π termination can be obtained in the similar fashion. In this case, instead of the admittance matrix, we use the impedance matrix for convenience, i. e.,

$$\begin{bmatrix} \boldsymbol{V}(0) \\ \boldsymbol{V}(\ell) \end{bmatrix} = \boldsymbol{Z} \begin{bmatrix} \boldsymbol{I}(0) \\ -\boldsymbol{I}(\ell) \end{bmatrix} = \begin{bmatrix} \boldsymbol{Z}_1 & \boldsymbol{Z}_2 \\ \boldsymbol{Z}_3 & \boldsymbol{Z}_4 \end{bmatrix} \begin{bmatrix} \boldsymbol{I}(0) \\ -\boldsymbol{I}(\ell) \end{bmatrix} \tag{11}$$

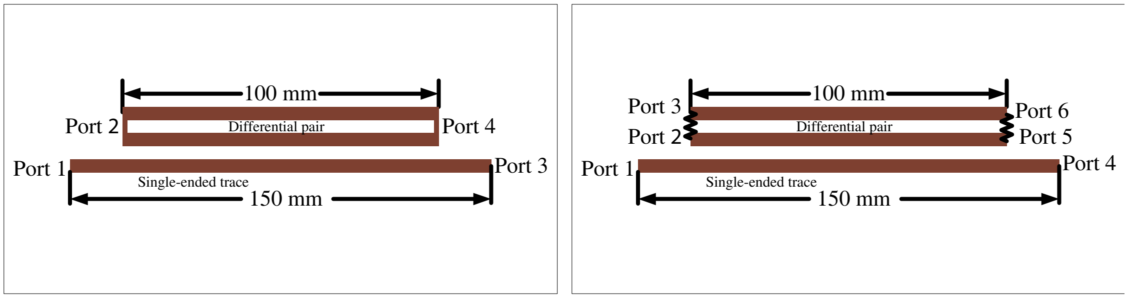
The impedance matrix can be obtained from the chain matrix (4). Let $\boldsymbol{I}_\Pi(k) = [I_1(0),\, I'_2(0),\, I'_3(0)]^T$, then the Π termination condition can be expressed as

$$\begin{bmatrix} \boldsymbol{I}_\Pi(0) \\ \boldsymbol{I}_\Pi(\ell) \end{bmatrix} = \begin{bmatrix} \boldsymbol{I}(0) \\ \boldsymbol{I}(\ell) \end{bmatrix} + \begin{bmatrix} \boldsymbol{G}_{p0} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{G}_{p\ell} \end{bmatrix} \begin{bmatrix} \boldsymbol{V}(0) \\ \boldsymbol{V}(\ell) \end{bmatrix} \tag{12}$$

where G_{p0} and $G_{p\ell}$ are the termination conductance matrices at the input and output. Substituting (11) into (12), we can obtain

$$\begin{bmatrix} \boldsymbol{I}_\Pi(0) \\ -\boldsymbol{I}_\Pi(\ell) \end{bmatrix} = \begin{bmatrix} \boldsymbol{U} + \boldsymbol{G}_{p0}\boldsymbol{Z}_1 & \boldsymbol{G}_{p0}\boldsymbol{Z}_2 \\ -\boldsymbol{G}_{pl}\boldsymbol{Z}_3 & \boldsymbol{U} - \boldsymbol{G}_{p\ell}\boldsymbol{Z}_4 \end{bmatrix} \begin{bmatrix} \boldsymbol{Y}_1 & \boldsymbol{Y}_2 \\ \boldsymbol{Y}_3 & \boldsymbol{Y}_4 \end{bmatrix} \begin{bmatrix} \boldsymbol{V}(0) \\ \boldsymbol{V}(\ell) \end{bmatrix} \equiv \boldsymbol{Y}_\Pi \begin{bmatrix} \boldsymbol{V}(0) \\ \boldsymbol{V}(\ell) \end{bmatrix} \tag{13}$$

Similar to the T termination case, from the above admittance matrix, $\boldsymbol{Y}_\Pi$, the scattering matrix can be derived, with which the crosstalk is analyzed.

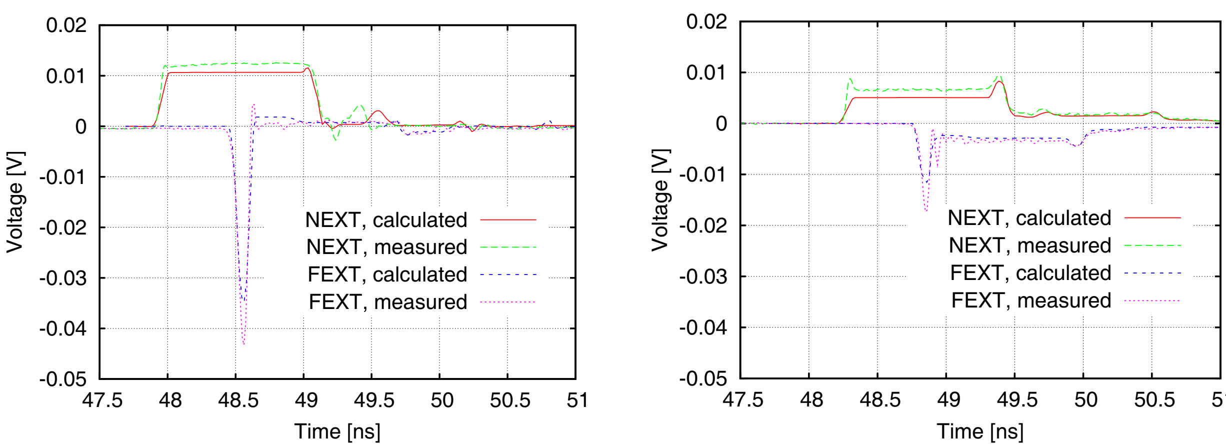


(a) (b)

Figure 3: Differential pair terminated by (a) T-type and (b) Π-type load circuit.

4. EXPERIMENTAL AND ANALYSIS RESULTS

To verify the above analysis, several test circuit boards were fabricated and the near- and far-end crosstalk are measured. The board geometries with the T and Π termination network are shown in Fig. 3. The board thickness is 1.6 mm and relative permittivity is $\varepsilon_r = 4.7$. All traces are with the same width of 2 mm and the edge-to-edge separations between the traces are 1 mm. The length of the differential pair traces are 100 mm. For soldering the SMA launch connectors, the single-ended trace is a little longer, with length of 150 mm. The effect of the extra sections of the single-ended line can be easily accounted by cascading the chain parameter matrices of the sections with the parameter matrix derived in Section 3. A digitizing oscilloscope with a TDR plug-in module was used in the measurements. The source of the scope head was used to provide a step signal to excite the single-ended trace. The step signal had an amplitude of 170 mV and a 10–90% risetime of about 40 ps. To simplify the structure and calculation, the resistors R_{t2k} and R_{t3k} were chosen to be $0\,\Omega$. In the test board, this is realized by simply connecting the two traces composing the differential pair with a conductor strap at both ends of the traces. SMA connectors were connected at the center of the straps, which are corresponding to ports 2 and 4 in Fig. 3(a). In the Π termination configuration, 50-Ohm chip resistors were connected between the two traces. SMA connectors were connected at the ends of each trace, which are corresponding to ports 2, 3, 5 and 6 in Fig. 3(b). For the single-ended trace, two SMA connectors were connected at both ends, which are corresponding to ports 1 and 3 in the T termination configuration, and ports 1 and 4 in the Π configuration. Further more, since the measurement was using a 50-Ohm system, the other resistors in the T and Π termination networks were 50 Ohm. The NEXT and FEXT responses were evaluated at ports 2 and 4 for T termination and ports 2, 3, 5, and 6 for Π termination by using the TDT features of the scope.



(a) (b)

Figure 4: Near- and far-end crosstalk while the differential pair is terminated with (a) T-network and (b) Π-network.

The measured and calculated near- and far-end crosstalks for the T and Π terminations are shown in Fig. 4. One can see the agreement between the measured and calculated results is very good. Some discrepancies do exist between the measurements and calculations, which is because in the calculation an ideal ramp step is used as the input signal, but in the measurement, the input signal has a small overshoot with some fluctuations. One can see also there are several similarities in the pulse shapes and widths for the two termination configurations. For both cases, the NEXT signals have a wide pulse-like shape. The amplitude of the NEXT signals increases from 0 with almost the same risetime as the driving pulse to certain level. After that the amplitudes rise no further. The pulse widths are twice as wide as the propagation time through the differential pair length, which is in fact the length of the coupled region. As for the far-end crosstalk, both cases have sharp negative-going blips that appear almost at the midpoint of the near-end pulse. This is because the far-end crosstalk appear exactly the same time as the driving pulse arriving at the end of the coupled region. On the other hand, the crosstalk for both case do have some differences that are deserved to mention. The amplitude of the crosstalk for the Π termination are obviously small than those for the T termination. Thus we can say the out-of-phase components on the differential pair are more immune to crosstalk from the adjacent single-ended line than the in-phase components do.

5. CONCLUSIONS

The crosstalk between single-ended and differential lines has been analyzed based on the telegrapher's equations for multiconductor lines. The differential pair was terminated respectively with T and Π termination networks, and the crosstalks have been evaluated at the grounded resistors. The analysis is simple, and the conventional single-ended digital oscilloscope or network analyzer can be used in the measurements.

REFERENCES

1. Paul, C. R., *Introduction to Electromagnetic Compatibility*, John Wiley &*Sons*, INC., New York, 1992.
2. Paul, C. R., *Analysis of Multiconductor Transmission Lines*, John Wiley & Sons, INC., New York, 1994.
3. Kami, Y. and R. Sato, "Coupling model of crossing transmission lines," *IEEE Trans. Electromagn. Compat.*, Vol. EMC-28, No. 4, 204–210, Nov. 1986.
4. Xiao, F., W. Liu, and Y. Kami, "Analysis of crosstalk between finite-length microstrip lines: FDTD approach and circuit-concept modeling," *IEEE Trans. Electromagn. Compat.*, Vol. 43, No. 4, 573–578, Nov. 2001.
5. Barmada, S., A. Musolino, and M. Raugi, "Wavelet-based time-domain solution of multiconductor transmission lines with skin and proximity effect," *IEEE Trans. Electromagn. Compat.*, Vol. 47, No. 4, 774–780, Nov. 2005.
6. Johnson, H. W. and M. Graham, *High Speed Digital Design: A Handbook of Black Magic*, NJ: Prentice Hall, 1993.
7. Bockelman, D. E. and W. R. Eisenstadt, "Direct measurement of crosstalk between integrated differential circuits," *IEEE Trans. Microwave Theory Tech.*, Vol. 48, 1410–1413, Aug. 2000.
8. Guo, W., G. Shiue, C. Lin, and R. Wu, "Comparisons between serpentine and flat spiral delay lines on transient reflection/transmission waveforms and eye diagrams," *IEEE Trans. Microwave Theory Tech.*, Vol. 54, 1379–1387, June 2006.
9. Makino, O., F. Xiao, and Y. kami, "Characteristics of electrically long two-conductor lines with inhomogeneous media," *IEICE Trans. Commun.*, Vol. E88-B, No. 7, 3028–3035, July 2005.

A Study of the Confirming Method on 20H Rule in High Speed PCB Design

Wei Chen, **Qiuyuan Huang**, **Peng Zhou**, **and Bin Wang**
School of Information Engineering, Wuhan University of Technology, Wuhan 430070, China

Abstract— As the operating frequency goes higher, the PCB designers have to overcome the ever-increasing EMI (electromagnetic interference) problems. 20H rule is a unproved experiential method used by the PCB designers which can reduce EMI in PCB design. In this paper, a practice measuring method is presented based on the analyses of the reflective coefficient theory and FDTD. With this method, the 20H rule can be confirmed and commonly implemented.

1. INTRODUCTION

The 20H rule was first presented by W. Michael King. In 1980, the 20H rule was recognized and used to reduce EMI. In 1996, the rule was embodied in a book [1] by Mr. Mark I Montrose. The 20H rule is about the relation of the power plane and the ground plane, but is only a still unproved rule established by usage. The different application of the 20H rule results different effects in PCB designing works, as it may reduce EMI or may increase EMI in different circumstances. So Confirming the 20H rule and studying its application circumstances has a practice importance.

2. 20H RULE

20H rule: If there are high-speed currents on the board, there are electromagnetic fields associated with them. At the edge of the planes (presumably at the edge of the board) these fields will fringe outward from the board as shown in Figure 1. If the ground plane is larger than the power plane, the energy can not radiate out, as we can see from Figure 1(b). Thus, outward EMI radiation is reduced and there is less chance for an external EMI problem.

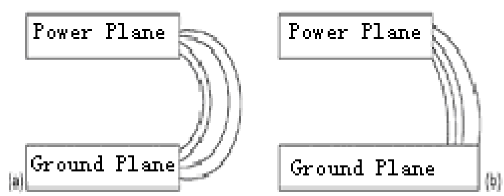


Figure 1: Recessing power plane may reduce outward fringing.

The reduction of edge radiation relates to the plane edge retraction. 20H role (H is the height between the power plane and the ground plane) take it that if the power plane edge is retracted 20H of the ground plane edge, 70% of the energy radiation can be repressed.

20H rule is experiential method without theoretically confirmation and demonstration. Some designers believe that the 20H rule will cause more radiation than not applying it [2, 3]. There also exist a saying that whether the 20H role is applicable depends on the dimension of the PCB and the frequency and the layers separation [4]. Others think that using discrete components instead of implementing the 20-H rule is another choice [5].

3. CONFIRMING METHOD

There is no such a method which can be used to accurately analyze 20H rule. In general, reflective coefficient, FDTD and antenna theories are used to simple analyze 20H rule.

The method of radiating coefficient is a way that analyzes the radiation through the PCB edge's reflective coefficient. FDTD has the advantage of analyzing the electromagnetic phenomenon, so it can be used to emulate the various conditions such as PCB with different dimension, or with different layer separation; different frequency; nH etc. using FDTD we can compare the result between the one using 20H rule and not using it that would give us more information about the 20H rule's application.

3.1. The Method of Reflective Coefficient

Building two 2-layers models (One use 20H rule, and another is normal). It can be seen in Figure 2.

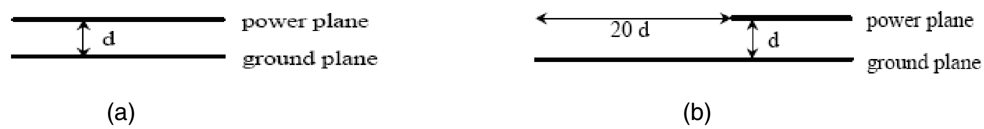


Figure 2: A model consist of two layers. (a) Two planes are of the same size, (b) 20H rule.

The radiation firmly relate to the reflection coefficient of the two layers' edge.

Take the plane thickness as 0, the edge reflection coefficient of the two planes [6]

$$R = |R|e^{-j\theta} \tag{1}$$

Amplitude

$$|R| = e^{-\pi q} \tag{2}$$

Phase:

$$\theta = 2q\left[1 - C + \ln\frac{2}{q} - \left(\frac{\sin^{-1} q}{q} - 1\right) - \sum_{m=1}^{x} A_{2m+1}\left(S_{2m+1} - 1\right) q^{2m}\right] \tag{3}$$

As the charge $q = CU$, C is the function of the distance between the two plates. So the amplitude and phase of the reflection coefficient is the function of the distance between the two plates.

The structure of 20H rule can be estimated by the power plane and the mirror plane in Figure 3. Because the distance between the power plane and the mirror plane is 2 times longer than the distance in Figure 2(a), the reflection coefficient in Figure 2(b) is small than in Figure 2(a). Therefore, more radiation is expected to come out of the edges of the board implemented with 20-H rule.

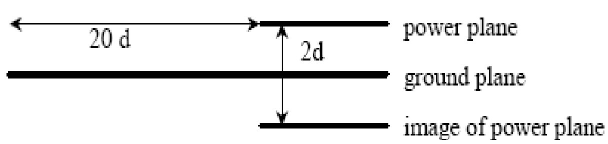


Figure 3: Image of the power plane in 20-H rule structure.

As we can understand that the model of this method is too simple and idealized; many practice factors have not been taken into account. So the conclusion is not scientific.

3.2. Finite Difference Time Domain Method (FDTD)

For researching 20H rule deeply, we can use more advanced method-Finite difference time domain method (FDTD) to analyze the radiations of applying 20H rule or not applying it. The focus is on the amplitude and distribution of the radiating field.

The following is how to build a model for analyzing and confirm the 20H rule by FDTD.

(1) The emulating model and excitation source of 2D plain structure.

For simplicity, we chose a 2D model to simulate the signal propagation between power plane and ground plane.

The PCB model has a power plane and a ground plane as shown in Figure 4. For simulating 20H rule, the following assumptions are made: 1. the power and ground planes have a good conductivity, satisfying the PEC condition; 2. both of two planes are infinitely thin in thickness; 3. the simulation is based on infinite space. In the following discussions, H stands for the distance between the layers of the PCB; E stands for electronic field.

As the power and ground planes satisfy the PEC condition (ideal conductivity), so two planes' tangential electric field should be zero. The blue mark in the figure stands for dipole excitation source.

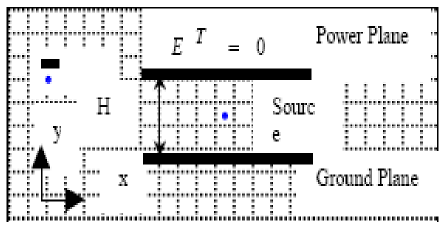


Figure 4: 2D planar structures.

As the limitation of the computer resources, we reduce the infinite space to a finite rectangular space. There is a appropriate matching layer fitted to the area edge. Taking the above condition into account, we build the simplest model.

They are three kinds of excitation sources which can be used in simulation. They are dipole excitation source, Uniform voltage excitation source and Gaussian pulse.

Here only introduce the dipole excitation source, usually in the center of structure, by enforcing the E field at this point, shown in the following equation:

$$E_{\text{center}} = E_0 \sin(2\pi f k dt) \tag{4}$$

where f = the frequency of the excitation source, k = time step number, dt = length of time step.

Different from the uniform voltage excitation source, the dipole source will not change with layers separation distance. This characteristic has an advantage in studying the 20H rule with different layer separation distance.

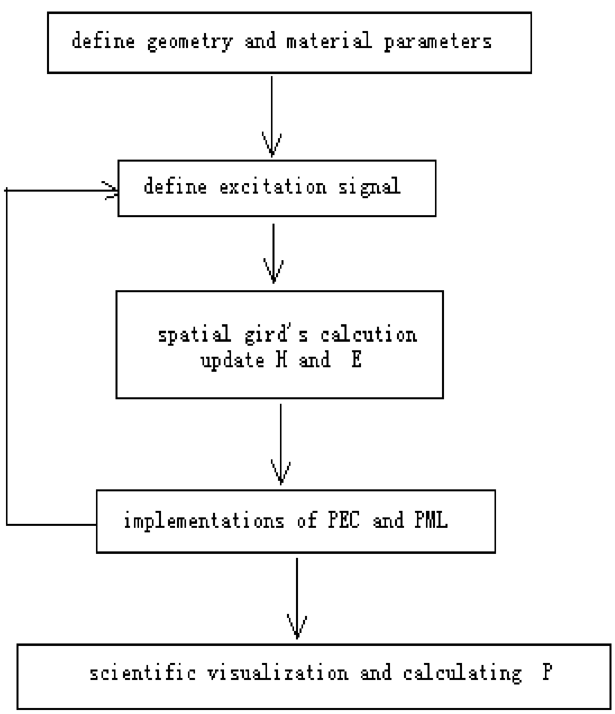


Figure 5: Flow chart of FDTD Method.

(2) FDTD Method

Figure 5 present a flow chart of FDTD algorithm. The first step of for the FDTD simulation of a microwave structure is to define geometry structure and material parameters. That is: the interface of the dielectric and the conduction interface are specified on the computational grid in the step. The second step is to establish excitation signal. When the excitation source is made sure, the FDTD method should update all the field components on the computational grid for each step. Using the discrete form of Maxwell equation, the renewed equation can be modified. As the time step length satisfy the stable condition, so it can be reduced to:

$$\Delta t = \frac{\Delta x}{2C_0} \tag{5}$$

As a good conduction layer on the interface of the dielectric, the tangent part of the electric field is set to zero. With perfectly matched layers (PML) is used as the absorbing boundary conditions, the tangent part of the field on the outer layers of the computational grids are updated.

(3) Energy Expression: *Poynting Vector*

For the study of the 20H rule effect, the PCB edge should be quantized. P represents the energy of the edge radiation. P vector stands for the density of electromagnetic energy which can be expressed as: $\overrightarrow{P} = \overrightarrow{E} \times \overrightarrow{H}$($\overrightarrow{E}$, $\overrightarrow{H}$ stand for electric field and magnetic field intensity).

As the E and H are all instantaneous vectors, so P is instantaneous vector too. If we integrate P along a definite surface, it just results the energy out flown from a closed surface. This can also be applicable in 2D simulation. The only thing should be changed is along a line to integrate but not a closed surface, that is along the edge AB, BC, CD and DA as shown in Figure 6.

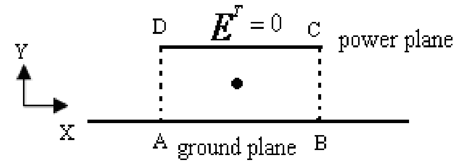


Figure 6: Integration route for the P.

The three components of the P:

$$P_x = E_y H_z - E_z H_y \tag{6}$$
$$P_y = E_z H_x - E_x H_z \tag{7}$$
$$P_z = E_x H_y - E_y H_z \tag{8}$$
$$\oint_{ABCD} P \cdot dl = -\int_{AB} P_y dl + \int_{BC} P_x dl + \int_{CD} P_y dl - \int_{DA} P_x dl \tag{9}$$

In time domain, for $\overrightarrow{P}$ vector, we can use different expression to represent different mode of the waves along this line (TE, or TM).

As the TE mode only has E_x, E_y and H_z along AB and CD, $E_x = 0$, so $P_y = 0$. Using Cartesian discretization, from expression 9 we can get:

$$\oint_{ABCD} \overrightarrow{P} \cdot dl = -\int_{DA} P_x dl + \int_{BC} P_x dl \tag{10}$$

For the same reason, TE mode only has E_z, H_x and H_y. Along AB and CD, we know $E_x = 0$, so $P_y = 0$. For TM mode, we can get the same result.

As the symmetric structure, the integral along the line can be simplified as:

$$\oint_{ABCD} P \cdot dl = 2\int_{BC} P_x dl \tag{11}$$

3.3. Practice Measuring Method

The theoretical analyze may differ from simulation. So we can use a real measure device to measure the edge radiation. The measure device is shown in Figure 7.

The results of measuring and simulation may also be not equal. This is because the real measure system is much complicated than simple model. On the other hand, the antenna, cable, measuring circumstances and the definition of the measuring devices and the near field effect all can be a factor affecting the measuring result. The radiation from the connector of excitation source may be greater than the edge radiation which make the measuring error larger than usual. All these factors can bring tremendous difficulties in the certification of 20H rules.

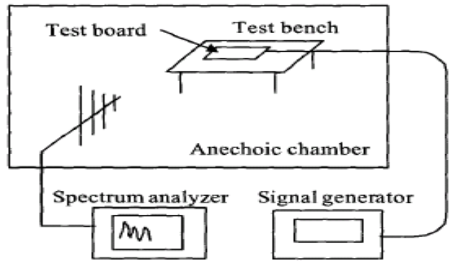


Figure 7: Measuring equipment.

4. CONCLUSIONS

As the above analyses are all based on bare PCB, unlike real PCB which has a lot dynamic digital components on, the real analyses can be very complex. Though we have not got the 20H rule's working mechanism, one thing is certain that he use of 20H rules resolved many EMC issues of PCB design.

Among the Confirming methods of 20H rule, the reflective coefficient is not a proper method for its limitation and incomprehensive. FDTD is an effective method. There are three major factors which will greatly affect the result when using FDTD. They are: operation frequency; the dimension and layer's separation of PCB. In addition, layer number, component's layout and layers setup would all be factors affecting the 20H rule's use. We need to build more precise model based on the above conclusion. Only more effective algorithm and program is promoted, can we analyze the 20H rule more efficiently.

ACKNOWLEDGMENT

This project is sponsored by Natural scientific fund, Hubei province (2004ABA045), and the authors would like to thanks Stephen H. Hall for the guide in the filed.

REFERENCES

1. Montrose, M. I., "Printed circuit board design techniques for EMC compliance," *IEEE Press*, 26–28, 1996.
2. Chen, H. and J. Fang, "Effects of 20-H rule and shielding vias on electromagnetic radiation from printed circuit boards," *Electrical Performance of Electronic Packaging, IEEE Conference*, 193–196, 23–25 October, 2000.
3. Shin, H. W. and T. Hubing, "20-H rule modeling and measurements," *IEEE International Symposium*, Vol. 2, 939–942, 13–17 August, 2001.
4. Yi, J., L.-W. Li, and E.-P. Li, "Design and analysis of printed circuit boards using FDTD method for the 20-H rule," 2002.
5. Montrose, M., E. Liu, and E.-P. Li, "Analysis on the effectiveness of printed circuit board edge termination using discrete components instead of implementing the 20-H rule," *International Symposium*, Vol. 1, 45–50, 9–13 August, 2004.

De-embedding Techniques for Passive Components Implemented on a 0.25 μm Digital CMOS Process

Marc D. Rosales, **Honee Lyn Tan**, **Louis P. Alarcon**, **and Delfin Jay Sabido IX**
Department of Electrical and Electronics Engineering
University of the Philippines, Diliman, Quezon Ctiy 1101, Philippines

Abstract— On wafer measurement and characterization of passive components implemented on CMOS technology is one of the initial activities in implementing RF circuits on CMOS. Using test fixtures is necessary in order to test inductors and capacitors at GHz frequencies. However, these fixtures introduce significant effects on measured parameters.

This study focuses on the OPEN and THREE step de-embedding techniques used for on wafer measurement of passive devices. Different values of inductors and capacitors were used as devices under test (DUT). For each component type, a set of structures for the two de-embedding were fabricated. All of these structures were fabricated on a 0.25 μm Digital CMOS process.

1. INTRODUCTION

Test fixtures are used to facilitate the on-chip probing and measurement of passive components implemented in CMOS devices. However, these test fixtures have significant effects on the measured parameters of the devices. These effects can be cancelled out by using different de-embedding techniques. These techniques will use several de-embedding structures and offline post processing of the measured data.

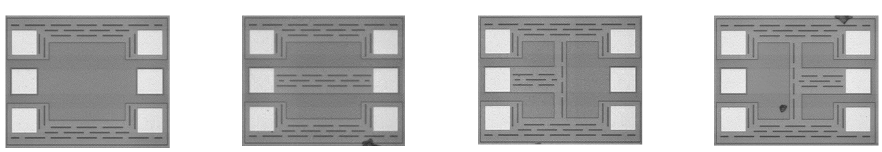
A typical on wafer characterization follows these series of activities: (a) preparing the measurement setup, (b) calibration, (c) on-chip probing and measurement, (d) de-embedding and (e) parameter extraction.

2. IMPLEMENTED DEVICES

To facilitate the study, inductors [1] and capacitors [2] are implemented on a 0.25 μm Digital CMOS process. Two inductor structures are used for this study: a square 1.8nH inductor and an octagonal 3.0 nH inductor. Two capacitor structures are also used: a 635 fF horizontal parallel plate capacitor (HPP) and a 584 fF woven capacitor structure.

In order to measure the devices under test, a Ground Signal Ground (GSG) fixture is used. This serves as an interface to the device port and probe tips. Different fixtures are used for each type of device under test. The GSG probe tip arrangement is advised for measurement beyond 18 GHz, and required beyond 26 GHz [3]. Some guidelines are followed in the design of fixtures [4].

Figure 1(a) shows the actual GSG fixture used to hold the inductors. The upper ground pads of the fixtures are joined with metal strips from metal layers 5, 4, 3, 2 and 1. The metal strips are joined together by dense via array. The ground structure provided the contact to ground of the substrate from the top. The same arrangement is followed in the lower ground pads. The signal pads are implemented by stacking metal layer 5, 4, 3, 2 and 1 together with the use of dense via array. Another GSG fixture is fabricated to facilitate the on-wafer measurement of the capacitor structures.



(a) (b) (c) (d)

Figure 1: (a) Open (b) Thru (c) Short 1 (d) Short 2.

3. MEASUREMENT SETUP

A typical setup for doing on-wafer measurements for RF application consists of several equipments: an IC probe station, Probe link arms, GSG probes, semi- rigid cables and Vector Network Analyzer.

The laboratory setup consists of an Agilent 8753ES Vector Network Analyzer (30 KHz–6 GHz), Micromanipulator IC Probing Station, Microwave Probe link arms, GSG Picoprobes 40AGSG-160P with 160 microns pitch. The silicon substrate is also grounded from the back side through the testing chuck.

The network analyzer is configured to perform broadband passive component measurements. The power setting is set to 0 dBm and the sweep frequency is set from 30KHz up to 6 GHz. A total of 201 points are used to represent the S-parameters of the whole sweep range. A Touchtone file format in real and imaginary components is chosen to be the output format of the network analyzer.

The measurement setup is calibrated using a GGB industries CS-5 calibration substrate. A Short-Open-Load-Thru (SOLT) technique is used for the measurement calibration.

4. DE-EMBEDDING

De-embedding is a process of eliminating the influence of the transition region between the probe, probe contact and the device under test. It is mathematically subtracting networks from the measured results. A pre-requisite for a correct deembedding is that certain test structures are available on a wafer together with the DUT itself. Depending on the selected de-embedding method, open short and thru structures are required and must be measured. Figure 1 shows the de-embedding structures used for the inductor measurements. Another set of de-embedding structures is also fabricated for the capacitor measurements.

4.1. Open De-embedding

Open de-embedding (OPD) is the most widely used because of its simplicity. Only one set of open pads are required to perform de-embedding. This method assumes that the parasitics leading to the DUT are parallel admittance y_p [4] .

It is performed by subtracting the Y-Parameter of the open fixture only to the Y-Parameters of the DUT inside the fixture to obtain the Y-parameters of the DUT only.

De-embedding the fixture from the DUT inside the fixture was performed by subtracting the admittance y_p from the admittance of the DUT attached to the fixture. The Y-parameters of the open fixture is estimated by the admittance y_p. (1) is performed for every frequency point.

4.2. Three Step De-embedding

Three step de-embedding (TRISD) offers a more general form of de-embedding and calibration. It uses a more complicated model topology for the test fixtures. All parasitics associated with probe pads and interconnect-metal lines can be represented and can be subtracted out of the measurement [5]. This requires the use of an open, thru, short 1, and short 2 structures as can be seen in Figures 1. Three-step de-embedding includes the following operations: (a) Subtract the on-wafer Y-parameters shunting the input and output ports. (b) Subtract the Z-parameters taken from the thru and shorts. (c) Subtract the coupling capacitance between the input and output lines. A detailed discussion of the computation of OPD and TRISD is in [5].

5. INDUCTOR MEASUREMENTS

The inductors inside the fixture are measured following the procedures discussed in the earlier sections. Figure 2 shows the measured S-parameters of the square and octagonal inductors together with the effects of the test fixture. The plot shows how the fixture influences the measured $S11$ of the inductors. The S-parameters of all the de-embedding structures for the inductor measurement are also measured. The measured S-parameters are processed offline using de-embedding functions written using MATLAB.

Figure 3 shows the open de-embedded and three step de-embedded S-parameters of the square inductor. The S-parameter plots of both techniques are similar. In the octagonal inductor, it was also observed that the results of both de-embedding techniques are similar.

Parameter extraction is the last stage in the postprocessing of the measured data. Using the deembedded S-parameters, a simple pi model of the inductor can be used to extract its inductance and quality factor.

$$L_s = \frac{\mathrm{Im}\left(\frac{1}{-y_{12de-\text{embeded}(\omega)}}\right)}{\omega} \tag{1}$$

$$Q = \frac{-\mathrm{Im}(y_{11de-\text{embeded}(\omega)})}{\mathrm{Re}(y_{11de-\text{embeded}(\omega)})} \tag{2}$$

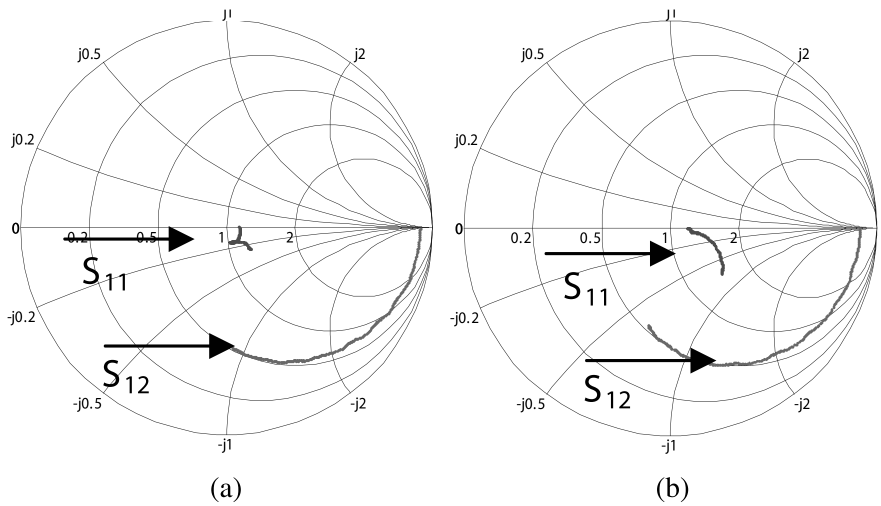


Figure 2: Measured S-Parameters with GSG Fixture (a) square inductor (b) octagonal inductor.

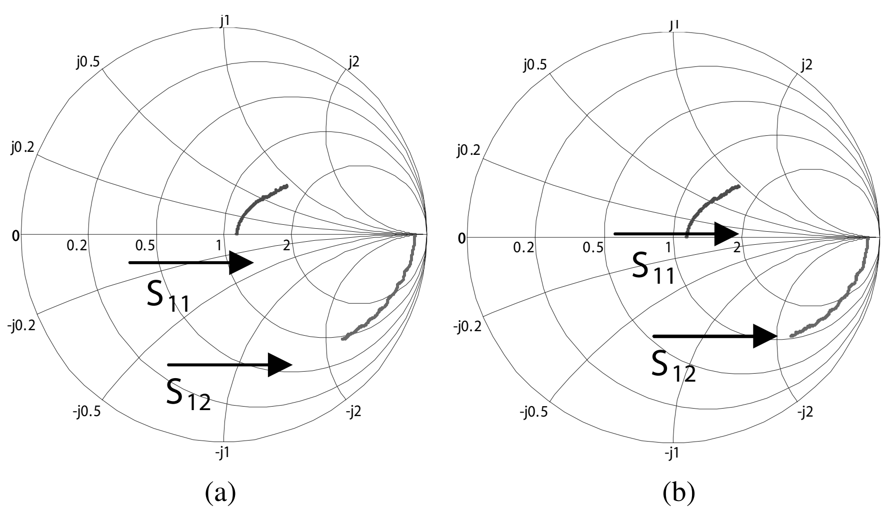


Figure 3: De-embedded S-parameters of the square inductor (a) OPD (b) TRISD.

The square data points are the extracted Q from an open de-embedded measurement while the circle data points are the extracted Q from a three-step deembedded measurement. Initial

Table 1: OPD VS. TRISD

INDUCTOR	**OPD**		**TRISD**	
	PEAK	FREQ (GHZ)	PEAK	FREQ (GHZ)
square	3.15	5.16	3.15	5.16
octagon	3.33	3.96	3.33	3.96

inspection shows that the extracted Q from both de-embedding methods have similar values. The extracted inductance of the three-step de-embedded measurements exhibits a slightly higher value at higher frequencies. The same trend is also observed for the octagonal inductor.

Table 1 shows a summary of the maximum Q of the inductors and the frequency it occurs. The values show the de-embedding technique used does not affect the extracted peak inductor Q.

6. CAPACITOR MEASUREMENTS

The S-parameters of the woven capacitor are illustrated in Figure 4. The CADENCE plots of the $S11$ and $S12$ parameters are the simulation results acquired from the capacitor without the test fixture. The 'With Fixture' plots represent the measured $S11$ and $S12$ parameters before de-embedding. The deembedded $S12$ resulting from the TRISD procedure follows the Cadence simulator's $S12$ plot, while the OPD $S12$ has a small deviation from the simulation result. Both the $S11$ results of the OPD and the TRISD techniques deviate from that of the Cadence's simulation result. The deviation becomes significant as the frequency increases. Nevertheless, it is seen that de-embedding techniques can eliminate the effect of the fixture in the measured S-parameters and approximate the measurement to the S-parameters from the Cadence simulation. The rest of the S-parameter plots of the implemented capacitors follow the same trend the extracted capacitance of the selected tests is obtained from $\Im m[1/\text{-}Y12]$ and plotted with respect to frequency. The effect of the two de-embedding techniques in the extraction of the capacitances is observed. The capacitances at 1.02 GHz and 2.4 GHz are extracted for each de-embedding technique as presented in Figure 5 and the resulting capacitance plots obtained from OPD and TRISD showed that the difference is minimal for the small valued capacitors.

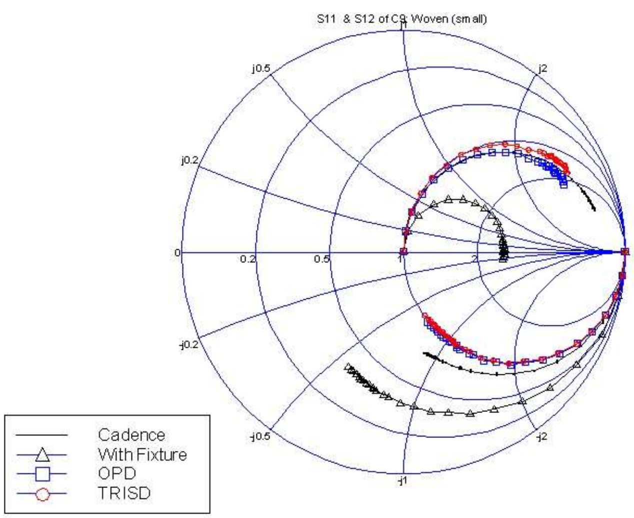


Figure 4: S-parameters of the woven capacitor.

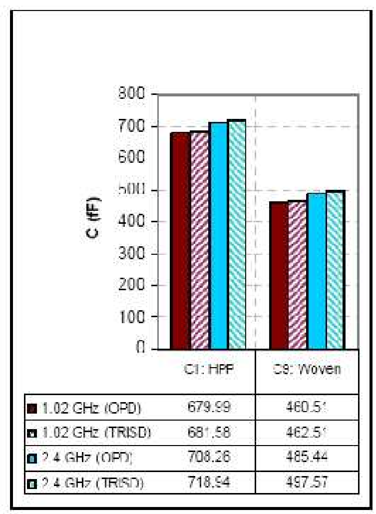


Figure 5: Capacitances at 1.02 GHz and 2.4 GHz.

Further analysis on the behavior of the capacitance is done by getting the capacitances at 2.34 GHz and 2.46 GHz and compared to the capacitance at 2.4 GHz. The percent change (%Δ) is computed, for both de-embedding methods, based on the center frequency of 2.4 GHz, given a bandwidth of 120 MHz. The %Δ metric shows the behavior of the capacitance operating at the frequency of interest within the given bandwidth of 120 MHz. The lower the %Δ the more accurate the capacitor value within the given frequency specification. The capacitances at 2.4 GHz are selected because this is the center frequency used for a Low Noise Amplifier (LNA) developed in the lab.

Table 2 enumerates the computation of the %Δ for both OPD and TRISD methods. The capacitor values generally show that all the fabricated capacitors, based on the capacitance at 2.4 GHz, have a %Δ of less than $\pm 1\%$. The computed %Δ resulting from the OPD data and TRISD data have comparable values implying that any of the two deembedding techniques can be used in extracting the capacitance.

The analysis is extended by computing %ΔCave given by (3). C_{opd} is the capacitance using OPD, C_{trisd} is the capacitance using TRISD, i is the capacitance at the ith frequency and j is the total frequency points from 270 MHz to 3 GHz This is the percent change of the capacitance

Table 2: OPD VS. TRISD

Capacitor	OPD %Δ		TRISD %Δ		$\%\Delta C_{ave}$ (%)
	2.34–2.4 GHz	2.4–2.46 GHz	2.34–2.4 GHz	2.4–2.46 GHz	
HPP	0.44	-0.56	0.49	-0.60	-0.869
Woven	0.27	-0.34	0.36	-0.42	-1.46

after the TRISD with respect to the capacitance obtained from the OPD. The frequency range of 270 MHz to 3 GHz is chosen since almost all of the capacitances within this range do not change abruptly as can be seen in the capacitance plots.

$$\%\Delta C_{ave} = \frac{\sum_{i=1}^{j}\left(\frac{C_{opd}-C_{trisd}}{C_{opd}}\right)}{j} \tag{3}$$

Equation (3) is applied to all structures and the results are listed also in Table 2. The smallest %ΔCave for the small capacitor structure is — 0.86% (the negative sign implies that the TRISD measurement is greater than that of OPD for the HPP capacitor structure. Since the highest %ΔCave is around 2%, it is therefore suggested that any of the two de-embedding techniques may be used to extract the capacitance.

7. CONCLUSION

A methodology of doing on wafer measurement of passive devices is presented. The inductor and capacitor test structures illustrated and validated the said methodology. Empirical results show that for measurements up to 6 GHz the variation extracted parameters using open and three step de-embedding method is nominal. The preferred de-embedding technique would be open de-embedding. The advantages OPD offers against the TRISD are the following:

(a) OPD mathematical approach is straight forward
(b) The concept and implementation is simple compared to TRISD
(c) The three-step de-embedding structures required for TRISD can be eliminated, resulting in a smaller chip area f.

ACKNOWLEDGMENT

The authors wish to acknowledge the assistance and support of Intel Philippines, DOST PCASTRD, IMI and Eazix Inc..

REFERENCES

1. Rosales, M. D., "Characterization, comparison and analysis of monolithic inductors in silicon for RF IC's," M.S. Thesis, University of the Philippines, April 2003.
2. Tan, H. B., "Characterization, comparison and analysis of monolithic capacitors in silicon for RF IC's," M.S. Thesis, University of the Philipines, April 2004.
3. "Layout rules for GHz-probing," *Application Note*, Cascade Microtech Inc.
4. Kolding, T. E., "On-wafer calibration tecniques for giga-hertz CMOS measrurments," presented at *Int'l Conf. on Microelectronic Test Structures*, 1999.
5. Cho, H. and D. Burk, "A Three-step Method for the De-embedding of high frequency S-parameter Measurements," *IEEE Transactions of Electron Devices*, 1991.

Extraction of Subterahertz Transmission-line Parameters of Coplanar Waveguides

Jingjing Zhang and Thomas Y. Hsiang

Department of Electrical and Computer Engineering and Laboratory for Laser Energetics
University of Rochester, Rochester, NY 14627-0231, USA

Abstract— Building on our previous work that compared computed propagation characteristics of picosecond pulses on coplanar waveguides (CPW's) with experiments, we extract transmission-line parameters, including resistance, inductance, conductance, and capacitance, from S-parameters up to subterahertz frequencies. The distributed-element circuit model is analyzed for applications in CPW-based circuit design.

1. INTRODUCTION

The coplanar waveguide (CPW) is commonly used as passive components and interconnects in high-speed circuits. The ideal CPW contains infinitely wide ground planes and can be exactly analyzed with conformal mapping [1]. Recent research interests, however, have focused on the practically important CPW's that have finite ground-plane widths and can only be analyzed by numerical simulations [2–6]. In our previous studies [5, 6], we used full-wave analysis to characterize the subterahertz propagation characteristics of both wide- and narrow-ground CPW's and compared the results with experimental measurements and, in the case of wide-ground lines, with closed-form analysis. Effects of ground-plane width and lateral line dimensions on the attenuation and dispersion characteristics were investigated up to subterahertz frequencies. The good agreement with experiments validates the numerical simulations. In this paper we extend the study to further extract the frequency-dependent transmission-line parameters from the numerical analysis. A distributedelement circuit model of CPW is described up to subterahertz frequencies.

2. TRANSMISSION-LINE PARAMETER EXTRACTION

A transmission line of length l, characteristic impedance Z_c, and propagation constant γ can be modeled as a two-port network, as shown in Fig. 1. The $ABCD$ matrix for the network is given as [6]

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} \cosh\gamma l & Z_c \sinh\gamma l \\ \frac{\sinh\gamma l}{Z_c} & \cosh\gamma l \end{bmatrix}. \tag{1}$$

Due to symmetry of the transmission line, the scattering matrix can be simplified with $S_{11} = S_{22}$ and $S_{12} = S_{21}$. Provided that both the input and output impedances equal the reference impedance Z_o, the relationship between the S-parameters and the $ABCD$ matrix becomes

$$\begin{aligned} A &= \frac{1 - S_{11}^2 + S_{21}^2}{2S_{21}} \\ B &= \frac{\left(1 + 2S_{11} + S_{11}^2 - S_{21}^2\right) Z_0}{2S_{21}} \\ C &= \frac{1 - 2S_{11} + S_{11}^2 - S_{21}^2}{2S_{21} Z_0} \\ D &= \frac{1 - S_{11}^2 + S_{21}^2}{2S_{21}} \end{aligned}, \tag{2}$$

Substituting A, B, C, and D into Eq. (1), the propagation constant and characteristic impedance are expressed in terms of S-parameters as below:

$$\begin{aligned} e^{-\gamma l} &= \frac{2S_{21}}{1 - S_{11}^2 + S_{21}^2 \pm \sqrt{\left(1 + S_{11}^2 - S_{21}^2\right)^2 - 4S_{11}^2}} \\ Z_c &= \pm Z_0 \sqrt{\frac{(1 + S_{11})^2 - S_{21}^2}{(1 - S_{11})^2 - S_{21}^2}} \end{aligned}. \tag{3}$$

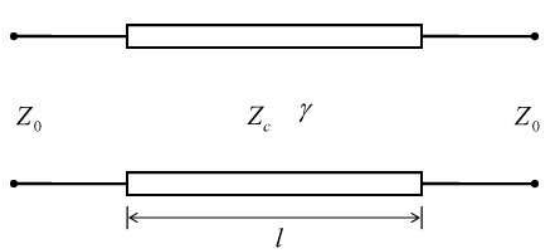


Figure 1: Two-port network of a transmission line.

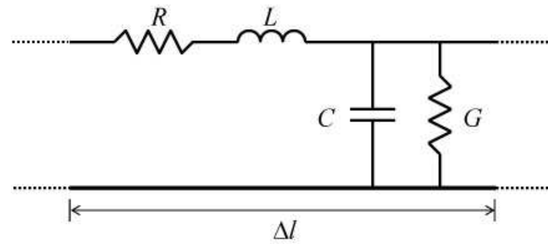


Figure 2: Distributed-element circuit model of a transmission line.

Experimentally, γ and Z are the measured quantities and are directly compared with the computed S parameters [5, 6].

In an alternative model, Fig. 2, which is more usefully for the circuit designer, the transmission line can be described with a distributed-element circuit model:

$$\begin{array}{l} \gamma = \sqrt{(R + j\omega L)\,(G + j\omega C)} = \alpha + j\beta \\ Z_c = \sqrt{\dfrac{R + j\omega L}{G + j\omega C}} \end{array}, \tag{4}$$

where R, L, G, C are the unit-length resistance, inductance, conductance, and capacitance; α and β are the attenuation and phase constants of signal propagation, respectively. Solving Eq. (4) yields the circuit-element parameters of the transmission line:

$$\begin{array}{l} R = \Re\{\gamma Z_c\} \\ L = \Im\{\gamma Z_c\}/\omega \\ G = \Re\{\gamma/Z_c\} \\ C = \Im\{\gamma/Z_c\}\,\omega \end{array}. \tag{5}$$

Since the S-parameters are readily obtained from our measurements or simulations, Eqs. (3-5) then provide the algorithm from which the circuit-model parameters, shown in Fig. 2, can be extracted.

3. RESULTS AND DISCUSSION

In our previous work [5, 6], two classes of CPW lines (made of Au on GaAs substrates) were studied experimentally and numerically. The first class contains ground planes at least 10 times wider than the center conductor or the conductor spacing (the "wg lines"), closely approximating an ideal CPW, while the second class uses narrow ground planes with the same width as the center electrode (the "ng lines"). Simulations of signal propagation that make use of full-wave analysis were compared with and verified by experimental data.

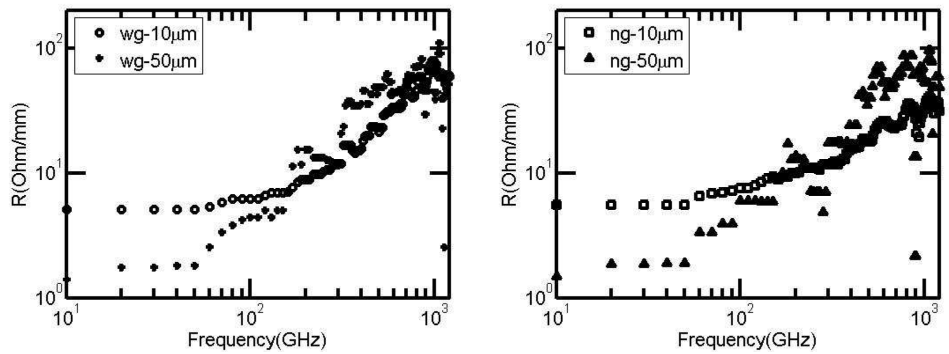
For computing the distributed-element circuit parameters, we chose to use the numerically simulated S-parameters as the starting point and solve Eqs. (3) and (5). Examples of these parameters are shown in Fig. 3 that correspond to the CPW's previously studied [5].

The strongest frequency-dependent parameter is the resistance term, R. At low frequencies, R approaches a constant, corresponding to the quasi-static limit where the resistance per unit length of the CPW can be calculated by

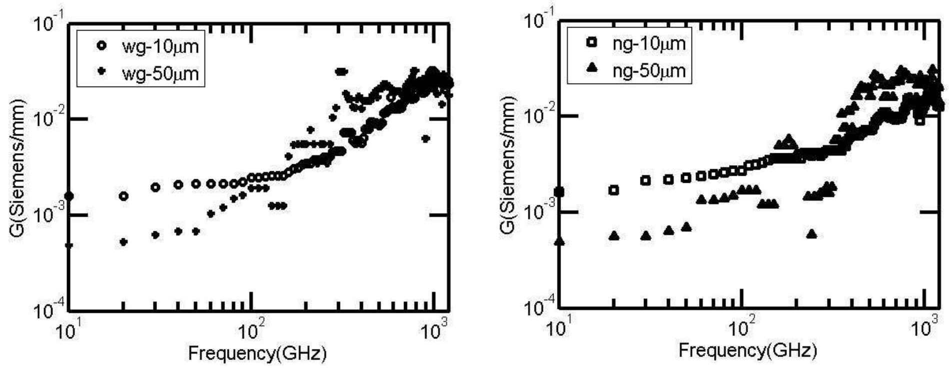
$$R = \frac{1}{\sigma W t}, \tag{6}$$

where σ is conductivity of the conductor metal, W and t are the width and thickness of the center conductor. For the 10-μm CPW, R is 7.63 Ω/mm and, for the 50-μm CPW, it is 1.53 Ω/mm; both closely matching the values in Fig. 3(a). As frequency increases (up to 100 GHz), the skin depth becomes smaller than the conductor thickness and surface resistance increases as the square-root of frequency. In this region where conductor loss dominates, wider lines would have lower loss compared with narrower lines, which can be seen from Fig. 3(a). At still higher frequencies, above 100 GHz, radiation loss becomes dominant. In this case, the narrower lines would have lower loss. The crossover of the value of R's for different line-widths as frequency is increased is prominently shown in the data.

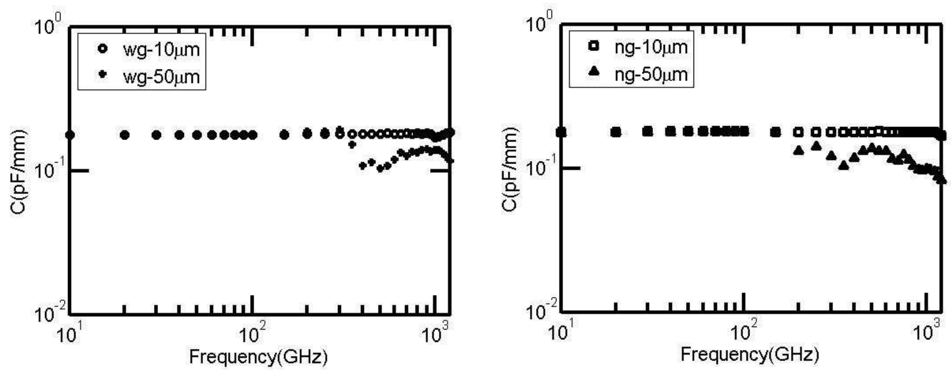
The width of the ground plane also has an effect on the value of R. Comparing the left and right curves of Fig. 3(a), it can be seen that at higher frequencies, the ng lines have a lower value



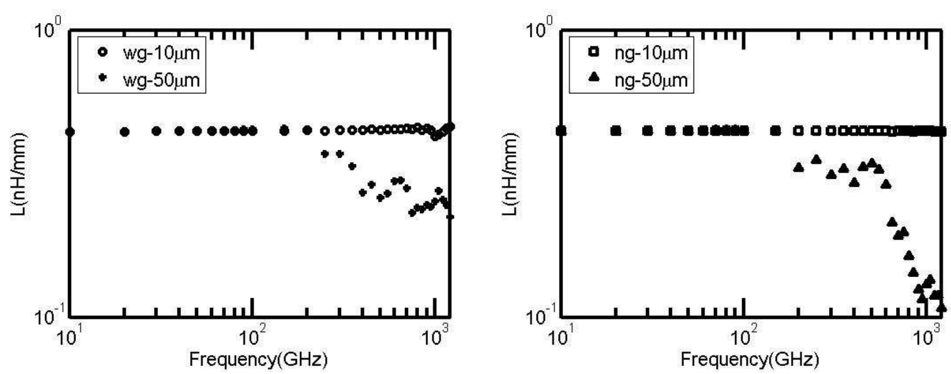
(a)



(b)



(c)



(d)

Figure 3: Transmission line parameters R, G, C, L of CPW's extracted from S-parameters. wg and ng refer to wide ground and narrow ground, respectively. 10 μm and 50 μm indicate the lateral line dimensions of the CPW's.

of R, as a result of the reduced coupling of surface modes with the ground conductors [3, 5]. This effect shows that the ng lines are preferred over the wg lines for high-frequency applications.

A secondary loss component for the CPW is its conductance term, G. In the quasi-static limit, G arises when the dielectric contains a finite loss tangent. The value of G increases with decreasing line separation. For each of the examples shown in Fig. 3(b), the line separation equals the line width, thus giving rise to a higher conductance for the narrower lines. In the 100's GHz range, we observe an increase in the value of G, followed by a cross-over in frequency for the two line widths shown, similar to that observed for R. A plausible explanation is that, in this region, part of the radiation loss contributes to the coupling between the conductors and increases G; since the loss due to the conductance is two orders of magnitude lower than that due to the resistance for the samples considered, a small part of the radiation loss dominates the dielectric loss at high frequencies.

The other two terms are the capacitive and inductive terms. In the low-frequency region and for the wg lines, capacitance is determined by [7]

$$C = (\varepsilon_r + 1)\varepsilon_0 2\frac{K(k)}{K'(k)}, \tag{7}$$

where ε_o is the vacuum permittivity, ε_r is relative permittivity of the substrate, and *K(k)/K'(k)* is a geometry coefficient that depends only on the line separation-to-width ratio. Using this expression, the capacitance of wg CPW's is found to be 0.157 pF/mm and closely approximates the quasi-static values in Fig. 3(c). We also observe that the ng lines have almost identical capacitance to that of the wg lines, showing that Eq. (7) is applicable irrespective of the ground-plane widths. Further, the capacitance for the 10-μm lines remains nearly constant to the highest frequency considered, showing that the surface modes which dominate the high-frequency losses have only a small effect on the overall field distribution. For the 50-μm lines, we observe a drop in C at higher frequencies; this is attributed to a computational artifact.

Inductance is proportional to the magnetic flux contained in the CPW. Its dominant component is the external inductance, associated with the magnetic flux between the electrodes, and can be expressed as:

$$L = \frac{1}{c^2 C_a}, \tag{8}$$

where c is speed of light and C_a is the capacitance with the substrate replaced by air. Using Eq. (7), we obtain $L = 0.49\,\mathrm{nH/mm}$, in agreement with the quasi-static values in Fig. 3(d). Secondary (and frequency-dependent) components in L include the magnetic flux in the conductor and the surface modes that modify the field distribution. These two effects appear to be negligible, as evidenced by the nearly flat frequency dependence. The drop in L at higher frequencies for the 50 μm lines is attributed to a computational artifact only.

4. CONCLUSIONS

In summary, we have extracted broadband transmission-line parameters *R, L, G, C* of CPW's from S-parameters computations, up to subterahertz frequencies. The origin, physics, and validation of the circuit model are discussed. Examples are given that demonstrate how these model parameters are affected by changing ground-plane widths and lateral dimensions.

REFERENCES

1. Gupta, K. C., R. Garg, and I. J. Bahl, *Microstrip Lines and Slotlines*, Artech House, Norwood, MA, 1979.
2. Heinrich, W., "Quasi-TEM description of MMIC coplanar lines including conductor-loss effects," *IEEE Trans. Microwave Theory Tech.*, Vol. 41, No. 1, 45–52, 2003.
3. Schnieder, F., T. Tischler, and W. Heinrich, "Modeling dispersion and radiation characteristics of conductor-backed CPW with finite ground width," *IEEE Trans. Microwave Theory Tech.*, Vol. 51, No. 1, 137–143, 2003.
4. Spiegel, S. J. and A. Madjar, "Light dependence of silicon FGCPW transmission lines," *Microwave Symposium Digest, 1999 IEEE MTT-S International*, Vol. 4, 1801–1804, 1999.
5. Zhang, J., S. Alexandrou, and T. Y. Hsiang, "Attenuation characteristics of coplanar waveguides at subterahertz frequencies," *IEEE Trans. Microwave Theory Tech.*, Vol. 53, No. 11, 3281–3287, 2005.

6. Zhang, J. and T. Y. Hsiang, "Dispersion characteristics of coplanar waveguides at subterahertz frequencies," *Journal of Electromagnetic Waves and Applications*, Vol. 20, No. 10, 1411–1417, 2006.
7. Gupta, K. C., R. Garg, and R. Chadha, *Computer-aided Design of Microwave Circuits*, Artech House, Dedham, MA, 1981.

Design of MEMS Controlled Phased Shifter Using SCT

N. Raveu[1], **E. Perret**[2], **H. Aubert**[1,3] **and H. Legay**[4]

[1]INPT-ENSEEIHT, France
[2]INPG-ESISAR, France
[3]LAAS-CNRS, France
[4]ALCATEL ALENIA SPACE, France

Abstract— The development of an original method called Scaled Changing Technique (SCT) has led to the fast analysis of planar reconfigurable phase-shifters presenting a great number of phase states. Phase-shifters have been designed for obtaining 360° phase-shift range and by taking into account the losses.

Reflectarrays using Radio-frequency Micro-mechanical Switches (RF-MEMS) is an emerging technology for reconfigurable and scanning antennas. In this paper, the planar phase-shifter is composed of 3 metallic patches and 10 RF-MEMS (see Fig. 1). The up and down states of the RF-MEMS control the phase-shift by tuning the slot length. A phase-shifter cell with 10 switches presents $2^{10} = 1024$ available configurations: in order to predict the phase-shift for each configuration, an electromagnetic simulation tool has been developed. This tool has been specifically dedicated to the study of these MEMS-controlled planar phase-shifter [1, 2]. Thanks to this specific tool, the fast design of phase-shifter cells covering the overall 360° phase-shift range has been performed.

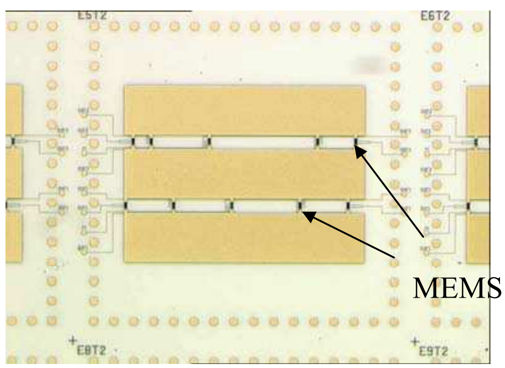


Figure 1: Planar MEMS-controlled phase-shifter.

The electromagnetic simulation tool is based on an original method called the *Scale Changing Technique* (SCT). In this technique, the discontinuity plane is first partitioned in several domains (see Fig. 2). All these domains are enclosed by perfect electric or magnetic boundary conditions. The transition between the domains is described by the so-called *scale changing network* derived from the modal analysis. Applied to the planar MEMS-controlled phase-shifter shown in Fig. 1 [1], the SCT allows the derivation of the equivalent network of the phase-shifter results from the cascade of 4 scale changing networks shunted by the RC impedance model of RF-MEMS (Fig. 3). The phase-shift φ is derived from the input impedance Z_{in} of this cascade.

Phase-shifters can be characterized in a metallic waveguide excited by the TE_{10}-mode. Electromagnetic simulations are performed here in these conditions. The losses are evaluated as the ratio between the power dissipated by the phase-shifter and the power of the incident TE_{10} mode [2]. The configurations that present high losses (higher than a given threshold) can then be identified and removed. The phase-shifter is designed for presenting 360° phase-shift range with a minimum phase-shift between successive RF-MEMS configurations and an attenuation less than 1dB. The performances of the phase-shifter are studied for many values of up-state capacitance C_{up} using the SCT. For $C_{\text{down}} = 1.5\,\text{pF}$ and $R = 1\,\Omega$, the ratio $C_{\text{down}}/C_{\text{up}}$ has to be higher than 9 in order to provide the desired 360° phase-shifts range (Fig. 4).

As indicated in Fig. 5, the losses imply a reduction of available configurations from 1024 to 892 if one wants an attenuation less than 1dB (with $R = 1\,\Omega$, $C_{\text{down}} = 1.5\,\text{pF}$ and $C_{\text{up}} = 15\,\text{fF}$): the

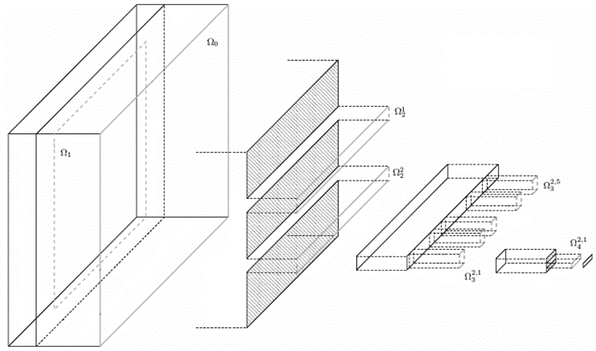


Figure 2: The multiple domains of the planar phase-shifter located in the cross section of a metallic waveguide.

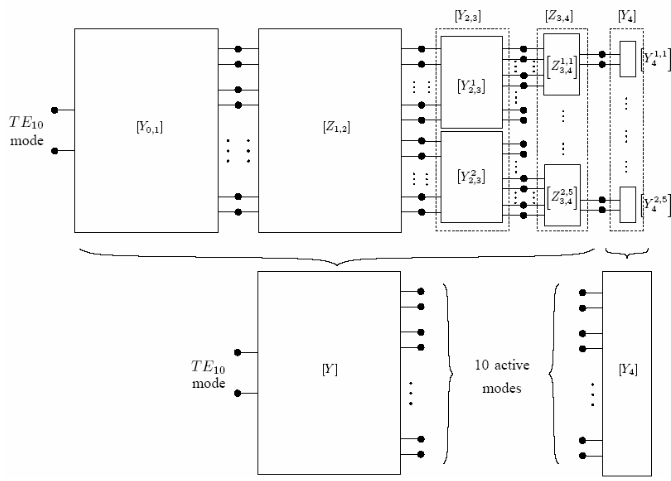


Figure 3: Equivalent network of the phase-shifter.

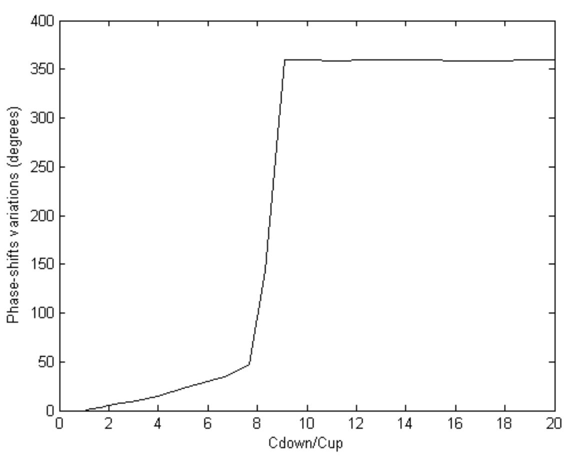


Figure 4: Phase-shift range at 11.7 GHz versus $C_{\mathrm{down}}/C_{\mathrm{up}}$.

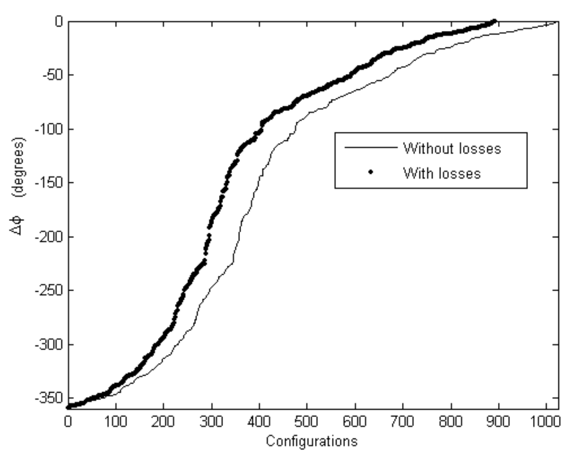


Figure 5: Phase-shift variation with and without losses at 11.7 GHz.

360° phase-shift range is still possible with no more than 15° between two successive phase-shifts (instead of 5° in the lossless case).

The SCT-based tool allows a fast identification of the RF-MEMS configurations that present high losses and a rapid electromagnetic analysis when the number and positions of RF-MEMS are tuned.

REFERENCES

1. Perret, E., H. Aubert, and H. Legay, "Scale-changing technique for the electromagnetic modeling of MEMS-controlled planar phase shifters," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 54, Issue 9, 3594–3601, Sept. 2006.
2. Perret, E., N. Raveu, H. Aubert, and H. Legay, "Scale changing technique for MEMS-controlled phase-shifters," *European Microwave Week*, Manchester, England, 10–15 Sept., 2006.

Waveguide Analysis Using Multiresolution Time Domain Method

Xiaoli Huang[1], **Xia Ling**[1], **and Haiyan Chen**[2]

[1]College of Electronics and Information, Xihua University, Chengdu, 610039, China
[2]School of Physics Science and Technology, Yangtze University, Jingzhou, 434023, China

Abstract— In this paper, we report on an analysis method using wavelet-based FDTD in numerical simulation of erbium-ytterbium co-doped waveguide by multi resolution analysis (MRA). The difference formulation for multiresolution time domain(MRTD) based on Daubechies wavelets is derived, and the propagating characteristics of the active waveguide is simulated by MRTD. It demonstrates that this new algorithm is valid.

1. INTRODUCTION

Optical waveguide amplifiers based on rare earth-doped materials have become the most attractive and efficient devices to regenerate optical signals being processed through optical multichannel systems[1−3]. This tendency has motivated dynamic research activity oriented to the precise numerical simulation. Comparison with erbium doped fiber amplifier (EDFA), one of characteristics of erbium doped waveguide amplifier(EDWA) fabricated is not readjusted, the precise numerical simulation is especially important. General computational methods are finite element method(FEM)[4], FD-BPM[5], FDTD[6], ADI-FDTD[7], and FDTD-based overlapping integral integral-RK method[8], etc. Among above algorithm FDTD is an efficient method, but it has some major drawbacks including its massive memory consumption and computational time. It is shown that the large-stencil finite-difference scheme based on Daubechies compactly supported orthogonal wavelets and the Deslauriers-Dubuc interpolating functions as biorthogonal wavelet bases is a efficient replacement for traditional FDTD[9, 10]. The advantage of the Daubechies wavelets and Deslauriers-Dubuc interpolating wavelets is that, by virtue of their interpolation property, their expansion coefficients represent direct physical values of the field.

In this paper, we apply wavelets based Finite-Differences Time-Domain Method, multiresolution time domain, to the numerical simulation of the gain characteristics of erbium and ytterbium co-doped phosphate glass waveguide amplifier. This algorithm takes advantage of compression properties of wavelet transforms and multi resolution analysis (MRA). Simulation results show the validity of the algorithm.

2. AMPLIFIER MODELING BASED ON MRTD

2.1. Computation of Lateral Profiles of Pump and Signal Lights

We consider Maxwell's curl equations in a erbium-ytterbium co-doped waveguide described in [7]. Using MRTD the time evolution equations can be obtained as:

$$H_x^{n+1/2}(i,\, j+1/2) \;=\; H_x^{n-1/2}(i,\, j+1/2) - \frac{\Delta t}{\mu \Delta y} \sum_{l=-2L+1}^{l=2L-1} a(l) E_z^n(i,\, j+l+1) \tag{1}$$

$$H_y^{n+1/2}(i+1/2,\, j) \;=\; H_y^{n-1/2}(i+1/2,\, j) - \frac{\Delta t}{\mu \Delta x} \sum_{l=-2L+1}^{l=2L-1} a(l) E_z^n(i+l+1,\, j) \tag{2}$$

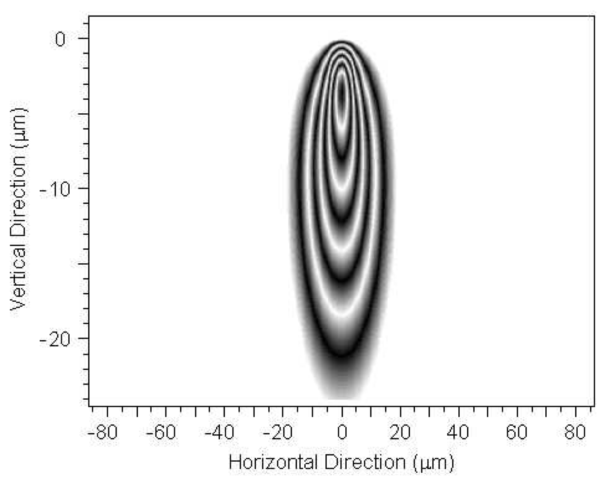
$$\begin{aligned} E_z^{n+1}(i,\, j) \;=\; & E_z^n(i,\, j) + \frac{\Delta t}{\varepsilon} \left| \sum_{l=-2L+1}^{l=2L-1} a(l) \right| \\ & \times \left| \frac{H_y^{n+1/2}(i+l+1/2,\, j)}{\Delta x} - \frac{H_x^{n+1/2}(i,\, j+l+1/2)}{\Delta y} \right| \end{aligned} \tag{3}$$

where E_z denotes the electronic field, H_x and H_y are the magnetic fields, ε and μ are the dielectric constant and the megenetic permeability of media, respectively. Δx and Δy are the x- and y-direction step, and Δt the time step, the connection coefficients $a(l)$ is listed in Table1 for the p-th order schemes.

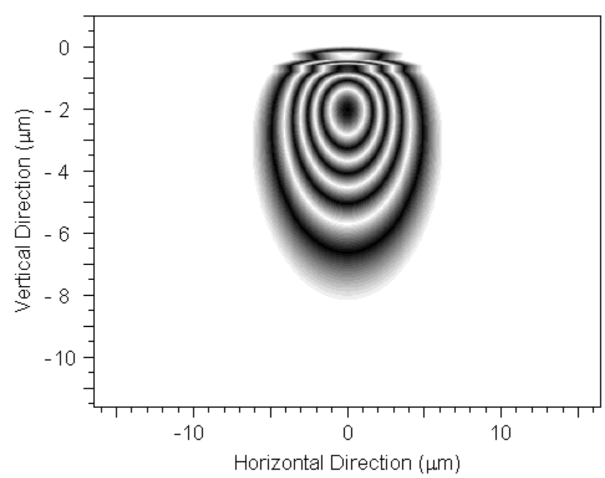
The absorption boundary conditions for TE mode are the same as [10], wavelet-collocation scheme is the 2th-order interpolation function. The space and the time increments were $\Delta x = \Delta y = 0.05\,\mu$m, and $\Delta t = 0.23$ fs. The width and depth $6\,\mu$m and $3\,\mu$m, the profile of refractive index is chosen as diffusion type for ion-exchange waveguides, the refractive index of substrate is 1.5288, and the difference of refractive is 0.01. the calculated fundamental mode s for signal (@1.532 μm) and pump(@0.98 μm) are shown in Fig. 1.

Table 1: Connection coefficients $a(l)$ for Deslauriers-Dubuc interpolation bases $a(-l) = -a(l-1)$

l	D_2	D_4	D_{10}
0	1.2291666667	1.3110340773	1.3033236013
1	−0.0937500000	−0.1560100710	−0.1636941766
2	0.0104166667	0.0419957460	0.0616127747
3		−0.0086543236	−0.0265749940
4		0.0008308695	0.0104549954
5		0.0000108999	−0.0034151723
6		−0.0000000041	0.0008675397
7			−0.0001583521
8			0.0000179275
9			−0.0000007829
10			−0.0000000260
11			−0.0000000072
12			0.0000000014
13			0.0000000001



(a)



(b)

Figure 1: Fundamental modes propagating in waveguides, (a) Signal light @1.532 μm, (b) Pump light @0.98 μm.

3. CONCLUSIONS

We gives a try on the application of MRTD analysis based on Daubechies wavelets on the calculation of the gain characteristics of erbium-ytterbium co-doped waveguide amplifier, results demonstrate that wavelets based FDTD is an efficient method.

ACKNOWLEDGMENT

This work is supported the Research Foundation of Education Bureau of Hubei Province, China (No. D200512011).

REFERENCES

1. Chen, H. Y., Y. Z. Liu, J. Z. Dai, and X. L. Huang, *IEEE 2002 International Conf. on Communications Circuits and Systems Proc.*, 827, 2002.
2. Chen, H. Y., Y. Z. Liu, J. Z. Dai, and X. L. Huang, *The 8th Optoelectronics and Communications Conference proc.*, 515, 2003.
3. Reichmann, K. C., P. P. Iannone, M. Birk, N. J. Frigo, D. Barbier, C. Cassagnettes, T. Garret, A. Verlucco, S. Perrier, and J. Philipsen, *IEEE Photon. Technol. Lett.*, Vol. 13, 1130, 2001.
4. Pasquale, F. D. and M. Zoboli, *IEEE J. lightwave Technol.*, Vol. 11, 1565, 1993.
5. Guzmán, A. M. and H. J. Kalinowski, *V State Symposium of Lasers Application*, 274, 1992.
6. Chen, H. Y., X. L. Huang, and Y. Z. Liu, *Acta Photonica Sinica*, Vol. 30, 214, 2001.
7. Chen, H. Y., J. Z. Dai, X. J. Huang, and Y. Z. Liu, *Computational Physics*, Vol. 20, 443, 2003.
8. Chen, H. Y., *International Journal of Infrared and Millimeter Waves*, Vol. 26, 555, 2005.
9. Fujii, M. and W. J. R. Hoefer, *IEEE Journal of Quantum Electronics*, Vol. 37, 1015, 2001.
10. Guo, Y. F., F. M. Kong, K. Li, Y. Liu, and J. Ling, *Journal of Optoelectronics Laser*, Vol. 15, 238, 2004.

Application of FDTD-overlapping Integral Method in Simulation of Waveguide Amplifier

J. Z. Wang[1,2], **P. Zhang**[2], **and Y. Z. Liu**[1]

[1]College of Optical-Electronic Information
University of Electronic Science and Technology of China, China
[2]Computer Teaching Center
Sichuan Normal University, Chengdu, China

Abstract— A novel algorithm, calling FDTD-based overlapping integral-RK method, is proposed to analyze Er-Yb co-doped phosphate glass IR waveguide amplifiers. This new method is derived from the combination of Finite-Difference Time-Domain (FDTD), overlapping integral and RK method. First the normalized eigen-fields of signal and pump lights in Er-Yb co-doped phosphate glass waveguide are calculated by FDTD algorithm, and then the overlapping integrals between light fields and the distribution of Er3+(Yb3+) concentrations are obtained, finally the distributions of powers of signal and pump along waveguide are calculated by RK method. Gain performance of IR waveguide amplifiers is obtained from calculated signal power. Comparison of the simulation result with an experiment is given, it demonstrates that this new algorithm is valid and shows good agreement with experimental results. Abstract should summarize the approach and the article's major contributions, emphasizing the importance and significance of the results.

1. INTRODUCTION

Recently, IR optical waveguide amplifiers based on Er-Yb co-doped phosphate glass have become the most attractive and efficient devices to regenerate optical signals being processed through optical multichannel systems [1–5]. Applications already foreseen are integrated compact power booster and preamplifiers, and lossless splitters for fiber-to-the-home (FTTH), multi-laser optical radar, optical phase-difference radar, phased-array laser radar, and optical phased-array radar.

The numerical modeling for IR waveguide amplifiers includes calculation both transverse profile of lights in waveguide and distribution of light power along waveguide. The general algorithms are Runge-Kutter (RK) combined with Finite- Element Method (FEM) [6], Beam Propagation Method(BPM) [7], Line Method [8], Finite Difference-Beam Propagation Method (FD-BPM) [9], and overlapping integral method [10]. In [6–9] the distributions of signal and pump lights in waveguide are repeatedly calculated while calculating the power distribution of signal and pump lights along waveguide, leading to increase of calculation workload. Although overlapping integral method [10] is simple, its precision is not enough high, because the key parameter-overlapping integral is estimated.

In this paper, we introduce FDTD into the calculation of Er-Yb co-doped phosphate glass integrated waveguide, and the eigen-fields propagating in the waveguide are given. The overlapping integrals between light fields and the distribution of Er3+(Yb3+) concentrations are obtained and gain performance of IR waveguide amplifiers is discussed by RK method using the calculated overlapping integrals. Comparison of the simulation result with an experiment is given.

2. FDTD-OVERLAPPING INTEGRAL-RK METHOD

2.1. Calculation of Eigen-fields Propagating in Waveguide Using FDTD

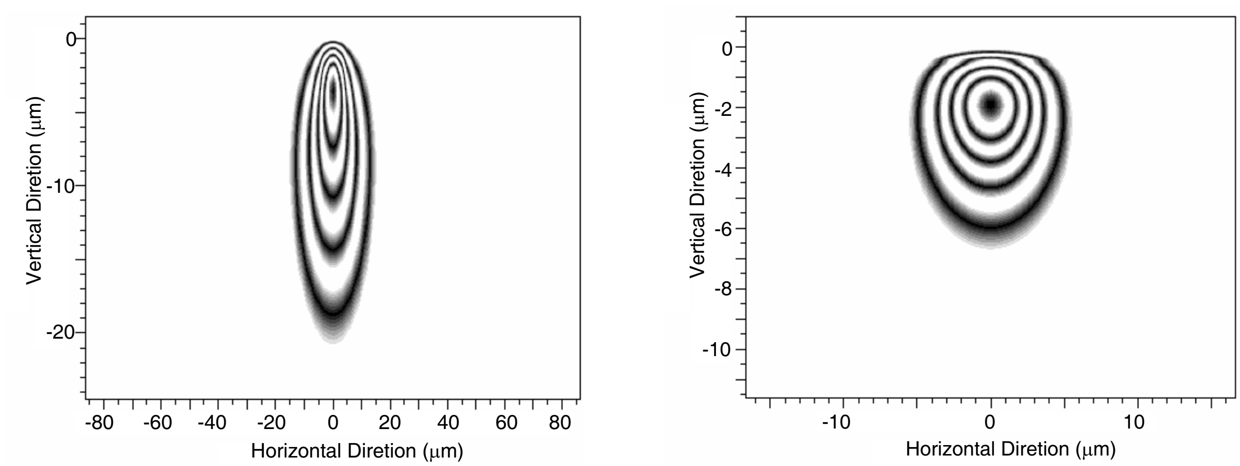
First we calculate the eigen-fields propagating in Er-Yb codoped phosphate glass waveguide using FDTD. We use Yee's scheme of a Cartesian mesh with proper field vectors at the discrete points. The vectors of one Yee-cell will be addressed with the same integer triple $(i,\, j,\, k)$ denoting the cell position. The relations between all field components look similar to the following one [11]:

$$\begin{aligned} E_x^{n+1}\left(1+\frac{1}{2},\, j,\, k\right) =& E_x^{n}\left(1+\frac{1}{2},\, j,\, k\right) + \frac{\Delta t}{\varepsilon_0\varepsilon_r}\left\{\frac{1}{\Delta y}\left[H_z^{n+1/2}\left(i+\frac{1}{2},\, j+\frac{1}{2},\, k\right)\right.\right. \\ & \left. -H_z^{n+1/2}\left(i+\frac{1}{2},\, j-\frac{1}{2},\, k\right)\right] - \frac{1}{\Delta z}\left[H_y^{n+1/2}\left(i+\frac{1}{2},\, j,\, k+\frac{1}{2}\right)\right. \\ & \left.\left. -H_y^{n+1/2}\left(i+\frac{1}{2},\, j,\, k-\frac{1}{2}\right)\right]\right\} \end{aligned} \tag{1}$$

where E_x denotes the electronic field, H_z and H_y are the magnetic fields, ε_0 and ε_r are the dielectric constant in vacuum and the relative permittivity of media, respectively. Δy and Δz are the y- and z-direction steps, n is the number of the time step Δt.

$$\Delta t = 1/c\sqrt{(1/\Delta x)^2 + (1/\Delta y)^2 + (1/\Delta z)^2} \tag{2}$$

where Δx are the x-direction step, c is the speed of light in free space. In calculation we use the absorption boundary conditions [11]. The space and the time increments are $\Delta x = \Delta y = \Delta z = 0.05\,\mu$m,and $\Delta t = 0.23$ fs. The width and depth of waveguide are $6\,\mu$m and $3\,\mu$m, the profile of refractive index is chosen as diffusion type for ion-exchange waveguides, the refractive index of substrate is 1.5288 and the difference of refractive index is 0.01. The calculated fundamental modes for signal (@1.535 μm) and pump (@0.98 μm) are shown in Fig. 1.



(a) Signal light @1.535 μm.

(b) Pump light @0.98μm.

Figure 1: Fundamental modes propagating in waveguide.

2.2. Calculation of Overlapping Integrals

The overlapping integrals between light fields and the distribution of Er3+(Yb3+) concentrations are defined as [2] :

$$\Gamma_p = \iint_A \psi_p(x,\,y) f(x,\,y) \mathrm{d}x\mathrm{d}y \tag{3}$$

$$\Gamma_s = \iint_A \psi_s(x,\,y) f(x,\,y) \mathrm{d}x\mathrm{d}y \tag{4}$$

where ψ_p, ψ_s and f, respectively, are normalized transverse field profiles of pump and signal, and transverse concentration profiles of dopant ions(Er3+ and Yb3+), A is the cross section area of active waveguide.

From (3), (4) and mode profiles propagating in Er-Yb co-doped phosphate glass waveguide we can obtain $\Gamma_p = 0.65$ and $\Gamma_s = 0.58$.

2.3. Solving Power Distributions along Waveguide by RK Method

A detailed description of the amplifier model used in this work can be found in [10]. Ignoring propagation loss of waveguide the depletion of the pump beam and the growth of the signal beam, as they propagate along the z-axis ($z = 0$ is the launch pump and signal point), can be described by the following set of differential equations[2]:

$$\mathrm{d}p_p(z)/\mathrm{d}z = -\Gamma_p\left(\sigma_{ap}n_3 - \sigma_{ep}n_4\right)p_p(z) \tag{5}$$

$$\mathrm{d}p_s\left(z,\,v_s\right)/\mathrm{d}z = \Gamma_s\left(\sigma_{es}n_2 - \sigma_{as}n_1\right)p_s(z,\,v_s) \tag{6}$$

With boundary condition

$$p_p(0) = P_{p0},\ p_s(0, v_s) = P_{s0}(v_s) \tag{7}$$

where P_{p0} and P_{s0} are input power of pump and signal, respectively. $Ps = A \times Is$, $Pp = A \times Ip$, Ps and Pp are the powers of signal and pump lights, and Is and Ip the intensities. σ_{es} and σ_{ep} are the emission cross section of the signal and pump photons, σ_{as} and σ_{ap} are the absorption cross section of the signal and pump photons, respectively. Level 1 and 2 represent the 4I15/2 and 4I13/2 of Er3+, and level 3 and 4 represent the 2F7/2 and 2F5/2 , respectively. $ni(i = 1 - 4)$ is the populations of level i.

The equations (6) and (7) can be solved to obtain gain performance of IR integrated waveguide amplifiers using the RK method.

The gain of amplifier is defined as

$$\text{Gain} = 100 \log_{10} p_s(z = L)/P_{s0}(z = 0) \tag{8}$$

where L is the length of amplifiers.

3. RESULT AND DISCUSSION

Simulation of slab IR waveguide amplifiers based on the algorithm of the preceding sections is performed. The IR waveguide amplifier parameters are shown as: substrate index is 1.5288, and the index transverse profile of waveguide is ion-exchange type, Er3+ concentration NEr $= 1.51 \times 1026\,/\text{m}^3$, Yb3+ concentration NYb $= 1.95 \times 1027\,/\text{m}^3$, $\sigma_{ap} = 1.0 \times 10 - 24\,\text{m}^2$, $\sigma_{as} = 6.8 \times 10 - 25\,\text{m}^2$, $\sigma_{es} = 7.9 \times 10 - 25\text{m}^2$, $\sigma_{ep} = 1.0 \times 10 - 24\,\text{m}^2$, signal wavelength is 1.535 μm, and pump wavelength 0.98 μm, $\Gamma p = 0.65$, $\Gamma s = 0.58$, $L = 1.8$ cm. Schematics of the amplification measurement setup in Fig. 2.

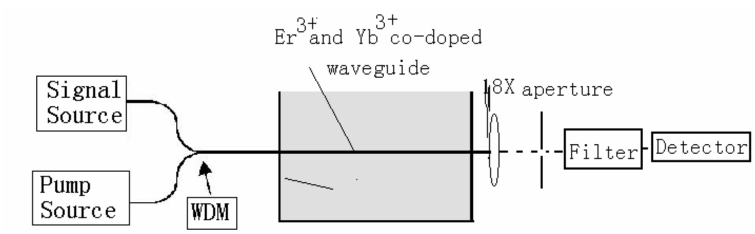


Figure 2: Schematics of the amplification measurement setup.

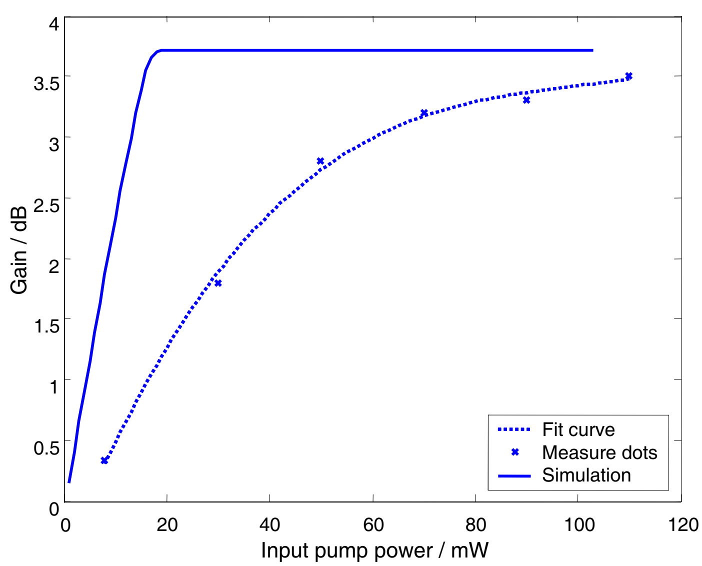


Figure 3: Small gain vs input pump power.

The small signal gain of IR waveguide amplifiers both theoretical calculation and experimental data fabricated by ion exchange technology are plotted in Fig. 3. The maximum gain difference between theoretical and experimental values is ~ 0.2 dB, it demonstrates that the new algorithm presented in this paper is valid and good agreement with experimental data.

From Fig. 3, we can see that there is some larger difference between theoretical analysis and experimental results, especially when pump power is ~ 20 mW. We think that this large difference is independent of this new algorithm, and is caused by the modeling used. Because we ignore the up-conversion caused by high concentration Er dopant in calculation, that the up-conversion exist leads to decrease of pump efficiency, though the sensitization of Yb3+ can reduce the effect of up-conversion and improve the pump efficiency, the effect of up conversion can not be completely eliminated, for example, we can see blue light during experiments, so compared with theoretical analysis the experimental result is smaller.

Furthermore, we ignore loss of ion-exchange waveguide, this is another case leading to the fact that experimental value of threshold is larger than theoretical one.

4. CONCLUSIONS

The gain characteristics of Er-Yb co-doped phosphate glass IR waveguide amplifiers is studied using FDTD-based overlapping integral-RK method. The calculated overlapping integrals for signal and pump are 0.58 and 0.65, the maximum gain simulated for a amplifier of 1.8 cm with pump power of 110 mW at 980 nm is ~ 3.75 dB, experimental value is ~ 3.5 dB. It demonstrates that the new algorithm presented in this paper is valid , another advantage of this new method is avoiding the repeated calculation of normalized light fields at every step power calculation, and thus saves calculating workloads.

REFERENCES

1. Wong, S. F., E. Y. B. Pun, and P. S. Chung, "Er3+-Yb3+ codoped phosphate glass waveguide amplifier using Ag^+-Li^+ ion exchange," *IEEE Photon. Technol. Lett.*, Vol. 14, 80–82, 2002.
2. Chen, H., Y. Liu, J. Dai, et al., "Modeling and simulation of saturation gain of ir integrated photonic amplifiers," *International Journal of Infrared and Millimeter Waves*, Vol. 25, 1791–1798, 2004.
3. Chen, H. Y., Y. Z. Liu, and J. Z. Dai, "Er3+/Yb3+ co-doped phosphate glass waveguide amplifier," *IEEE 2002 International Conference on Communications Circuits and Systems Proceedings(ICCCAS'02)*, 827–829, Chengdu, China, 2002.
4. Chen, H., Y. Liu, J. Dai, and Y. Yang, "High gain and broad band optical waveguide amplifiers," *Microwave and Optical Technology Letters*, Vol. 42, 64–66, 2004
5. Reichmann, K. C., P. P. Iannone, M. Birk, N. J. Frigo, D. Barbier, C. Cassagnettes, T. Garret, A. Verlucco, S. Perrier, and J. Philipsen, "A eight-wavelength 160-km transparent metro WDM ring network featuring cascaded erbium-doped waveguide amplifiers," *IEEE Photon. Technol. Lett.*, Vol. 13, 1130–1132, 2001.
6. Fabrizio, D. P. and M. Zoboli, "Analysis of erbium-doped waveguide amplifiers by a full-vectorial finite-element method," *IEEE Lightwave Technol.*, Vol. 11, 1565–1573, 1993.
7. Caccavale, F., F. Segato, and I. Mansour, "A numerical study of erbium doped active LiNbO3 waveguides by the beam propagation method," *Lightwave. Technol.*, Vol. 15, 2294–2300, 1997.
8. Huang, W. and R. R. A. Syms, "Analysis of folded erbium-doped planar waveguide amplifiers by the method of lines," *Lightwave Technol.*, Vol. 17, 2658–2664, 1999.
9. Yu, Z., L. M. Gao, W. Wei, et al., "Numerical analysis of amplification characteristic of erbium-doped waveguide amplifier by FD-BPM," *Optical and Quantum Electron.*, Vol. 36, 321–330, 2004.
10. Vallés, J. A., J. A. Lázaro, and M. A. Rebolledo, "Modeling of integrated erbium-doped waveguide amplifiers with operlapping factors methods," *IEEE J. Quantum Electron.*, Vol. 32, 1685–1694, 1996.
11. Kawano, K. and T. Kitoh, *Introduction to Optical Waveguide Analysis: Solving Maxwell's Equations and Schrödinger Equation*, 233–248, John Wiley & Sons, Inc, 2001.

Reliability Analysis of the Circuit and FM Modulation Parameters for the First Harmonic Level Reduction of the Forward Switching Power Supplies

Shahram Hosseinzadeh and Nader Samsunchi
Electrical Engineering Group, Department of Engineering
Azarbijan University of Teachers Education, 35th Km Tabriz Maragheh Road, Tabriz, Iran

Abstract— In this paper, the sensitivity of first harmonic of switching power supplies with the FM modulation of switching signal is investigated. Inspections are done with respect to the circuit and FM modulation parameters. In this regard, sets of orthogonal experiments are designed based on the Latin Cubes concept. Then the results of the experiments (i.e., the level of the 1st harmonic) are interpolated on a response surface. The response surface shows the self and the mutual effects of the factors on the level of the 1st harmonic.

It is seen that the minor changes of the circuit and modulation parameters has a major effect in the level of the 1st harmonic. Consequently, after circuit design process, sensitivity analysis must be done with respect to the circuit parameters and the level of 1st harmonic can be reduced by small changes in the circuit parameters. It is seen that the response surface can be readily used for analysis of the effects of parameters and with the small amounts of experiments, suitable results are achievable on the level of the 1st harmonic.

The harmonic level is more sensitive to the reactive elements and the load than the parameters of modulation. Therefore in the design for low EMI, reactive elements must be selected with care. Dominant interference depends on the load. The effect of load has positive sign in response surface. Therefore increasing the load intensifies the harmonic level.

1. INTRODUCTION

Power electronic has a history that is much older than many of peoples practicing in the field today are likely to realize [1]. The evolution of this branch of science, grows the other branches of science. Its progression also depends on the other branches of science (like the semiconductors technology). Doubtless, matureness of the power electronic is due to introduction of switching power supplies. Today the efficiency of the switching power supplies are beyond the 90 percent [2]. Without any aggression, switching power supplies become an important part of mankind life. They are power supplies of most electrical and electronic equipments.

Recently, attentions are paid to reduction of the size of electrical and electronic equipments. Consequently, increasing the power density and the integrations of switching power supplies as a principal element of these equipments are considered. The main factor that prevents the shrinking of switching power supplies is the size of the reactive elements. This factor can be overcome by increasing of the switching frequency.

The principle operation of switching power supplies is based on the switching of a DC voltage and passing it through a filter with the proper frequency characteristics. Filter is responsible to attenuate the unwanted harmonics. Fig. 1 shows basic structure of a forward switching power supply. The transistor T_1 acts as a switch element of the input voltage. When the switch is on, inductor L_1 and capacitor C_2 are charged by the V_{in} voltage. The load power also is provided by power supply. In the off state of the switch, the reactive elements are discharged to the load. The diode closes the path for the inductor current. Output voltage depends on the duty cycle of the T_1 switching pulse.

Switching power supplies are sources of interfering harmonics. Increasing the switching frequency eases the coupling the harmonics to the neighborhood circuits. Therefore EMC/EMI aspects must be seriously considered in the design phase of such power supplies.

In this paper, the sensitivity of interference signal due to the first harmonic of the switching power supplies with the FM modulation of switching signal is investigated. Inspections are done with respect to the circuit and FM modulation parameters. In this regard, sets of orthogonal experiments are designed based on the Latin Cubes concept. Then the results of the experiments (i.e., the level of the 1st harmonic) are interpolated on a response surface. The response surface shows the self and the mutual effects of the factors on the level of the 1st harmonic.

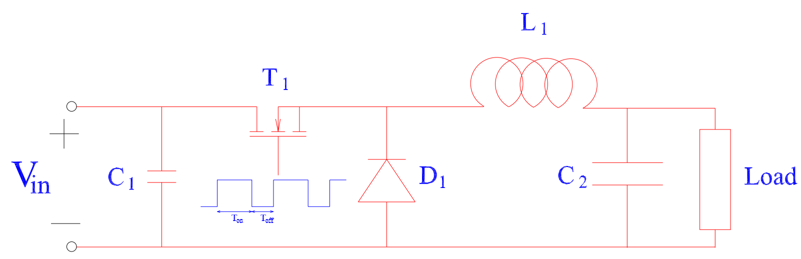


Figure 1: Forward Switching Power Supply.

The paper is organized as follows. The second section review the EMC aspects of switching power supplies. Then in the third section, the methodology for the design and analysis of experiments are presented. The fourth section of the paper is devoted on the sensitivity analysis of the forward switching power supply. Finally, the conclusion of the paper is provided.

2. EMC/EMI OF SWITCHING POWER SUPPLIES

Switching power supplies are sources of harmonics. Increasing the switching frequency eases the coupling the harmonics to the neighborhood circuits. Therefore EMC/EMI aspects must be seriously considered in the design phase of such power supplies.

Electromagnetic interference from an electromagnetic equipment to another one can be readily shown by the simple source-transmission path-victim model. The produced signal in the source is transferred to the victim by the transmission path. With regard to this model, EMI can be aviated by decreasing the interfering signal in the source, attenuating the transmission path and decreasing the sensitivity of victim, EMI can by aviated. It can be achieved by the proper design of PCBs, selection of proper ground in order to prevent the conductive interference and shielding [4].

The main interference source in switching power supplies are harmonics. In regular switching power supplies, harmonics concentrates on the integer multiplies of the switching frequencies, as shown in Fig. 3. By increasing the switching frequency, the frequency of the 1st harmonic, which is dominant harmonic, raises and decreases the attenuation of transmission path for this harmonic. The most usual method for decreasing the harmonic level, is to spread the power of harmonics in the whole frequency band. Consequently the level of harmonics is decreased. Frequency modulation, spread spectrum, $\Sigma\Delta$ and chaos have been used for this purpose [5–11].

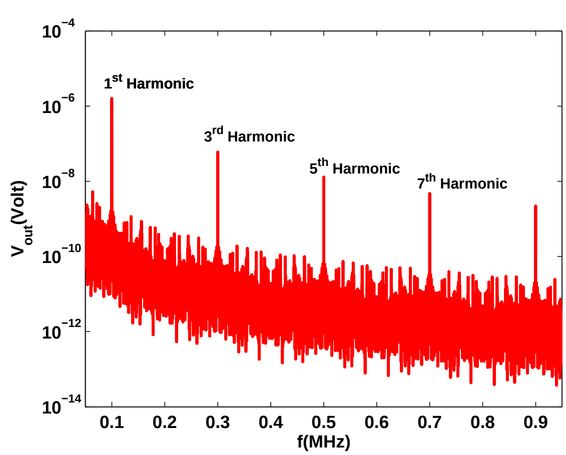


Figure 2: A sample of harmonic distribution in the forward switching power supply.

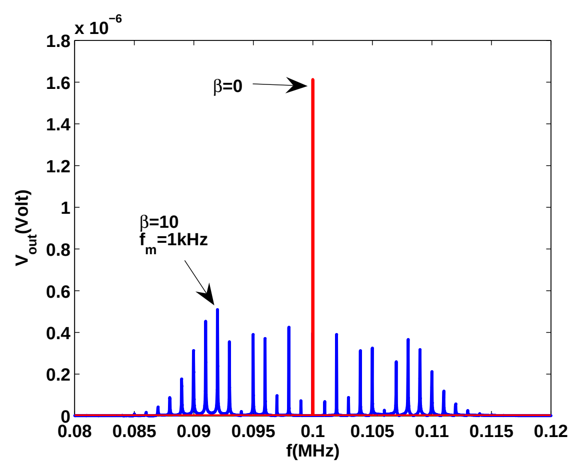


Figure 3: The effect of FM modulation on the harmonic distribution around the first harmonic.

3. EXPERIMENT DESIGN

Generally it seems that the improvement of the quality of switching power supply can be achieved by the following methods.

1- Selection of Proper Topology for the switching power supply
2- Proper parameters design for the selected topology
3- Proper sensitivity selection for the parameters

This paper considers only the third method for harmonic level reduction. In the usual method for analyzing of a response to the factors, only one factor varies and the other factors remain constant. This method do not consider the interactions between factors and consequently do not give suitable view of the response characteristics due to the effective factors.

Table 1: Typical values of the parameters.

L	1 mH
C	500 µF
R	20 Ω
T_s	1 µs
f_m	1 kHzs
β	10 s

Table 2: The effects for 7 levels of the parameters (effects are multiplied to 1000).

Factor	1	L	C	R_t	f_m
Effect	-0.1922	0.3874	-0.0249	0.0161	0.0043
Factor	β	L^2	LC	LR_t	Lf_m
Effect	0.0080	-0.1874	0.0024	-0.0081	-0.0012
Factor	$L\beta$	C^2	CR_t	Cf_m	$C\beta$
Effect	-0.0016	0.0104	-0.0034	0.0027	-0.0012
Factor	R_tf_m	$R_t\beta$	$f_m\beta$		
Effect	-0.0024	-0.0019	-0.0033		

For obtaining the most information from experiment, for all parameters, all parameters must be considered in the experiments. In this method we have a large amounts of experiment, which is usually impossible.

Therefore, we are looking for the maximum information from minimum amount of experiments. This is possible by Latin Cube concept, by which, design of orthogonal experiments are possible. In the orthogonal experiments, the interaction of the certain level of certain factor by the certain level of the another factor only is considered in the one of experiments [12].

4. SENSITIVITY ANALYSIS

The responses of the designed orthogonal experiments experiment can be fitted on the response surface as

$$y = \beta_0 + \sum_{i=1}^{N} \beta_i x_i + \sum_{i=1}^{N}\sum_{i=1}^{N} \beta_{i,j} x_i x_j \tag{1}$$

where y is response, x_i is the ith factor, β_i is the effect of ith factor and $\beta_{i,j}$ is the interaction of ith and jth factors.

The effects and interactions can be readily found from experiments as follows

$$\underline{\beta} = (\underline{\underline{X}}^t \underline{\underline{X}})^{-1} \underline{\underline{X}}^t \underline{y} \tag{2}$$

wherein $\underline{\beta}$ and $\underline{y}$ are vectors contain the effects (self and mutual) and the response from the designed experiments respectively. $\underline{\underline{X}}$ is a matrix contains the factors [13].

The sensitivity of the response to the factors can be readily found as

$$\frac{\partial y}{\partial x_i} = \beta_i + \sum_{j=1}^{N} \beta_{i,j} x_j \tag{3}$$

5. NUMERICAL EXAMPLE

As an example for application of procedure, sensitivity analysis of a forward switching power supply (which is shown in Fig. 1) is considered. Table 1 shows the typical values of the circuit parameters.

Figure 3 and Fig. 4 show the frequency spectrum around the 1st and 9st harmonic. Spreading of harmonics with the aid of the FM modulation can be observed. Comparison of Fig. 3 and Fig. 4 shows the harmonics spread in wider frequency band for the higher harmonics than the low harmonics.

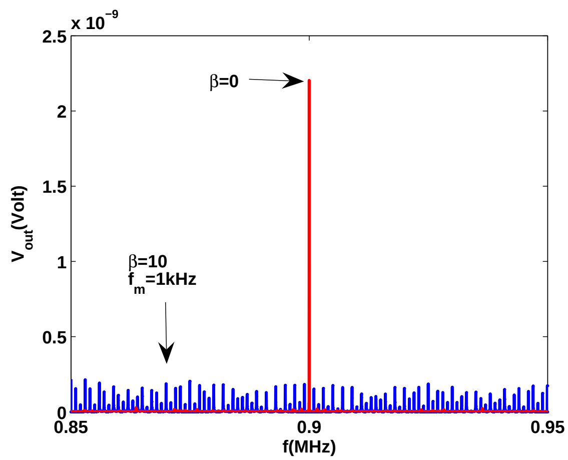


Figure 4: The effect of FM modulation on the harmonic distribution around the ninth harmonic.

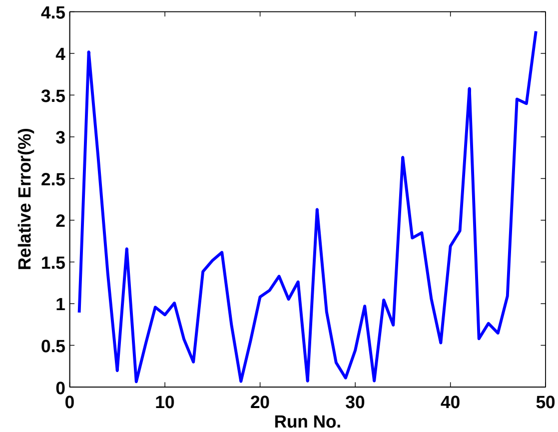


Figure 5: Error of the response surface method for 7 level of parameters.

A frequency window around the 1st harmonic(around 100 kHz in this example) is considered and the maximum level of the harmonic inside this window, is considered as the level of the first harmonic.

Experiments are designed based on the Latin cube concept for 5 factors (i.e., the inductance, the capacitance, the load impedance and parameters of FM). Experiments are designed for several levels of factors.

Table 3: The effects for 11 levels of the parameters (effects are multiplied to 1000).

$Factor$	1	L	C	R_t	f_m
$Effect$	-0.1918	0.3900	-0.0213	0.0113	0.0063
$Factor$	β	L^2	LC	LR_t	Lf_m
$Effect$	0.0039	-0.1891	0.0002	-0.0058	-0.0010
$Factor$	$L\beta$	C^2	CR_t	Cf_m	$C\beta$
$Effect$	-0.0011	0.0092	-0.0015	0.0006	-0.0003
$Factor$	R_tf_m	$R_t\beta$	$f_m\beta$		
$Effect$	-0.0032	-0.0004	-0.0025		

Experiments show that by 10 percent variations around the typical values, 19.2 dB variations in the harmonic level can be observed. Therefore slight variations of circuit and FM parameters cause large variations of harmonic levels.

Figure 5 shows the error of response surface for the experiments with 7 level of parameters. Table 2 shows the effect and the interactions. It is seen that the error do not exceed 7 percent. Experiments for the 11 and 21 level give similar results (see Fig. 6, Fig. 7, Table 3 and Table 4). Experiments show that considering the effect of reactive elements by the first order terms causes error in the response surface. Therefore in the response surface the effects of reactive elements is considered with second order terms. It can be seen that increase in the levels of parameters, does not reduce the computational error and with the least experiments, the effect of parameter variations on the harmonic level can be discussed.

Table 4: The effects for 21 levels of the parameters (effects are multiplied to 1000).

$Factor$	1	L	C	R_t	f_m
$Effect$	-0.1913	0.3913	-0.0152	0.0069	0.0048
$Factor$	β	L^2	LC	LR_t	Lf_m
$Effect$	0.0015	-0.1907	-0.0019	-0.0034	-0.0004
$Factor$	$L\beta$	C^2	CR_t	Cf_m	$C\beta$
$Effect$	-0.0001	0.0071	-0.0003	-0.0003	-0.0002
$Factor$	R_tf_m	$R_t\beta$	$f_m\beta$		
$Effect$	0.0026	-0.0003	-0.0013		

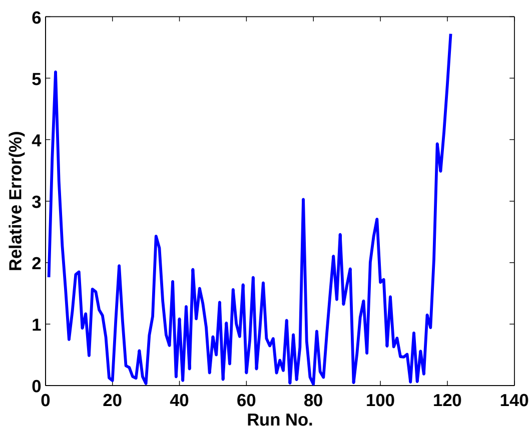


Figure 6: Error of the response surface method for 11 level of parameters.

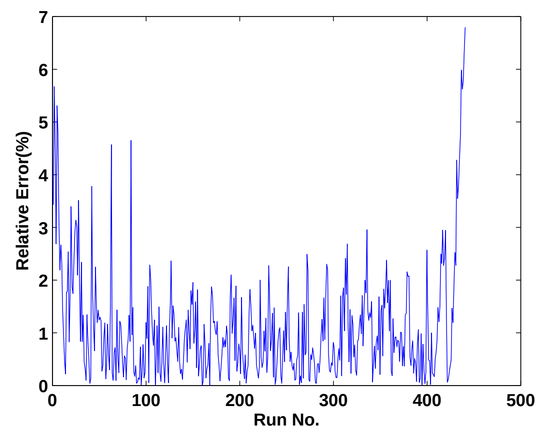


Figure 7: Error of the response surface method for 21 level of parameters.

Investigations of the effects in the Tables 2, 3 and 4 show that the harmonic level is more sensitive to the reactive elements ant the load than the parameters of modulation. Therefore in the design for low EMI, reactive elements must be chosen with care. Although FM plays undeniable role in harmonic level reductions, variation of its parameters has minor effect on the level of the harmonic.

6. CONCLUSION

In this paper, the sensitivity of interference signal due to the first harmonic of the switching power supplies with the FM modulation of switching signal is investigated with respect to the circuit and FM modulation parameters. In this regard, sets of orthogonal experiments are designed based on the Latin Cubes concept. Then the results of the experiments (i.e., the level of the 1st harmonic) are interpolated on a response surface. The response surface shows the self and the mutual effects of the factors on the level of the 1st harmonic. It is seen that the slight changes of the circuit and modulation parameters, causes large increases of the level of the 1st harmonic. Consequently, after circuit design process, sensitivity analysis must be done with respect to the circuit parameters and with the small changes in the circuit parameters; the level of 1st harmonic can be reduced. It is seen that the response surface can be readily used for analysis of the effects of parameters and with the minor experiments, suitable results are achievable on the level of the 1st harmonic. The harmonic level is more sensitive to the reactive elements ant the load than the parameters of modulation. Therefore in the design for low EMI, reactive elements must be chosen with care. Dominant interference depends on the load. With due regard to the positive sign of the effect of the load in the response surface, increase the load increase the harmonic level.

REFERENCES

1. Wilson, T. G., "The evolution of power electronics," *IEEE Transactions on Power Electronics*, Vol. 15, No. 3, 439–446, May 2000.
2. Ohashi, H., I. Omura, S. Matsumoto, Y. Sato, H. Tadano, and I. Ishii, "Power electronic innovation with next generation advanced power devices," *IEICE Transactions on Communications*, Vol. E87-B, 3422–3429, Dec. 2004.
3. Ferreira, J. A. and J. D. V. Wyk, "Electromagnetic energy propagation in power electronic converters: toward future electromagnetic integration," *Proceeding of the IEEE*, Vol. 89, No. 6, 876–889, June 2001.
4. Sengupta, D. L. and V. V. Liepa, *Applied Electromagnetics and Electromagnetic Compatibility*, Wily Interscience, 2006.
5. Balcells, J., A. Santolaria, A. Orlandi, D. Gonzalez, and J. Gago, "EMI reduction in switched mode power converters using frequency modulation techniques," *IEEE Transactions on Electromagnetic Compatibility*, Vol. 47, No. 3, 569–576, August 2005.
6. Tse, K. K., H. S. Chung, S. Y. Hui, and H. C. So, "A comparative study of carrier frequency modulation techniques for conducted EMI suppression in PWM converters," *IEEE Transactions on Power Electronics*, Vol. 49, No. 3, 618–627, June 2002.
7. Lee, Y. and R. Mittra, "Electromagnetic interference mitigation using a spread spectrum approach," *IEEE Transactions on Electromagnetic Compatibility*, Vol. 44, No. 2, 380–385, May 2002.
8. Tse, K. K., H. S. Chung, S. Y. Hui, and H. C. So, "Analysis and spectral characteristics of a spread spectrum technique for conducted EMI suppression," *IEEE Transactions on Power Electronics*, Vol. 15, No. 2, 399–410, March 2000.
9. Paramesh, J. and A. Von Jouanne, "Use of sigma-delta modulation to control EMI from switch mode power supplies," *IEEE Transactions on Industrial Electronics*, Vol. 48, No. 1, 111–117, Feb. 2001.
10. Tse, K. K., R. Wai-Man, H. S. Chung, and S. Y. R. Hui, "An evaluation of the spectral characteristics of switching converters with chaotic carrier frequency modulation," *IEEE Transactions on Industrial Electronics*, Vol. 50, No. 1, 171–182, Feb. 2000.
11. Banerjee, S., A. L. Baranovski, J. L. R. Marrero, and O. Woywode, "Minimizing electromagnetic interference problems with chaos," *IEICE Transactions on Fundamentals*, Vol. E87-A, No. 8, 2100–2109, August 2004.
12. Dugue, D. and M. Girault, *Analyse de Variance et Planse d' Experience*, Dounod Paris, 1969.
13. Brisset, S., F. Gillon, S. Vivier, and P. Brochet, "Optimization with experimental design: an approach using Taguchies methodology and finite element simulations," *IEEE Transactions on Magnetics*, Vol. 37-A, No. 5, 3530–3533, Sep. 2001.

Beam-wave Coupling in a Double-beam Gyrotron Traveling Wave Amplifier

Chong-Qing Jiao[1,2] **and Ji-Run Luo**[1]

[1]Institute of Electronics, Chinese Academy of Sciences
P. O. Box 2652, Beijing 100080, China
[2]Graduate School of the Chinese Academy of Sciences
Beijing 100039, China

Abstract— A linear theory is used to study the beam-wave coupling in a double-beam gyrotron traveling wave amplifier (Gyro-TWA). Numerical results show that the beam-wave coupling in this type of amplifier excites six waves, which include one constant-amplitude backward wave, and five forward waves consisting of two growing waves, two decaying waves and one constant-amplitude wave. The growing wave, whose growth rate is much larger than that of the other, has a larger operating bandwidth which is approximately the superposition of the interaction bandwidths of the two single-beams in the Gyro-TWA.

1. INTRODUCTION

High power and broad bandwidth capability of the gyrotron traveling wave amplifier makes it an attractive source of coherent radiation in the millimeter and sub-millimeter wavelength ranges. Gyro-TWAs are applied very broadly, especially in radar and communication systems. Steady progress in theory and experiment has been made over an extended period of time. Major advances in Gyro-TWA performances have been reported in the review papers [1–3].

The Gyro-TWA interaction comes from the coupling between the electron beam cyclotron mode and the waveguide mode, which respectively satisfy

$$\omega - k_z v_z - s\Omega = 0 \tag{1}$$

$$\omega^2 = c^2 k_{mn}^2 + c^2 k_z^2 \tag{2}$$

where, $\Omega = eB_0/\gamma m_0$ is the cyclotron frequency of the electron with a relativistic factor γ in a uniform static magnetic field B_0, k_{mn} is the cut-off wave number of the waveguide mode and s is the cyclotron harmonic number. The operating region of a Gyro-TWA normally locates near the grazing intersection where the dispersion line of the beam mode is tangential to that of the waveguide mode.

The concept of using two electron beams in a Gyro-TWA was proposed in Ref. [4] to increase the bandwidth of the amplifier. In such amplifier, two beam mode lines corresponding to the two beams are tangential to the same waveguide mode dispersion curve at different points (see Fig. 1). By choosing appropriately the operating parameters of the two beams, the operating regions of the two beams can be merged into a wider one and therefore wideband operation can be realized. It should be pointed out that one of the key points for the realization of double-beam Gyro-TWA is the generation of two beams with different energy and velocity, which was ever presented in Ref. [4].

Although large-signal numerical calculations had been done for a double-beam amplifier in Ref. [4], the small-signal analysis of such amplifier is still helpful for providing us clear physical insights for the double-beam Gyro-TWA interaction and giving us effective guidance for time-consuming large-signal simulations. In this paper, the small-signal characteristics of the double-beam Gyro-TWA are analyzed based on the linear theory reported in Ref. [5]. The rest of this paper is organized as follows. The model and analytic theory are presented in Section 2. Results and discussion are made in Section 3. This work is summarized in the final section.

2. MODEL AND ANALYTIC THEORY

For the double-beam Gyro-TWA, supposing that the electron beam consists of two annular cold beams, the equilibrium distribution function of the double-beam is the sum of these of the two single-beam and written as [5, 6]

$$f_0 = \sum_{j=1}^{2} \frac{I_{bj}}{2\pi r_{bj} e v_{zj}} \delta\left(r_b - r_{bj}\right) \delta\left(p_z - p_{zj}\right) \frac{1}{2\pi p_{tj}} \delta\left(p_t - p_{tj}\right) \tag{3}$$

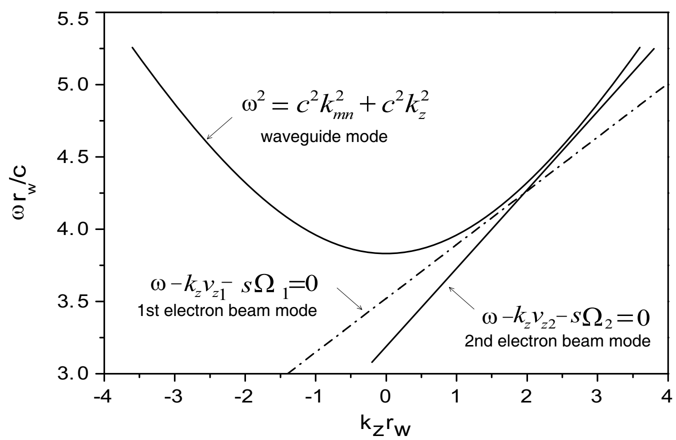


Figure 1: The waveguide mode (the cured line) and the two electron beam modes (the solid and dotted straight line) for the double-beam Gyro-TWT.

where, I_{bj}, r_{bj}, p_{zj} and p_{tj} are, respectively, the beam current, guiding center radius, axial and transverse momentum of the jth beam. Substituting Eq. (3) into the general dispersion relation of gyro-TWT interaction [5, 6], the dispersion relation of the double-beam Gyro-TWA with TEmn as its operating mode can be expressed in the SI unit system as [5–7]

$$D(\omega,\, k_z) = \frac{\omega^2}{c^2} - k_{mn}^2 - k_z^2 + \frac{e\mu_0}{\pi r_w^2 m_0 K_{mn}} \sum_{j=1}^{2} \frac{I_{bj}}{\gamma_j v_{zj}} \times \left(\frac{\beta_{tj}^2 \left(\omega^2 - k_z^2 c^2\right)}{\left(\omega - s\Omega_j - k_z v_{zj}\right)^2} H_{sm}\left(k_{mn} r_{bj},\, k_{mn} r_{Lj}\right) \right.$$
$$\left. - \frac{\left(\omega - k_z v_{zj}\right)}{\left(\omega - s\Omega_j - k_z v_{zj}\right)} T_{sm}\left(k_{mn} r_{bj},\, k_{mn} r_{Lj}\right) + \frac{k_{mn} v_{tj}}{\left(\omega - s\Omega_j - k_z v_{zj}\right)} U_{sm}\left(k_{mn} r_{bj},\, k_{mn} r_{Lj}\right) \right) = 0 \quad (4)$$

where k_z is the axial wave number, $k_{mn} = x_{mn}/r_w$, x_{mn} is the nth root of $J_m'(x)$, $K_{mn} = \left(1 - m^2/x_{mn}^2\right) J_m^2(x_{mn})$, r_w is the waveguide radius, r_L is the electron Larmor radius, $v_{zj} = p_{zj}/\gamma m_0$, $v_{tj} = p_{tj}/\gamma m_0$, $\beta_{zj} = v_{zj}/c$, and $\beta_{tj} = v_{tj/c}$. The definition of H_{sm}, T_{sm}, and U_{sm} are provides as follows:

$$\begin{aligned} H_{sm}(x,\, y) &= J_s'^2(y) J_{m-s}^2(x) \\ T_{sm} &= 2H_{sm} + yJ_s'(y) \left\{ 2J_s''(y) J_{m-s}^2(x) \right. \\ &\quad \left. - J_s(y) \left[\frac{1}{x} J_{m-s}'(x) J_{m-s}(x) + J_{m-s}''(x) J_{m-s}(x) + J'^2_{m-s}(x) \right] \right\} \\ U_{sm} &= -\frac{y}{2} J_s'(y) \left\{ J_{s+1}(y) \left(J_{m-s-1}^2(x) - J_{m-s}^2(x) \right) + J_{s-1}(y) \left(J_{m-s+1}^2(x) - J_{m-s}^2(x) \right) \right\} \end{aligned} \quad (5)$$

It can be seen that the dispersion relation of the double-beam Gyro-TWT (Eq. (4)) is a sixth-order polynomial on k_z, which has six roots corresponding to six waves. The evolution of the RF field amplitude along the waveguide axis can be expressed by the superimposition of the six waves, and is given by

$$f(z) = -i \sum_j \frac{f(0) N\left(k_{zj}\right) + f'(0)}{D'\left(k_{zj}\right)} e^{-ik_{zj} z} \quad (6)$$

where, $f(0)$ and $f'(0)$ are the amplitude of the wave and its first derivative at $z = 0$, where the electron beam enters the waveguide. k_{zj} is the jth root of the dispersion relation Eq. (4).

$$D'\left(k_z\right) = \frac{dD\left(k_z\right)}{dk_z} \quad (7)$$

$$N\left(k_z\right) = i \frac{e\mu_0}{\pi m_0 r_w^2 K_{mn}} \sum_{j=1}^{2} \frac{I_{bj}}{v_{zj}\gamma_j} \left(\frac{v_{zj} T_{sm}\left(k_{mn} r_{bj},\, k_{mn} r_{Lj}\right)}{\left(\omega - s\Omega_j - k_z v_{zj}\right)} - \frac{v_{tj}^2 k_z H_{sm}\left(k_{mn} r_{bj},\, k_{mn} r_{Lj}\right)}{\left(\omega - s\Omega_j - k_z v_{zj}\right)^2} \right) - ik_z \quad (8)$$

The power gain along the waveguide axis can be calculated by

$$G(z) = 10 \log_{10} \left(\frac{\mathrm{Im}\left(f(z) f'^{*}(z)\right)}{\mathrm{Im}\left(f(0) f'^{*}(0)\right)} \right) \tag{9}$$

3. RESULTS AND DISCUSSIONS

The operating parameters of the double-beam Gyro-TWA to be analyzed are as follows [4]: TE_{01} mode, $s = 1$, $r_w = 5.7\,\mathrm{mm}$ (corresponding to the cutoff frequency of 32.1 GHz), $V_1 = 90\,\mathrm{kv}$, $V_2 = 153\,\mathrm{kV}$, $\alpha_1 = 1.0$, $\alpha_2 = 0.621$, $I_{b1} = I_{b2} = 10A$, $r_{b1} = 0.75 r_w$, and $r_{b2} = 0.5 r_w$. Where $\alpha_i (i = 1 \,\mathrm{or}\, 2)$ is beam velocity ratio of the ith electron beam and the selected beam voltages and velocity ratios could keep the two beam mode lines are simultaneous tangential to the waveguide mode curve with same magnetic field B_0.

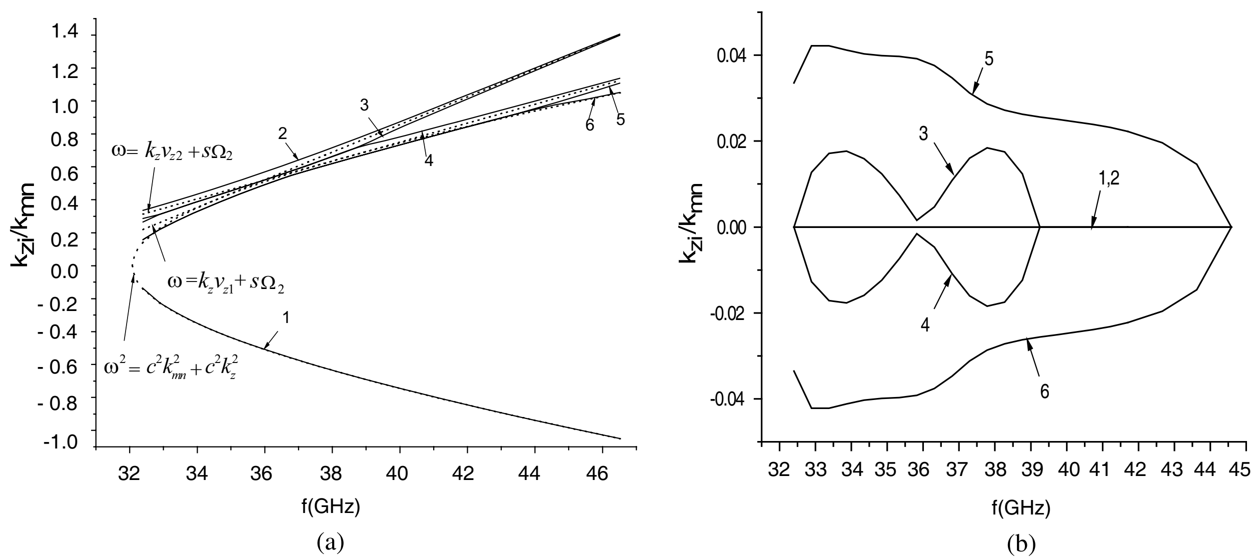


Figure 2: The real part (a) and imaginary part (b) of the axial wave number versus frequency for the double-beam Gyro-TWT at $B_0 = B_g$. The three dotted curves in (a) represent one waveguide mode and two beam modes respectively.

Figure 2 illustrates the roots of the dispersion relation of the above Gyro-TWA as a function of frequency with $B_0 = B_g$ (where B_g is the grazing magnetic field at which the beam mode is tangent to the waveguide mode). These roots include one positive real root, one negative real root, and two pairs of complex conjugates. The real parts (k_{zr}) are plotted in Fig. 2(a), representing the phase constants of the waves. Fig. 2(b) shows the imaginary parts (k_{zi}), corresponding to the spatial growth rates of the waves. It can be seen from Fig. 2 that the interaction among the two electron beams and the waveguide mode results in six waves, consisting of one constant-amplitude backward wave ("1" in Fig. 2) and five forward waves. Among the forward waves, there are one constant-amplitude wave ("2"), two growing waves ("3"and "5"), and two decaying waves ("4" and "6"). Further analysis by comparing the axial wave numbers of the six waves and the three cold modes, which include one waveguide mode and two beam modes (represented respectively by the three dotted curves in Fig. 2(a)), shows that the backward wave may come from the waveguide mode because their axial wave numbers are very close to each other; the constant-amplitude forward wave has the largest axial wave number in the six waves; the pair of conjugate waves ("3" and "4") may be caused by the coupling of the two beam modes because the values of their wave numbers are between these of the two beam modes; similarly, the two conjugate waves ("5" and "6") may result from the coupling among the waveguide mode and the two beam modes. The wave "5", whose growth rate is the much larger than that of the wave "3", has a larger bandwidth.

Figure 3 shows that growth rate of the wave "5" in the double-beam Gyro-TWT is close to that of the first beam in low-frequency region and close to that of the second beam in high-frequency region. Moreover, the bandwidth of the wave "5" is approximately the superposition of the bandwidths generated by the two single-beams in the Gyro-TWT.

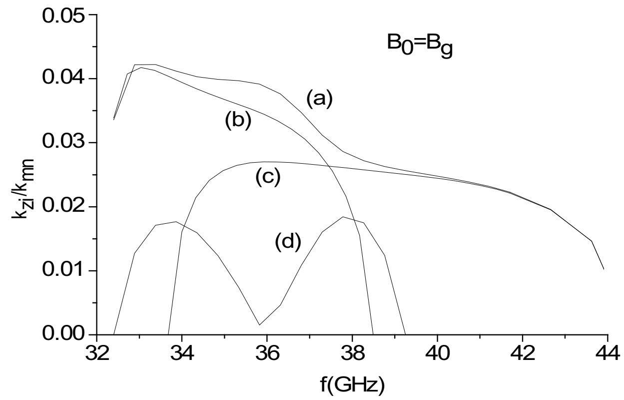


Figure 3: Growth rate as function of frequency. Curve (a) and (d) correspond to, respectively, the waves "5" and "3" in Fig. 2. Curve (b) corresponds to the first beam in the Gyro-TWT. Curve (c) corresponds to the second beam.

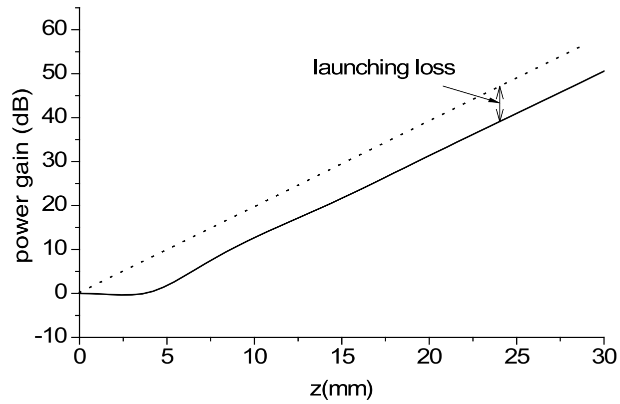


Figure 4: The spatial evolution of the power gain of a 36.4 GHz wave in the double-beam Gyro-TWT with $B_0 = 0.99B_g$.

Figure 4 is plotted to demonstrate the launching loss of the double-beam Gyro-TWT. In this figure, the solid curve represents the evolution of power gain of the RF field along the waveguide axis, and the dashed curve shows the growth of the RF field if the input signal is assumed to be totally coupled into the wave "5". The difference between the two curves is defined as the launching loss as indicated in Fig. 4. The solid curve illustrates that the RF field starts to exponentially grow after a certain interaction length. Obviously, this length is approximately the distance for the wave "5" starting to become dominant in the six waves. Therefore, the linear growth rate of the amplifier can be represented by the growth rate of the wave "5". It can be seen that the launching loss for the double-beam amplifier is about -9.54 dB, as is similar to the traditional Gyro-TWT with single-beam.

4. CONCLUSIONS

The beam-wave coupling in a double-beam Gyro-TWT leads to the generations of six waves, which include one constant-amplitude backward wave, and five forward waves consisting of two growing waves, two decaying waves and one constant-amplitude wave. The growth rate of one of the forward growing waves, which is much larger than that of the other, can be used to represent the linear growth rate of the amplifier. The double beam can obviously increase the bandwidth of a Gyro-TWT. The launching loss of the double-beam Gyro-TWT is also about -9.54 dB.

REFERENCES

1. Felch, K. L., B. G. Danly, H. R. Jory, et al., "Characteristics and applications of fast-wave gyro-devices," *Proceedings of the IEEE*, Vol. 87, No. 5, 752–781, 1999.
2. Chu, K. R.,"Overview of research on the gyrotron traveling-wave amplifier," *IEEE Trans. Plasma Sci.*, Vol. 30, No. 3, 903–908, 2002.
3. Granatstein, V. L., B. Levush, B. G. Danly, and R. K. Parker, "A quarter century of gyrotron research and development," *IEEE Trans. Plasma Sci.*, Vol. 25, No. 6, 1322–1334, 1997.
4. Yang, Y. and D. Wu, "A two-stream gyrotron traveling wave tube amplifier," *Phys. Plasmas*, Vol. 6, No. 11, 4328–4332, 1999.
5. Kou, C. S., Q. S. Wang, D. B. McDermott, et al., "High-power harmonic gyro-TWAT'spart I: linear theory and oscillation study," *IEEE Trans. Plasma Sci.*, Vol. 20, No. 3, 155–162, 1992.
6. Chu, K. R., "Gain and bandwidth of the Gyro-TWT and CARM amplifiers," *IEEE Trans. Plasma Sci.*, Vol. 16, No. 2, 90–104, 1988.
7. Fliflet, A. W., "Linear and non-linear theory of the Doppler-shifted ycylotron resonance maser based on TE and TM waveguide modes," *Int. J. Electronics*, Vol. 61, No. 6, 1049–1080, 1986.

BOR-FDTD Analysis of Spherical Lens Multi-beam Antenna

Y. H. Li and W. B. Dou
State Key Laboratory of Millimeter Waves, Southeast University
Nanjing 210096, China

Abstract— The multi-beam antenna system has important applications in the aeronautics and astronautics field, which has caught much attention in recent years. This paper analyze a multi-beam antenna consisted of spherical lens and multi feeds. Firstly, the single feed's performance of the spherical-lens antenna is analyzed. In view of the axial symmetry of the antenna, the BOR-FDTD method is used to reduce the three-dimensional solution to a two-dimensional one, and the number of the variable required by the solution will decrease greatly. As a result, it reduces the requirement of computer memory and makes it possible to calculate the electrically large object. The results calculated for the electrically small antenna agree well with that of the HFSS simulation, which proves the method is effective. Finally, the computation results for the electrically large antenna are presented.

1. INTRODUCTION

The multi-beam antenna system has important applications in the aeronautics and astronautics field, for example, multi-beam antenna has been used widely in satellite communications. There are several structures of multi-beam antenna such as reflector, multi-elements array and lens antennas [1].

Lens antenna has good optic properties and no obstruction. When it is used in millimeter and sub-millimeter wavelengths, the disadvantage of larger mass has been reduced or eliminated. So using a lens to get multi-beam antenna at millimeter waves is a suitable approach.

The researches to spherical lens have lasted for several decades. The most conventional method used is geometrical optics method (GO). However, GO is only an approximation method. Recently, based on numerical method, electrical larger lens may be analyzed by full wave method. In this paper, BOR-FDTD (Body-of-Revolution Finite-Difference-Time-Domain) method is used to analyze an electrical larger spherical lens antenna at millimeter wavelengths, and calculation results are presented.

2. BOR-FDTD FORMULA

Analysis for BOR has emerged since 1960's. BOR-FDTD is used by Dennis W. Prather and Shouyuan Shi in 1999 to deal with binary lens [3]. In cylindrical coordinate, permittivity $\varepsilon_r(\rho, z)$ and permeability $\mu_r(\rho, z)$ of BOR has no variance along circle direction Φ. If excite source is also symmetric of revolution, the fields E_ρ, E_Φ, E_z, H_ρ, H_Φ and H_z can be expanded with Fourier series as follows:

$$E_\eta(\rho,\phi,z,t) = \sum_{k=0}^{\infty} E1_{\eta,k}(\rho,z,t)\cos(k\phi) + \sum_{k=0}^{\infty} E2_{\eta,k}(\rho,z,t)\sin(k\phi) \tag{1}$$

$$H_\eta(\rho,\phi,z,t) = \sum_{k=0}^{\infty} H1_{\eta,k}(\rho,z,t)\cos(k\phi) + \sum_{k=0}^{\infty} H2_{\eta,k}(\rho,z,t)\sin(k\phi) \tag{2}$$

where $\eta = \rho, \Phi, z$, $k = 0, 1, 2, \ldots$, $E1_{\eta,k}$, $E2_{\eta,k}$ and $H1_{\eta,k}$, $H2_{\eta,k}$ are parameters to be determined. Because of the orthogonality the unknowns can be obtained by solving problem for each mode.

For the situation we considered, Maxwell's rotation equations can be written as

$$\mu\frac{\partial H2_{\rho,k}}{\partial t} = \frac{k}{\rho}E1_{z,k} + \frac{\partial E2_{\phi,k}}{\partial z} \tag{3.1}$$

$$\varepsilon\frac{\partial E1_{\rho,k}}{\partial t} + \sigma E1_{\rho,k} = \frac{k}{\rho}H2_{z,k} - \frac{\partial H1_{\phi,k}}{\partial z} \tag{3.2}$$

$$\mu\frac{\partial H1_{\phi,k}}{\partial t} = -\frac{\partial E1_{\rho,k}}{\partial z} + \frac{\partial E1_{z,k}}{\partial \rho} \tag{3.3}$$

$$\varepsilon\frac{\partial E2_{\phi,k}}{\partial t} + \sigma E2_{\phi,k} = \frac{\partial H2_{\rho,k}}{\partial z} - \frac{\partial H2_{z,k}}{\partial \rho} \tag{3.4}$$

$$\mu\frac{\partial H2_{z,k}}{\partial t} = -\frac{1}{\rho}\frac{\partial(\rho E2_{\phi,k})}{\partial \rho} - \frac{k}{\rho}E1_{\rho,k} \tag{3.5}$$

$$\varepsilon\frac{\partial E1_{z,k}}{\partial t} + \sigma E1_{z,k} = \frac{1}{\rho}\frac{\partial(\rho H1_{\phi,k})}{\partial \rho} - \frac{k}{\rho}H2_{\rho,k} \tag{3.6}$$

Then difference form is used instead of (3.1)∼(3.6), BOR-FDTD form can be obtained. Thus 3 dimensional problem is reduced to 2 dimensional problem and can be solved in PC. Singularity at origin can be solved by using Faraday law [3].

Configuration for BOR-FDTD is shown in Fig. 1, lens and feed is symmetric of revolution. The radius of the Lens is R, the distance between lens and feed is d, relative dielectric constant of lens is 2.1, the relative permeability is 1, and frequency is 37.5 GHz. PML is used to truncate the computing region. The feed is waveguide end, which is excited by TE_{11} mode, so $k = 1$ in Equations (3.1)–(3.6).

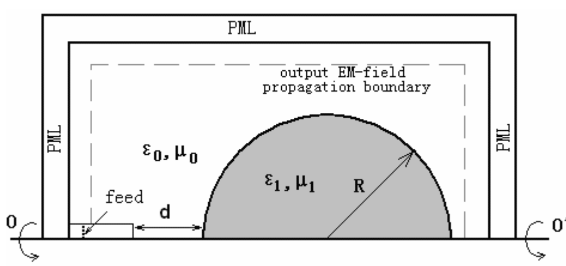


Figure 1: Configuration for BOR-FDTD.

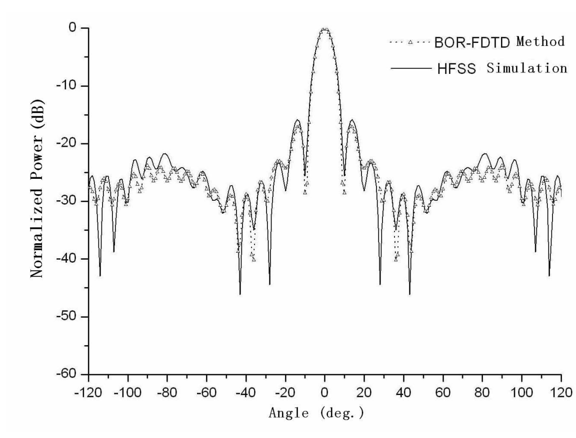


Figure 2: E-plane radiation pattern ($d = 8$ mm, $R = 30$ mm).

3. CALCULATION RESULTS

Firstly, the radiation pattern of a relative electrical small lens antenna (lens radius is 30 mm) is computed by BOR-FDTD and the results are compared with that of HFSS, which is given in Fig. 2. Good agreement has been observed, indicating the correcting of our method. However, the calculation time HFSS is beyond 3 time of BOR-FDTD. In order to get higher directivity, lens radius is increased form 30 mm into 50 mm, for this lens size, HFSS is unable to calculate. Fig. 3 and 4 show the calculation results of radiation pattern for different d. From Fig. 3 and 4, it can be seen that when d/R is near 0.40 the side lobe level (SLL) of the lens antenna may get its minimum.

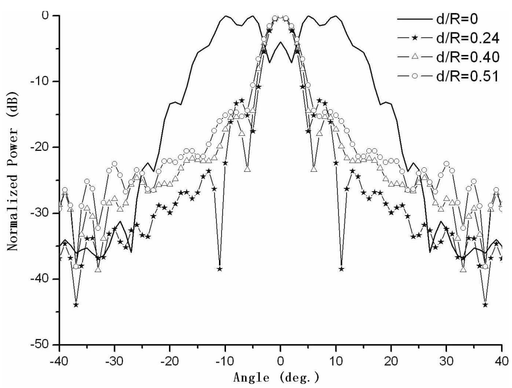


Figure 3: E-plane radiation pattern ($R = 50$ mm).

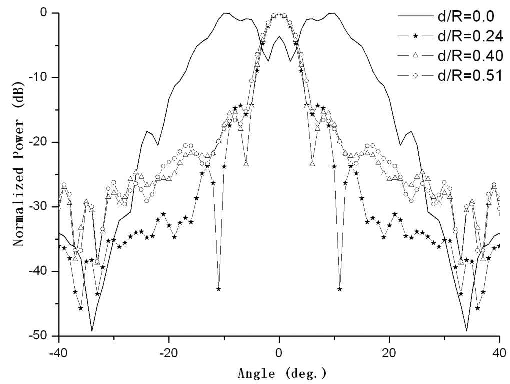


Figure 4: H-plane radiation pattern ($R = 50$ mm).

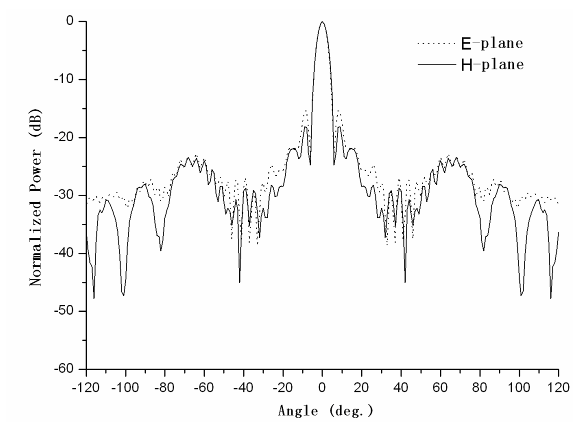


Figure 5: Radiation pattern for $d/R = 0.398$.

Figure 5 shows the radiation pattern for $d/R = 0.398$. Where half power beam width is 5.2°, SLL in E plane is -15.4 dB, in H plane is -18.2 dB.

4. CONCLUSION

BOR-FDTD is used to analyze an electrical larger spherical lens antenna. It reduces the problem from 3 dimensions to 2 dimensions. Radiation pattern of lens antenna for different d/R are presented, and good performance is obtained when d/R is near 0.4.

ACKNOWLEDGMENT

This work is supported partly by NSFC(60571028).

REFERENCES

1. Xie, C. J. and H. Z. Wang, "Reviews on multi beam antennas," *Space Science and Technology of China*, (in Chinese), 1995.
2. Ge, D. B., *FDTD of Electromagnetic Waves*, (in Chinese), Xidian University Press.
3. Prather, D. W. and S. Shi, "Formulation and application of the finite-difference time-domain method for the analysis of axially symmetric diffractive optical element," *J. Opt. Soc. Am.*, Vol. 16, No. 5, 1999.
4. Schoenlinner, B., X. Wu, J. P. Ebling, G. V. Eleftheriades, and G. M. Rebeiz, "Wide-scan spherical-lens antennas for automotive radars," *IEEE Trans. on MTT*, Vol. 50, No. 9, 2002.
5. Andreasen, M. G., "Scattering from bodies of revolution," *IEEE Trans. Antennas Propag.*, Vol. AP-13, 303–310, 1965.

Rolled Dipole Antenna for Low-resolution GPR

A. A. Lestari[1,2], D. Yulian[3], A. B. Suksmono[1,4], E. Bharata[1,4]
A. G. Yarovoy[2], and L. P. Ligthart[2]

[1]International Research Centre for Telecom and Radar–Indonesian Branch (IRCTR-IB), Indonesia
[2]International Research Centre for Telecom and Radar (IRCTR)–TU Delft, The Netherlands
[3]Radar and Communication Systems (RCS), Indonesia
[4]Bandung Institute of Technology (ITB), Indonesia

Abstract— In this paper a rolled dipole antenna for low-resolution impulse ground penetrating radar (GPR) is theoretically investigated. The antenna is designed for transmission of monocycle pulses with duration of 5 ns (200 MHz central frequency) suitable for low-resolution GPR applications. The dipole is rolled to considerably reduce it length and resistive loading with Wu-King profile is applied for suppression of late-time ringing important for GPR. Using NEC-2 as the numerical tool it is shown that the antenna radiates the pulse with no late-time ringing. Furthermore, by rolling the wires the antenna length is reduced by a factor of 4 with no evident negative impact on the antenna's characteristics.

1. INTRODUCTION

Ground penetrating radar (GPR) with low resolution is useful for detection of large buried objects or structures such as subsurface caves, bunkers and rivers or canals. Low-resolution GPR generally transmits relatively long transient pulses with central frequency of 200 MHz or lower for obtaining low down-range resolution and sufficient penetration depth required for detecting large subsurface structures. Consequently, large antennas are normally needed for efficient transmission of those pulses. If the dimensions of the antennas are reduced, generally one suffers from degraded efficiency and increased late-time ringing as a result of smaller antenna bandwidth. Minimal late-time ringing is important to avoid masking of radar returns when the targets are shallowly buried. In this paper we investigate a method to substantially reduce the length of a dipole antenna for low-resolution GPR by rolling the wires. Resistive loading with Wu-King profile is then applied to eliminate late-time ringing. Similar approach has been used in [1] to increase antenna bandwidth by slightly rolling the edges of a V-dipole antenna for high-resolution GPR. Here, nearly the whole length of the dipole is rolled resulting in a spiral-like shape.

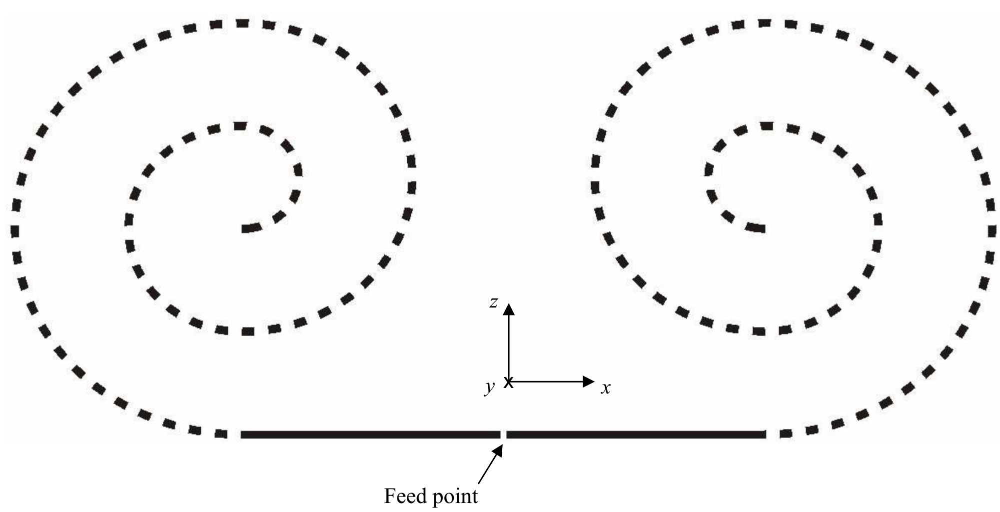


Figure 1: Geometry of the rolled dipole. The gaps are the locations of resistors for antenna loading. In practice the resistors are soldered across the gaps. The total length of the wire is 296 cm while the length of the rolled dipole is 75 cm, giving a reduction factor of nearly 4. Its height is 31.5 cm. The dipole will be realized as a printed antenna on an FR-4 substrate.

2. ANTENNA DESIGN

In this work the antenna is designed for excitation with monocycle pulses with duration of 5 ns (having a central frequency of 200 MHz) suitable for low-resolution GPR applications. The geometry of the proposed rolled dipole antenna is shown in Figure 1. The total length of the wire is 296 cm and by rolling the wire as shown in the figure the antenna length is reduced by a factor of nearly 4, resulting in antenna length of only 75 cm. The fraction of the wires starting from 20 cm from the feed point is resistively loaded with resistors according to the Wu-King profile. In each arm of the dipole we employ 65 resistors with 1 cm separation. This number of resistors should be sufficient for proper implementation of the Wu-King profile. The first resistor near the feed point has a value of 200 Ω and at the same time also functions as a secondary source of radiation due to the discontinuity it introduces. It has been demonstrated in [2] that when the distance between the feed point and the first resistor is chosen to be $c/(4f_c\sqrt{\varepsilon_{rs}})$, where c is the speed of light, f_c is the central frequency of the exciting pulse and ε_{rs} is the relative permittivity of the substrate, in the broadside direction of the antenna radiation from the secondary source combines constructively with radiation from the feed point, resulting in significant increase in the amplitude of the transmitted pulses.

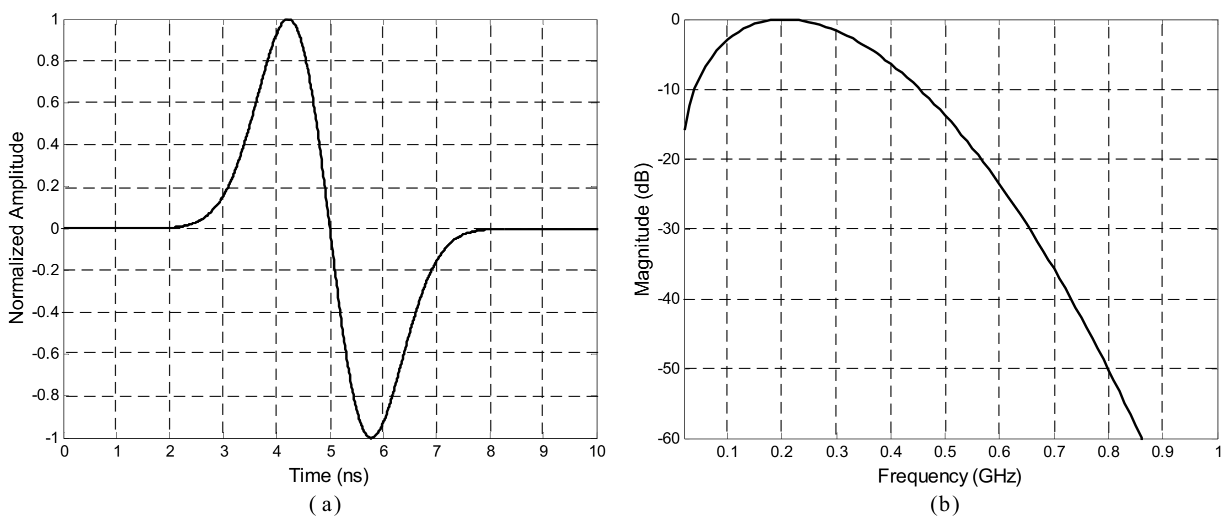


Figure 2: Exciting pulse: monocycle with duration of 5 ns. (a) waveform, (b) spectrum.

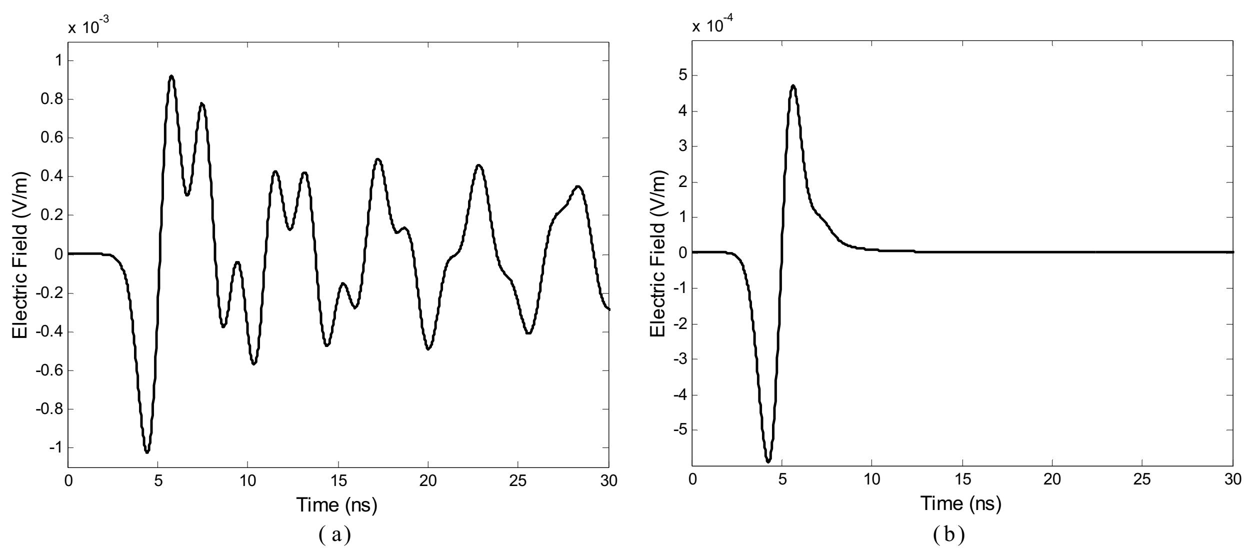
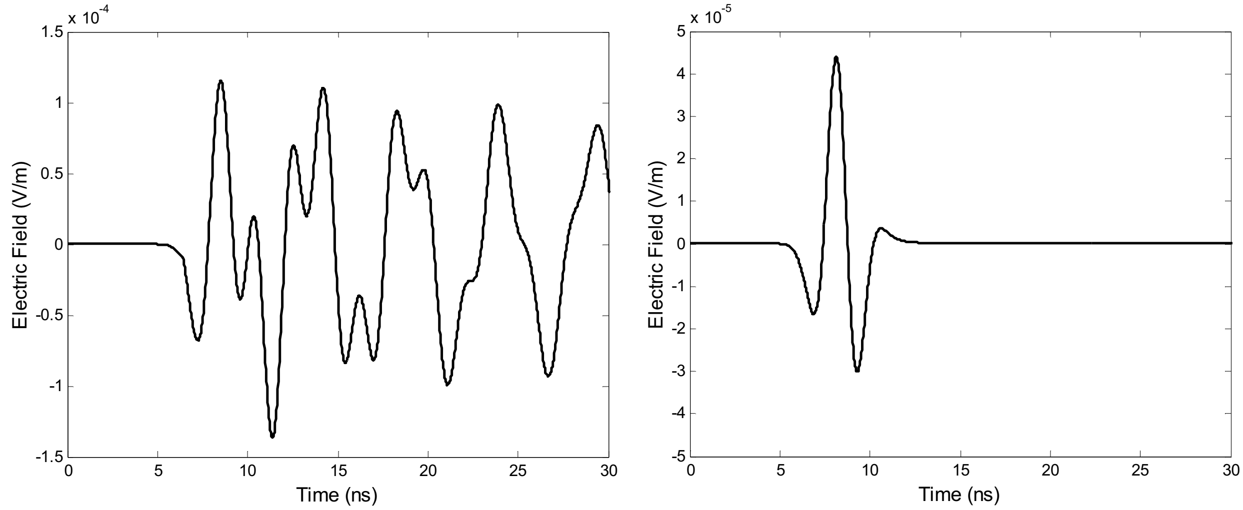


Figure 3: Near-field transmit waveforms of the rolled dipole in free space: (a) without loading, (b) with loading (Wu-King profile). Observation point is at a distance of 50 cm in the broadside direction from the antenna.

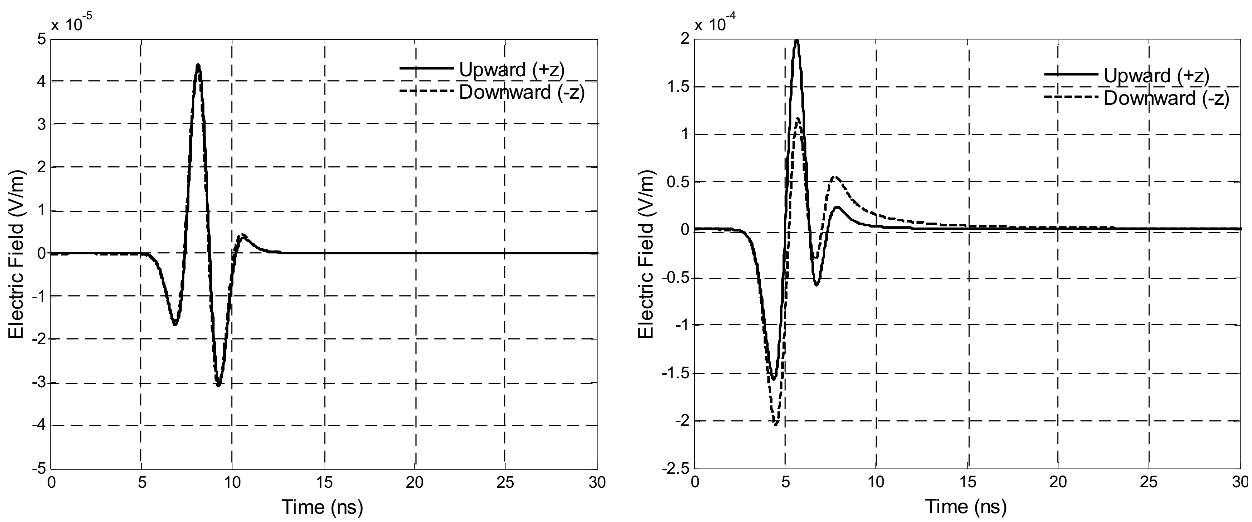
3. SIMULATIONS

The proposed design shown in Figure 1 will be realized as a printed antenna on an FR-4 substrate. By means of the equivalent radius formula introduced in [3] the antenna structure can be approximated with thin wires, for which the one-dimensional integral equation method is well-suited to perform numerical analysis. Therefore, the well-known NEC-2 code is here employed. Influence of the substrate is taken into account by scaling up the antenna dimensions by a factor of the square root of the substrate's relative permittivity. Time-domain responses of the antenna are obtained by performing frequency sweep over the spectrum of the 5-ns monocycle pulse followed by inverse FFT operation. The frequency sweep was performed from 1 MHz to 1 GHz with 1 MHz step. In Figure 2 the waveform and spectrum of the pulse are presented.



(a) (b)

Figure 4: Far-field transmit waveforms of the rolled dipole in free space: (a) without loading, (b) with loading (Wu-King profile). Observation point is at a distance of 5 m in the broadside direction from the antenna.

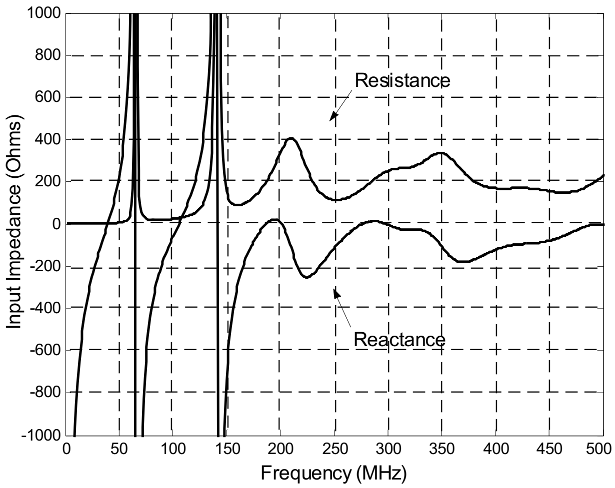


(a) (b)

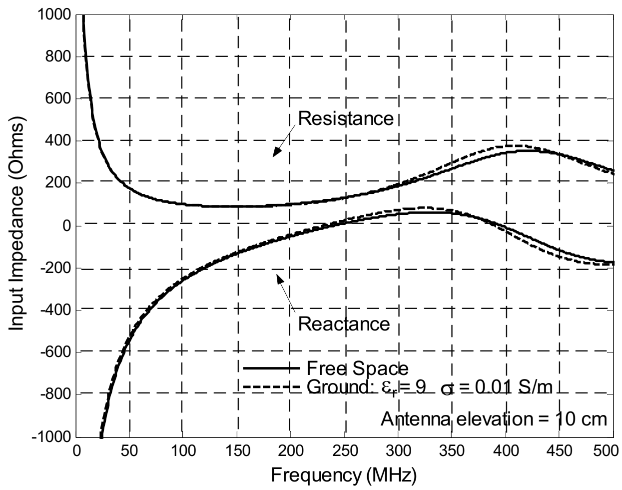
Figure 5: (a) Far-field and (b) near-field waveforms transmitted by the rolled dipole in the broadside direction, downwards (into the ground) and upwards (into the air).

Transmit waveforms of the antenna have been computed and the near-field waveforms are plotted in Figure 3. The near-field waveforms of the antenna without and with loading are given in Figures 3(a) and 3(b), respectively. It is shown that when the antenna is not loaded the waveform is dominated by internal reflections as expected. We note that in the near-field region the first reflection from the antenna ends is still clearly separated from the main pulse radiated from the feed point. When resistive loading with Wu-King profile is applied, it is demonstrated in Figure 3(b) that all reflections are eliminated, leaving the main pulse only. In the far-field region the main pulse and the first end reflection are not anymore separable as observed in Figure 4(a). When resistive loading with Wu-King profile is applied, similarly internal reflections are eliminated, leaving the main pulse with the shape of a triplet as seen in Figure 4(b). Furthermore, in Figure 5 we demonstrate that the influence of the rolled part of the wires on transmit waveforms is noticeable only in the near-field region.

The computed input impedance of the antenna is presented in Figure 6. It is demonstrated that the resonance behavior seen in Figure 6(a) is suppressed by the loading. Moreover, we observe in Figure 6(b) that soil has insignificant impact on the input impedance when the antenna is elevated at least 10 cm as generally is the case in real low-resolution GPR surveys.



(a)



(b)

Figure 6: Free-space input impedance of the rolled dipole: (a) without loading, (b) with loading (Wu-King profile). In (b) the input impedance of the rolled dipole situated 10 cm above a lossy soil is also shown, with the soil parameters: $\varepsilon_r = 9$, $\sigma = 0.01\,\mathrm{S/m}$.

Currently we are working on feeding and impedance matching techniques for the proposed antenna. In addition, an optimal shield for the antenna is being designed. These works and experimental verifications of the antenna in laboratory and field conditions will be reported in our future papers.

REFERENCES

1. Morrow, I. L., J. Persijn, and P. van Genderen, "Rolled edge ultra-wideband dipole antenna for GPR application," *IEEE APS Int. Symp. Digest*, Vol. 3, 484–487, 2002.
2. Lestari, A. A., A. G. Yarovoy, and L. P. Ligthart, "*RC* loaded bow-tie antenna for improved pulse radiation," *IEEE Trans. Antennas Propagat.*, Vol. 52, No. 10, 2555–2563, Oct. 2004.
3. Butler, C. M., "The equivalent radius of a narrow conducting strip," *IEEE Trans. Antennas Propagat.*, Vol. 30, No. 4, 755–758, July 1982.

Viability of Convex-modulated Exponential Serraions for Improved Performance of CATRs

T. Venkata Rama Krishna[1], **P. Siddaiah**[1], **and B. Prabhakara Rao**[2]

[1]Department of Electronics and Communication Engg., K L College of Engineering, Guntur, A. P., India
[2]Department of Electronics and Communication Engg., College of Engineering
Jawaharlal Nehru Technological University, A. P., India

Abstract— Compact Antenna Test Range Reflector (CATR) offers many advantages over other types of ranges and consequently a lot of effort is being directed towards the improvement of compact range performance. CATR performance can be further improved by using serrations. Serrating the edges of the reflector in order to redistribute the diffracted rays in many directions away from the test zone area. Application of serrated edges has been shown to be a good method to control diffraction at the edges of the reflectors. Serrated edge CATR is used to reduce the diffraction effect. However, in order to get some insight into the positive effect of serrated edges a less rigorous analysis technique known as Physical Optics (PO) may be used. In this paper, the sides of the reflector are considered to be equipped with convex-modulated exponential serrations. Fresnel zone field is evaluated at a distance of 64λ from the reflector surface. Ripple free and enhanced quiet zone width are observed for specific values of width modulation factors.

1. INTRODUCTION

The antenna is the interface between the transmission lines of our design and the unconstrained regions of space. To achieve a uniform field at the test antenna site with in the small distance compact antenna test range method is most appropriate. A quiet zone is desired over which the amplitude is constant to a fraction of dB. Using hybrid serrated geometry can reduce diffraction effect at the test site [1–4].

2. METHOD OF ANALYSIS

Beeckman has presented the PO analysis of a serrated edge reflector to predict the Fresnel zone field [5]. A square aperture reflector of $45\lambda \times 45\lambda$ is equipped with convex-modulated exponential serrations as shown in Figure 1. A recourse is taken to decompose the aperture area S into three parts S_1, S_2 and S_3 such that $S = S_1 + S_2 - S_3$ as shown in Figure 2. A quasi-analytical expression can be derived for the Fresnel field [6–12]. The boundary functions $g^+(y')$ and $g^-(y')$ are expressed as a Fourier series of convex serrations with rate of rise/fall 'a' and $h^+(x')$ and $h^-(x')$ are described as the Fourier series of modulated exponential serrations.

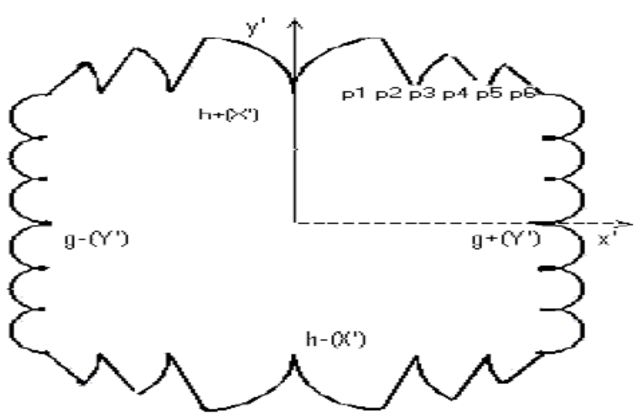


Figure 1: Square aperture reflector with convex-modulated exponential serrations.

2.1. Fourier series of convex serrations

$$g^+(y') = \frac{a_0}{2} + t\left[e^{(-ap/2)+1}\right] + (2t/ap)\left[e^{(-ap/2)-1}\right] - 4tapZ(y') \tag{1}$$

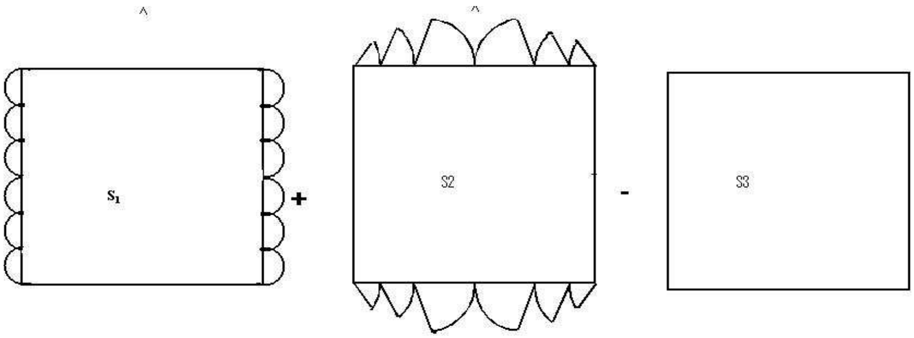


Figure 2: Decomposition of Figure 1.

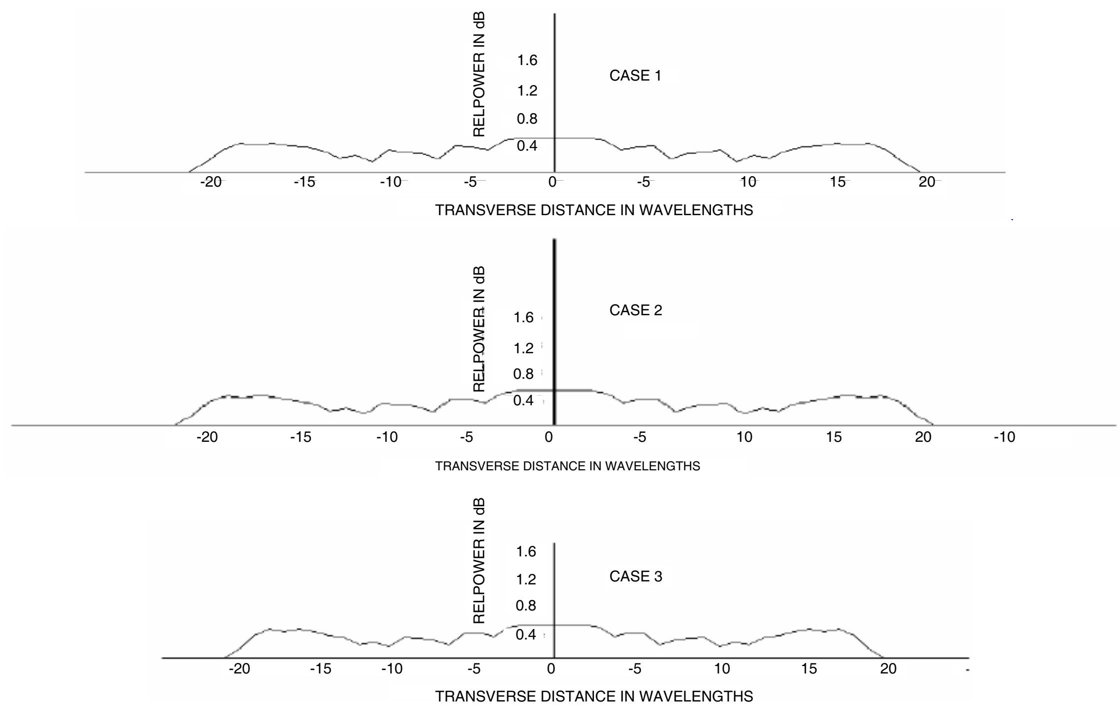


Figure 3: Fresnel zone field for $45\lambda \times 45\lambda$ convex-modulated exponential serrated CATR for cases1, 2 and 3.

where $Z(y') = \sum_{n=1}^{\infty} \frac{1 - e^{(-ap/2)\cos(n\Pi)}}{4n^2\Pi^2 + a^2p^2} \cos(2\Pi n y'/p')$

2.2. Fourier series of modulated exponential serrations

$$h^+(x') = \frac{a_0}{2} + \frac{t_1}{p_6}\left[\frac{(p_2+p_1)}{2} + \frac{1}{a_1}\left(e^{(-a_1p_1)} - 1\right)\right] + \frac{t_2}{p_6}\left[\frac{(p_4+p_3-2p_2)}{2} + \frac{1}{a_2}\left(e^{(-a_2(p_3-p_2))} - 1\right)\right]$$
$$+ \frac{t_3}{p_6}\left[\frac{(p_6+p_5-2p_4)}{2} + \frac{1}{a_3}\left(e^{(-a_3(p_5-p_4))} - 1\right)\right] + \frac{2}{p_6}Z(x') \qquad (2)$$

where

$$
\begin{aligned}
Z(x') = &\sum_{n=1}^{\infty} t_1 \left\{ \frac{\sin(q_1)}{q} - e^{(-a_1 p_1) b_1} (-a_1 \cos(q_1) + q(\sin(q_1))) + a_1 b_1 \right\} \\
&- \frac{t_1}{(p_2 - p_1)} \left\{ \left(\frac{1}{q^2} \right) (\cos(q_2) - \cos(q_1)) + \left(\frac{\sin(q_1)}{q} (p_2 - p_1) \right) \right\} \\
&+ t_2 \left\{ (\sin q_3 - \sin q_2)/q - e^{(-a_3(p_3 - p_2))} b_2(-a_2 \cos q_3 + q(\sin q_3)) + b_2(-a_2 \cos q_2 + q(\sin q_2)) \right\} \\
&- \frac{t_2}{(p_4 - p_3)} \{ (\cos q_4 - \cos q_3)/q^2 + (\sin q_3/q)(p_4 - p_3) \} \\
&+ t_3 \left\{ (\sin q_5 - \sin q_4)/q - e^{(-a_3(p_5 - p_4))} b_3(-a_3 \cos q_5 + q(\sin q_5)) + b_3(-a_3 \cos q_4 + q(\sin q_4)) \right\} \\
&- \frac{t_3}{(p_6 - p_5)} \{ 1/q^2((-1)^n - \cos q_5) + (\sin q_5/q)(p_6 - p_5) \} \cos(qy')
\end{aligned}
$$

where $q = n\pi/p_6$; $q^i = qp_i$; $b_i = \dfrac{1}{a_i^2 + q^2}$

The above Fourier series representations in conjunction with analytical expression of Beeckman [5] gives the Fresnel zone field.

Table 1: Width modulation factors.

CASE	P	P_1/p	P_2/p	P_3/p	P_4/p	P_5/p	P_6/p
1	$(a_0/2)/30$	4	6	12.67	15.33	25.33	30
3	$(a_0/2)/22.5$	3	4	10	12	20	22.5
6	$(a_0/2)/7$	0.75	1	2.5	3	6	7

3. RESULTS AND CONCLUSION

Fresnel field calculations are made at a distance of 64λ along the z-axis. The variation of relative power in dB with transverse distance in wavelengths is furnished in Figure 3. From figure it is observed that, by proper selection of width modulation factors (Table 1) lesser ripple and enhanced quiet zone width are observed. It is noticed that ripple free quiet zone is obtained at the center.

ACKNOWLEDGMENT

The authors place on record their grateful thanks to the authorities of the K L College of Engineering, India and Jawaharlal Nehru Technological University, College of Engineering, India for providing the facilities.

REFERENCES

1. Evans, G. E., *Antenna Measurement Techniques*, Artech House, Inc., 1990.
2. Srivastava, K. K., *IEEE Workshop on Design, Measurements and Evaluation of Microwave Antenna Systems*, New Delhi, September 28–30, 1978.
3. Hollis, J. S., T. J. Lyon, and L. Clayton, *Microwave Antenna Measurements*, Scientific Atlanta, Inc., Atlanta, Georgia, USA, November 1985.
4. Kummer, W. H. and E. S. Gillespie, "Antenna measurements-1978," *Proc. IEEE*, Vol. 66, No. 4, 483–507, April 1978.
5. Beeckman, P. A., "Prediction of the fresnel region field of a compact antenna test range with serrated edges," *IEE Proc.*, Vol. 133, Pt. H. No. 2, 108–114, April 1986.
6. Siddaiah, P. and P. V. Subbaiah, "Performance of compact antenna test range reflectors employing width and height modulated triangular serrations," *Journal of the Institution of Engineers*, ET, Vol. 84, India, July 2003.
7. Siddaiah, P. and P. V. Subbaiah, "Viability of on-off triangular and on off rectangular serrations for improved performance of compact antenna test ranges," *AMSE Journal*, Vol. 76, No. 6, Modeling A, France, 2003.

8. Siddaiah, P. and P. V. Subbaiah, "Combination of convex and concave serrations for improved performance of compact antenna test range reflectors," *NCSSS-2002, PSG College of Technology*, Coimbattore, 1–2 March 2002.
9. Siddaiah, P. and P. V. Subbaiah, "Performance augmentation of CATRs using width modulated exponential and width and height modulated triangular serrations," *MMTA-2002*, School of Environmental Sciences, JNU, New Delhi, 4–6 February 2002.
10. Siddaiah, P. and P. V. Subbaiah, "Analysis of width and height modulated exponential serrated compact antenna test range reflectors," Recent Trends in Communication Technology, Jiwaji University, Gwalior, 13–14 April, 2002.
11. Siddaiah, P. and P. V. Subbaiah, "Performance evaluation of WHME and WME serrated CATR," Recent Trends in Communication Technology, Jiwaji University, Gwalior, 13–14 April, 2002
12. Siddaiah, P. and P. V. Subbaiah, "A comparison between WMET & WHMET serrated CATRs," *IETE Journal of Research*, Vol. 49, No. 4, July–August 2003.
13. Siddaiah, P., P. V. Subbaiah, and T. V. R. Krishna, "Performance comparison between HMCE & IMCE serrated CATRs," Solid Sate Physics Laboratory, Lucknow Road, DRDO, New Delhi, Feb. 2004.

Antenna Design for Ultra Wideband Application Using a New Multilayer Structure

Yashar Zehforoosh, **Changiz Ghobadi**, and **Javad Nourinia**
Department of Electrical Engineering, Urmia University, Iran

Abstract— As wireless communication applications require more and more bandwidth, the demand for wideband antennas increases as well. For instance, the ultra wideband radio (UWB) utilizes the frequency band of 3.1–10.6 GHz.This paper presents a work carried out within ULTRAWAVES in the area of antenna design and analysis. A new multilayer microstrip antenna is introduced using Stacked Multiresonator patches. In order to achieve suitable bandwidth, the antenna size is fine for mobile applications. The antenna is designed, optimized and simulated using Ansoft designer. In addition results and conclusions are presented.

1. INTRODUCTION

Ultra Wideband Radio (UWB) is a potentially revolutionary approach to wireless communication in that it transmits and receives pulse based waveforms compressed in time rather than sinusoidal waveforms compressed in frequency [1]. This is contrary to the traditional convention of transmitting over a very narrow bandwidth of frequency, typical of standard narrowband systems such as 802.11a, b, and Bluetooth. This enables transmission over a wide swath of frequencies such that a very low power spectral density can be successfully received [2]. The recent allocation of the 3.1–10.6 GHz frequency spectrum by the Federal Communications Commission (FCC) for Ultra Wideband radio applications has presented a myriad of exciting opportunities and challenges for antenna designers [3]. Pulsed UWB, by definition, refers to any radio or wireless device that uses narrow pulses (on the order of a few nanoseconds or less) for sensing and communication. This requires sufficient impedance matching, proper return loss and VSWR<2 throughout the entire bandwidth. In this paper a new low profile, small stacked multiresonator microstrip antenna is presented for UWB application. The bandwidth of a microstrip antenna increases with an increase in substrate thickness and decreases in the dielectric constant also, the bandwidth of the antenna increases when multiresonators are coupled in planar or stacked configurations. In this paper we use three patches placed in the bottom layer, and multiresonators taken on the top layers. One of the bottom patches is excited by a coaxial-fed. Ease of construction, suitable radiation pattern and better characteristics are advantages of this antenna over the previously presented antennas.

2. ULTRA WIDEBAND COMMUNICATION

Ultra-Wideband (UWB) technology has been around since the 1980s, but it has been mainly used for radarbased applications until now, because of the wideband nature of the signal that results in very accurate timing information. However, due to recent developments in high-speed switching technology, UWB is becoming more attractive for lowcost consumer communications applications. ULTRA-WIDEBAND communications involves the transmission of short pulses with a relatively large fractional bandwidth [4] and [5]. More specifically, these pulses possess a -10 dB bandwidth which exceeds 500 MHz or 20% of their center frequency [6] and is typically on the order of one to several gigahertz. The Federal Communications Commission's Report and Order (R&O), issued on Feb. 2002, defines UWB as any signal that occupies more than 500 MHz in the 3.1–10.6 GHz band and that it meets the spectrum mask shown in Figure 1. A comparison with the other unlicensed bands currently available in the US is shown in Table 1.

This definition, which replaces the previous one that expressed UWB in terms of fractional bandwidth, opened up a new way of thinking for several leaders in the UWB community.

Given the recent spectral allocation and the new definition of UWB adopted by the FCC, UWB is not considered a technology anymore, but available spectrum for unlicensed use. This means that any transmission signal that meets the FCC requirements for UWB spectrum can be considered UWB technology. This, of course, is not just restricted to impulse radios or high speed spread spectrum radios pioneered by companies so far, but to any technology that utilizes more than 500 MHz spectrum in the allowed spectral mask and with the current emission limit's restrictions.

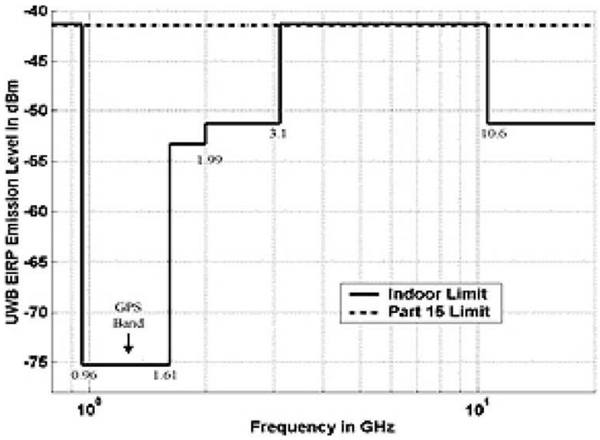


Figure 1: UWB spectral mask for indoor communication systems [1].

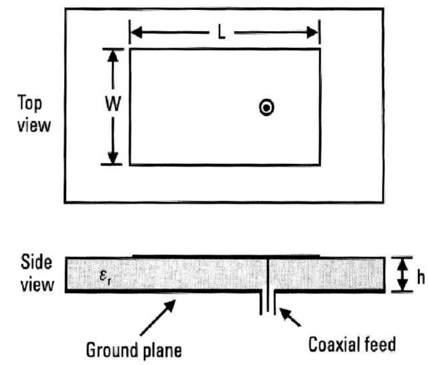


Figure 2: Top and side views of RMSA.

Table 1: Comparison between bands.

Unlicensed bands	Frequency of operation	Bandwidth
ISM at 2.4 GHz	2.4000-2.4835	83.5 MHz
U-NII at 5 GHz	5.15-5.35 GHz 5.75-5.85 GHz	300 MHz
UWB	3.1-10.6 GHz	7,500 MHz

3. MICROSTRIP ANTENNAS

3.1. History of MSAs

Deschamps first proposed the concept of microstrip antenna (MSA) in 1953 [8]. However practical antennas were developed by Munson [9, 10] and Howel [11] in 1970. The numerous advantages of MSA, such as its low weight, small volume, and ease of fabrication using printed circuit technology, led to the design of several configurations for various applications [12–15]. With increasing requirements for personal and mobile communications, the demand for smaller and low-profile antennas has brought the MSA on the forefront. An MSA in its simplest form consists of a radiating patch on one side of a dielectric substrate and a ground plane on the other side. The top and side views of a rectangular MSA (RMSA) are shown in Figure 2. However, other shapes such as the square, circular, triangular, semicircular, sectoral, and annular ring shapes are also used.

3.2. Characteristics of MSAs

The MSA has proved to be an excellent radiator for many applications because of its several advantages, but it also has some disadvantages.

The main advantages of MSAs are listed as follows:

- They are lightweight and have a small volume and low profile planar configuration.
- They can be made conformal to the host surface.
- Their ease of mass production using printed-circuit technology leads to a low fabrication cost.
- They are easier to integrate with other MICs on the same substrate.
- They allow both linear polarization and circular polarization.
- They can be made compact for use in personal mobile communication.
- They allow for dual-band and triple frequency operations.

MSAs suffer from some disadvantages as compared to conventional microwave antennas. They are the following:

- Narrow BW.
- Lower gain.

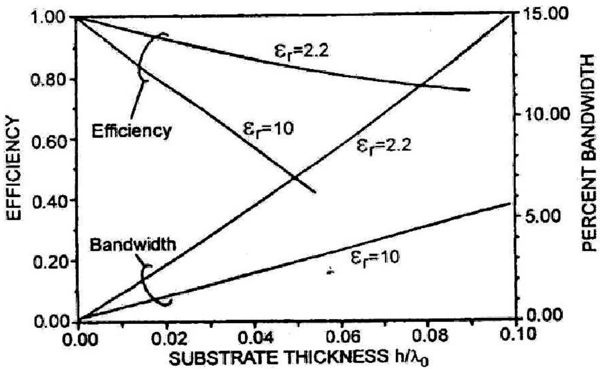


Figure 3: Effect of substrate thickness and dielectric constant on the impedance BW (VSWR<2) and radiation efficiency[12].

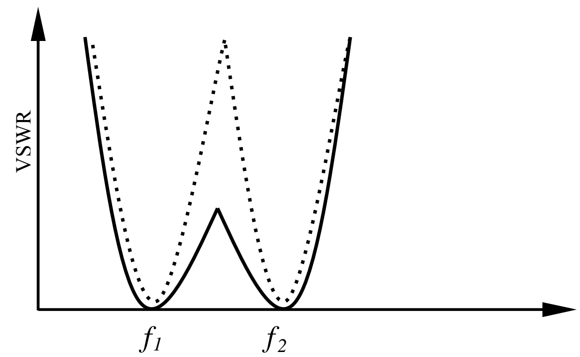


Figure 4: VSWR plot of two coupled resonators having narrow bandwidth(...) individual resonators and(—).

- Low power-handling capacity.

MSAs have narrow BW, typically 1–5%, which is the major limiting factor for the widespread application of these antennas. Increasing the bandwidth of MSAs has been the major thrust of research in this field.

3.2.1. Definition of BW

The VSWR or impedance BW of the MSA is defined as the frequency range over which it is matched with that of the feed line within specified limits. The BW of the MSA is inversely proportional to its quality factor Q and is given by [16].

$$\mathrm{BW} = \frac{\mathrm{VSWR} - 1}{Q\sqrt{\mathrm{VSWR}}}$$

where VSWR is defined in terms of the input reflection coefficient Γ as:

$$\mathrm{VSWR} = \frac{1 + |\Gamma|}{1 - |\Gamma|}$$

The Γ is a measure of reflected signal at the feed point of the antenna. It is defined in terms of input impedance Z_{in} of the antenna and the characteristic impedance Z_0 of the feed line as given below:

$$\Gamma = \frac{Z_{in} - Z_0}{Z_{in} + Z_0}$$

The BW is usually specified as frequency range over which VSWR is less than 2 (which corresponds to a return loss of 9.5 dB or 11% reflected power). Some times for stringent applications, the VSWR requirement is specified to be less than 1.5 (which corresponds to a return loss of 14 dB or 4% reflected power). In this paper we use VSWR<2.

3.2.2. Effects of Substrate Parameters on BW

Impedance BW of a patch antenna varies inversely as quality factor Q of the patch antenna. Therefore substrate parameters such as dielectric constant and thickness can be varied to obtain different Q, and ultimately the increase in impedance BW. Q of a resonator is defined as

$$Q = \frac{\text{energy stored}}{\text{power lost}}$$

Figure 3 shows the effect of substrate thickness on impedance BW and efficiency for two values of dielectric constants. Note that the BW increases monotonically with thickness. Also, a decrease ε_r in value increases the BW. This behavior can be explained from the change in Q value.

In conclusion, we can say that the increase in h and decrease in ε_r can be used to increase the impedance BW of the antenna. However, this approach is helpful up to $h \leq 0.02\lambda$ only. The disadvantages of using thick and high dielectric constant substrates are many, including these:

- Surface wave power increases, resulting in poor radiation efficiency.
- The radiation from surface waves may lead to pattern degradation near end-fire.
- Thick substrates with microstrip edge feed will give rise to increased spurious radiation from the microstrip step-in-width and other discontinuities. Radiation from the probe feed will also increase.
- Substrates thicker than 0.11λ for $\varepsilon_r = 2.2$ makes the impedance locus of the probe fed patch antenna increasingly inductive in nature, resulting in impedance matching problems.
- Higher order modes along the thickness may develop, giving rise to distortions in the radiation patterns and impedance characteristics. This is a limiting factor in achieving an octave BW.

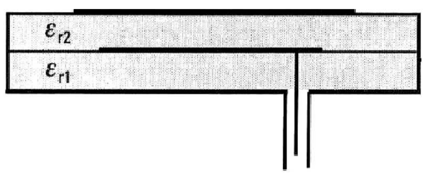


Figure 5: Multilayer MSA.

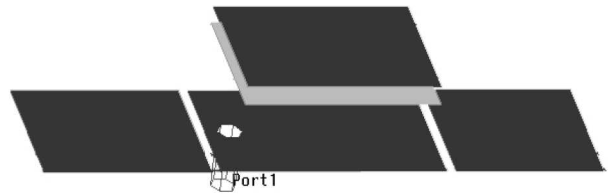


Figure 6: The geometry of the antenna.

Table 2: Dimensions of the antenna.

L	L_1	L_2	L_3	W	h_1	h_2	h_3	ε_{r1}	ε_{r2}	ε_{r3}	x
18 mm	15.4 mm	13.4 mm	11.7 mm	10.8 mm	1.2 mm	5 mm	4 mm	2.1 mm	2.1 mm	2.1 mm	7 mm

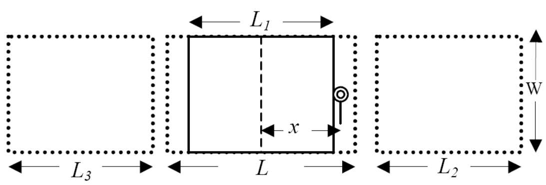


Figure 7: Top view of the antenna.

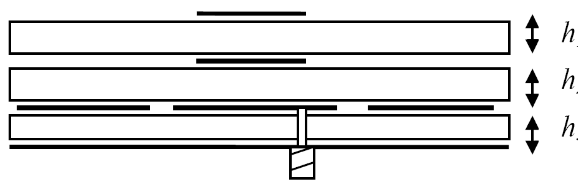


Figure 8: Side view of the antenna.

4. EFFECT OF PARASITIC PATCHES

A patch placed close to the fed patch gets excited through the coupling between the patches [4]. Such a patch is known as a parasitic patch. If the resonance frequencies f_1 and f_2 of these two patches are close to each other, then broad bandwidth is obtained as shown in Figure 4. The overall input VSWR will be the superposition of the responses of the two resonators resulting in a wide bandwidth [7, 8]. If the bandwidth is narrow for the individual patch, then the difference between f_1 and f_2 should be small and If the bandwidth of the individual patch is large, then the difference in the two frequencies should be large to yield an overall wide bandwidth.

5. MULTILAYER CONFIGURATIONS

In the multilayer configuration, two or more patches on different layers of the dielectric substrate are stacked on each other. Based on the coupling mechanism, these configurations are categorized as electromagnetically coupled or aperture coupled MSA.

In the electromagnetically coupled MSA, one or more patches at the different dielectric layers are electromagnetically coupled to the feed line located at the bottom dielectric layer as shown in Figure 5. Alternatively, one of the patches is fed by a coaxial probe and the other patch is electromagnetically coupled. The patches can be fabricated on different substrates, and accordingly the patch dimensions are to be optimized so that the resonance frequencies of the patches are close to each other to yield broad BW. These two layers may be separated by either air gap or foam [8].

The multilayer broadband MSAs, unlike single layer configurations, show a very small degradation in radiation pattern over the complete VSWR BW. The drawback of these structures is the increased height; which is not desirable for conformal applications and increased back radiation.

Planar and stacked multiresonators techniques are combined to yield a wide bandwidth with a higher gain. In this paper we use two different configurations for this type of antennas.

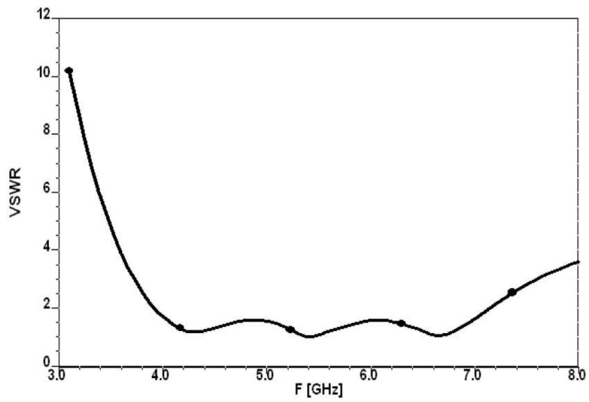


Figure 9: VSWR plot of antenna.

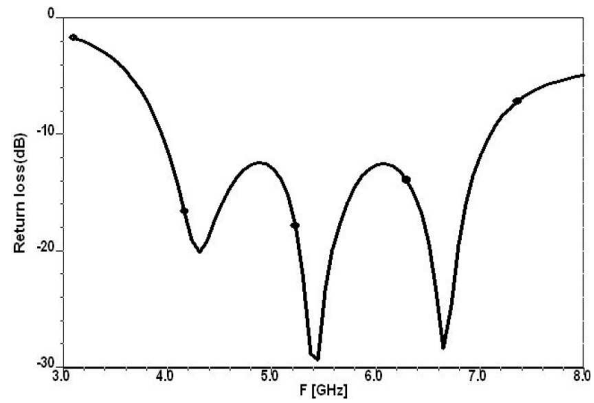


Figure 10: Return loss plot of antenna.

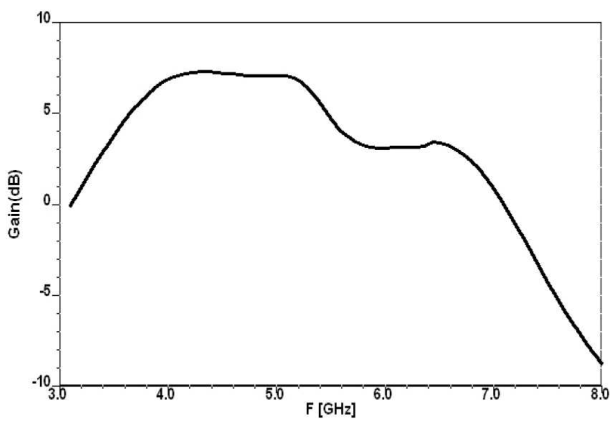


Figure 11: Gain plot of antenna.

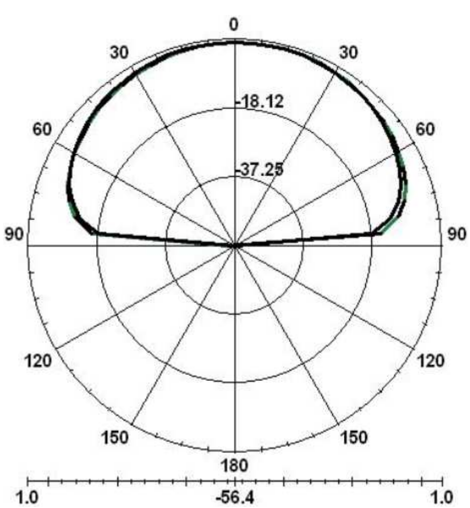


Figure 12: Radiation pattern of antenna.

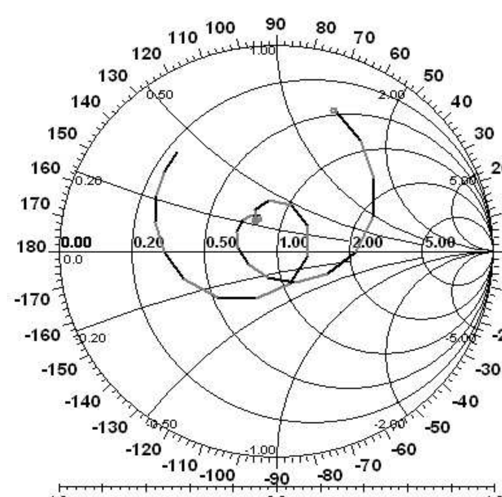


Figure 13: Smith chart of antenna.

6. ANTENNA DESIGN

As it is shown in Figure 6 planar and stacked multiresonators techniques are combined to yield a wide bandwidth with a higher gain.

The antenna has three rectangular patches at the bottom and two patches on the top layers exciting the bottom patch by coaxial feed. The two top patches are the same in size but the two patches beside the exited patch are different in size.

Only the bottom patch is fed and the other patches electromagnetically coupled as shown in Figure 6.

In Figure 7 the top view and in Figure 8 the side view of the antenna are shown. The patch on the bottom layer is shown in dotted lines and the patches on the top layer are shown in solid lines. Because of the multilayer configuration of the antenna and its especial structure the number of parameters that are to be optimized is increased. Referring to the antenna geometry from Figures 7 and 8, the dimensions of the antenna are presented in Table 2.

The antenna is designed, optimized and simulated using Ansoft designer software. The bandwidth obtained for the antenna is 3.25 GHz. The radiation is in the broad side direction, and the variation in the pattern is very small over the entire bandwidth. At 4.3 GHz, the gain is 7.5 dB. As shown in figures 9-13 the bandwidth and return loss are proper for ultra wideband applications and antenna dimensions are suitable for mobile devices.

7. CONCLUSIONS

In this paper, a new small microstrip antenna for ultra wideband applications is designed, optimized simulated. There was a great success in finding a suitable structure for mobile applications. Also obtaining bandwidth about 50% and maximum gain about 7.5 dB shows that this structure can be mentioned as a useful design for ultra wideband products. However acquired results show that the antenna design and structure need more refinement in order to achieve the ultimate design with a smaller physical profile and better performance.

ACKNOWLEDGMENT

The authors would like to thank Iran telecommunication research center (ITRC) for its financial supports.

REFERENCES

1. First Report and Order (FCC 02-48), Action by the Commission February 14, 2002. New Public Safety Applications and Broadband internet access among uses envisioned by FCC authorization of Ultra-Wideband Technology.
2. James, J. R. and P. S. Hall, *Handbook of Microstrip Antennas*, Vol. 1, Peter Peregrinus Ltd., London, 1989.
3. Derneyd, A. G. and A. G. Lind, "Extended analysis of rectangular microstrip resonator antennas," *IEEE Trans. Antennas Propagation*, Vol. AP-27, 846–849, November 1979.
4. James, J. R., P. S. Hall, and C. Wood, *Microstrip Antennas Theory and Design*, Peter Peregrinus, London, 1981.
5. Taylor, J. D., *Introduction to Ultra-Wideband Systems*, CRC Press, Ann Arbor, MI, 1995.
6. Astanin, L. Y. and A. A. Kostylev, *Ultrawideband Radar Measurements Analysis and Processing*, IEE, London, U.K., 1997.
7. Kumar, G. and K. P. Ray, "Stacked gap-coupled multiresonator rectangular microstrip antennas," *IEEE AP-S Int.Symp. Digest*, Bostan, MA, 514–517, July 2001,
8. Deschamps, G. A., "Microstrip microwave antennas," *Proc. 3rd USAF symposium on antennas*, 1953.
9. Munson, R. E., "Single slot cavity antennas assembly," U.S Patent No. 3713462, January 23, 1973.
10. Munson, R. E., "Conformal microstrip antennas and microstrip phased arrays," *IEEE Trans. Antennas Propagate*, Vol. AP-22, 74–78, 1974.
11. Howell, Q. E., "Microstrip antennas," *IEEE Trans. Antennas Propagat.*, Vol. ap-23, 90-93, January 1975.
12. Bahl, R. E. and P. Bhartia, "Microstrip antennas," Artech House, Dedham, MA, 1980.
13. From Pozar, D. M., "Microstrip antennas," *Proc. IEEE*, Vol. 80, 79–91, 1992.
14. Carver, K. R. and J. W. Mink, "Microstrip antennas technology," *IEEE Trans. Antennas Propagate.*, Vol. AP-29, 2–24, January 1981.
15. Mailloux, R. J., et al., "Microstrip array technology," *IEEE Trans. Antennas Propagate.*, Vol. AP-29, 25–37, January 1981.

16. Pozar, D. M. and D. H. Schaubert, *Microstrip Antennas: The Analysis and Design of Microstrip Antennas and Arrays*, IEEE Press, New York, 1995.
17. Luk, K. M. and K. F. Lee, "Circular U-slot patch with dielectric superstrate," *Electronics Letters*, Vol. 33, No. 12, 1001–1002, 1997.
18. Sabban, A., "A new broadband stacked two-layer microstrip antenna," *IEEE AP-S Int. Symp. Digest*, 63–66, June 1983.
19. Chen, A. H., A. Tulintseff, and R. M. Sorbello, "Broadband two layer microstrip antenna," *IEEE AP-S Int. Symp. Digest*, Vol. 2, 251–254, June 1984.

Dual-band CPW-fed G-shaped Monopole Antenna for 2.4/5 GHz WLAN Application

Wen-Chung Liu and Chao-Ming Wu

Formosa University, Taiwan, China

Abstract— A novel planar monopole antenna with dualband operation suitable for WLAN application is presented in this paper. The antenna resembles as a "G" shape from combining two folded strips and is fed by a coplanar waveguide (CPW) transmission line. By optimally selecting the antenna dimensions, dualband resonant modes with a much wider impedance matching at the higher band can be produced. Prototypes of the obtained optimized design have been constructed and tested. The measured results explore good dualband operation with −10 dB impedance bandwidths of 9.7% and 62.8% at bands of 2.43 and 4.3 GHz, respectively, which cover the 2.4/5.2/5.8 GHz WLAN operating bands, and show good agreement with the numerical prediction. Good antenna performances such as radiation patterns and antenna gains over the operating bands have also been observed.

1. INTRODUCTION

Recently, with the wireless communications, such as the wireless local area network (WLAN), having evolved at astonishing rate, it has been well known that the future communication technology pressingly demands integration of more than one communication system into a limited equipment space. This indicates the future communication terminal antennas will not only be desired to be low-profile lightweight, flush mounted, and single-feed but also need to meet the requirements of dual- or multiband operation for sufficiently covering the possible operating bands. To comply with this requirement, a compact planar high performance antenna with function of dual- or multiband operation, and good radiation characteristics is therefore needed. However, the difficulty of antenna design increases when the number of operating frequency bands increases. So far, many antenna designs with enhanced dual- or multiband operation capabilities to satisfy the IEEE 802.11 WLAN standards in the 2.4/5.2/5.8 GHz operating bands have been developed and presented [1–6]. Among these antennas, the CPW-fed monopole antennas have especially received much more interest than others owing to their potential in providing various required radiation features of dual- or multiband, broad bandwidth, simple structure of a single metallic layer, and easy integration with system circuit board. However, for the available CPW-fed antenna designs capable of dual- or multiband operation for use in a WLAN system, as references of [4–6], some of them are with a complex structure and some are large in antenna sizes.

In this study, we propose a CPW-fed G-shaped planar monopole antenna with optimal dualband operation simultaneously suitable for use in the 2.4/5.2 GHz wireless local area network (WLAN) applications. The geometry parameters of the proposed antenna, including the dimensions of the G-shaped strip structure, the sizes of the coplanar ground planes, and the spaces between the ground plane to either the G-shaped strips or the CPW feeding line, were all carefully selected by using moment method code IE3D to achieve good dualband operation. Prototypes of the obtained antenna for dualband operation were constructed and the predicted and measured antenna performance such as input return loss, impedance bandwidths, radiation patterns, and gains are presented and discussed.

2. DESIGN OF ANTENNA

Figure 1 shows the geometrical configuration of the proposed CPW-fed planar monopole antenna for achieving dualband operation. The antenna is printed on only one side of an FR4 microwave substrate with the substrate thickness of 1.6 mm and the dielectric constant of 4.4. The main structure of the proposed antenna comprises two folded strips, denoted as L_1 and L_2, respectively, and a CPW feeding line. The strips L_1 and L_2 are both have a fixed strip width of s and are folded to resemble the antenna in a "G" shape. For the smaller folded strip L_1, it includes two horizontal sections with lengths of $w - w_f/2$ and ℓ_1 for the lower and upper sections, respectively, and one vertical section of length of $d_1 + s + s$. As for the larger folded strip, L_2, it includes two vertical and two horizontal sections, and can be determined by using only three parameters, which are the distances d_2 and d_3 from the upper horizontal section of the smaller folded strip to the left and right vertical section, respectively, of the larger folded strip, and the length ℓ_2 of the right vertical section of the larger folded strip. The major function of the two folded strips of unequal lengths is to produce two different current paths and thus expected to effectively excite dual resonant modes. A $50\,\Omega$ CPW feeding line with a fixed signal strip thickness of w_f and a gap distance of g between the signal strip and ground is used for centrally feeding the G-shaped antenna from its bottom edge. Two equal finite ground planes, each with dimensions of width W_g and length L_g, are situated symmetrically on each side of the CPW feeding line. The G-shaped radiating structure has a vertical spacing of h away from the ground plane. In this investigation, for trying to obtain good dualband impedance matching by controlling the current distribution on the G-shaped stripline and compensation between the capacitive and inductive effects caused from the electromagnetic coupling effects of the finite ground planes and both the feeding line and the G-shaped stripline at the desired various operating bands, the geometry parameters were all carefully examined and finally the proper values for these parameters were obtained to be as those listed in Table 1.

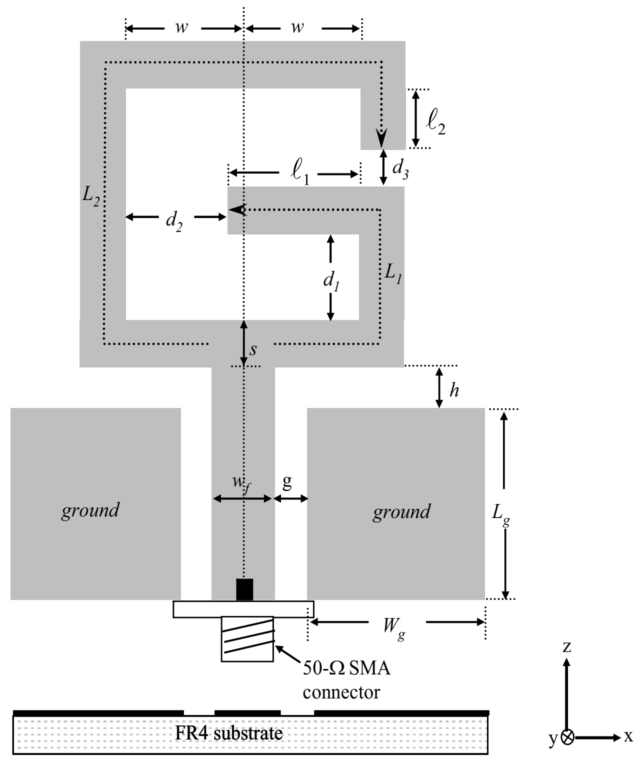


Figure 1: Geometrical configuration of proposed CPW-fed G-shaped monopole antenna for dualband operation.

Table 1: Antenna parameters (in mm).

ℓ_1	ℓ_2	d_1	d_2	d_3	s	w_f	W_f	L_g	g	h
10	1.2	10	3.3	10	3	4.75	5	10.75	1.35	4.69

3. RESULTS AND DISCUSSION

The obtained dual-band CPW-fed G-shaped monopole antenna was constructed and experimentally studied. Figure 2 shows the measurement and simulation frequency response of the return loss for the proposed antenna. Obviously, the simulation results show that except the two resonant modes at frequencies of 2.45 and 5.2 GHz two additional resonant modes are also excited at frequencies

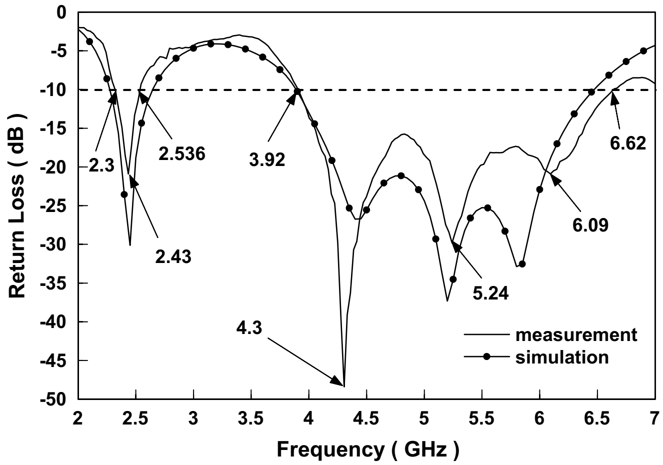


Figure 2: Return loss against frequency of proposed antenna from simulation and measurement: $\ell_1 = 10\,\mathrm{mm}$, $\ell_2 = 1.2\,\mathrm{mm}$, $d_1 = 10\,\mathrm{mm}$, $d_2 = 3.3\,\mathrm{mm}$, $d_3 = 10\,\mathrm{mm}$, $s = 3\,\mathrm{mm}$, $w_f = 4.75\,\mathrm{mm}$, $W_g = 5\,\mathrm{mm}$, $L_g = 10.75\,\mathrm{mm}$, $g = 1.35\,\mathrm{mm}$, and $h = 4.69\,\mathrm{mm}$.

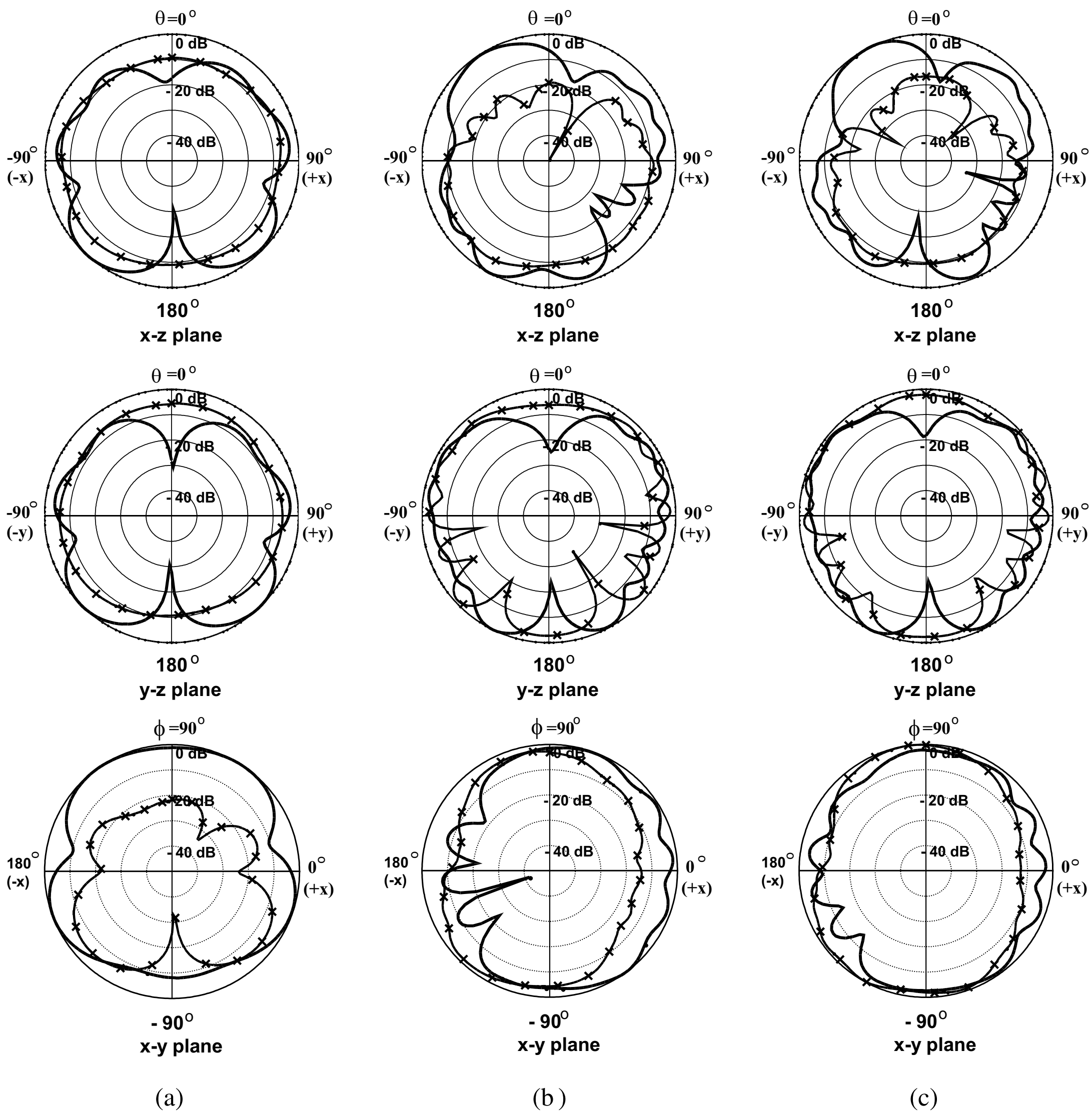
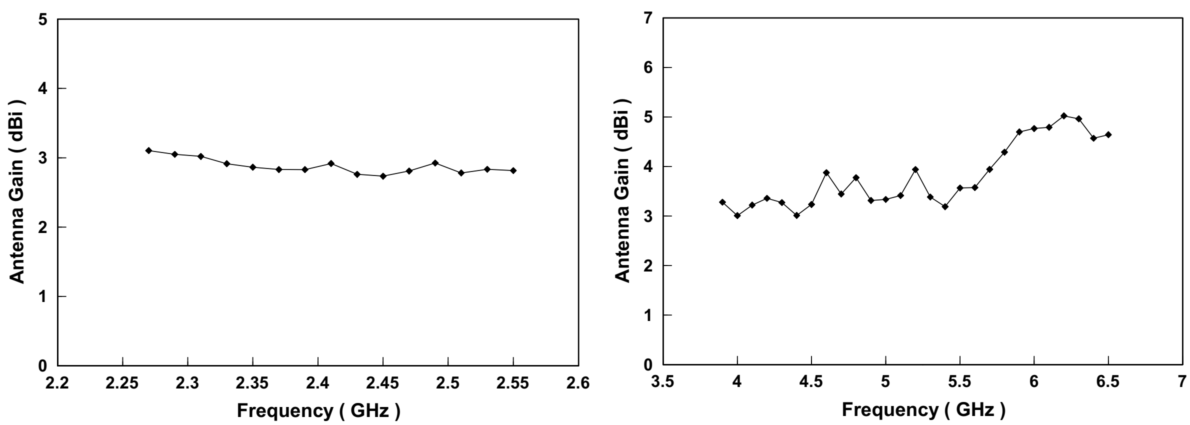


Figure 3: Measured radiation patterns for the proposed antenna at (a) $f = 2.45\,\mathrm{GHz}$, (b) $f = 5.25\,\mathrm{GHz}$, and (c) $f = 5.75\,\mathrm{GHz}$. (—E_θ; -x-x-xE_ϕ)

about 4.4 and 5.8 GHz. Especially, the simulated 10 dB impedance bandwidths across the three excited resonant bands at 4.4, 5.2, and 5.8 GHz are sufficient to produce a much broader continuous bandwidth from 3.91 to 6.46 GHz. As for the measured results, also, four resonant modes at frequency of 2.43, 4.3, 5.24, and 6.09 GHz were obtained. The lower mode has an impedance bandwidth (10 dB return loss) of 236 MHz (2.3–2.536 GHz), or about 9.7% with respect to the center frequency at 2.43 GHz, while for the higher mode, a broader continuous bandwidth produced from the three close resonant modes at 4.3, 5.24, and 6.09 GHz, has been reached to be 2.7 GHz (3.92–6.62 GHz), or about 62.8% referred to the best resonance frequency at 4.3 GHz. Obviously, the agreement between simulation and measurement seems very good. In addition, the obtained bandwidths can sufficiently cover the WLAN standards in the 2.4 GHz (2.4–2.484 GHz), 5.2 GHz (5.15–5.35 GHz), and 5.8 GHz (5.725–5.825 GHz) bands.

Radiation characteristics of the constructed antenna are also studied. Figure 3 plots the measured radiation patterns including the vertical (E_θ) and the horizontal (E_ϕ) polarization patterns in the elevation direction (x-z and y-z planes) and azimuthal direction (x-y plane) at 2.45, 5.25, and 5.75 GHz for the proposed antenna. Due to the asymmetry in the G-shaped structure, unsymmetrical radiation patterns are seen in the three cuts as depicted in the plots. In addition, general monopole-like radiation patterns in the x-z and y-z planes and nearly omnidirectional radiation in the azimuthal plane are observed. However, it is also found that the E_θ and E_ϕ components of the patterns in both the x-z and y-z planes are seemed to be much comparable. This electromagnetic phenomenon is probably a result of the strong horizontal components of the surface current at the two folded-strip of the G-shaped structure. Also note that very stable radiation patterns have been obtained for the proposed antenna from measurements at other operating frequencies across the bandwidth of each band.



(a) (b)

Figure 4: Measured peak antenna gain for frequencies across (a) the lower band (2.3-2.536 GHz) and (b) the higher band (3.92–6.62 GHz) for the proposed antenna studied in Figure 2.

The peak antenna gain of the proposed antenna for frequencies across the dual bands was measured and shown in Figure 4. The ranges of antenna gain at the lower band of around 2.45 GHz is about 2.7–3.1 dBi with a very flat gain curve, while that at the higher operating band ranged from 3.9 to 6.6 GHz is about 3–5 dBi also with a small gain variation around 5 GHz band.

4. CONCLUSION

In this study, a novel CPW-fed G-shaped planar monopole antenna with dualband operation is presented. The constructed prototype of the proposed antenna with dual impedance bandwidths of 9.7% and 62.8% at bands of 2.43 and 4.3 GHz, respectively, sufficiently covering the bandwidth requirements of the WLAN system in the 2.4/5.2/5.8 GHz standards, have been studied. Also, good antenna performances have been obtained and shown to match well with the numerical prediction.

ACKNOWLEDGMENT

The author wishes to thank Mr. C. F. Hsu for his contribution in the elaboration of the experimental data. This work was supported by the National Science Council of Republic of China under grant

NSC 95-2221-E150-019-

REFERENCES

1. Song, C. T. P., P. S. Hall, H. Ghafouri-Shiraz, and D. Wake, "Triple band planar inverted F antennas for handheld devices," *Electron. Lett.*, Vol. 36, No. 2, 112–114, 2002.
2. Choi, W., S. Kwon, and B. Lee, "Ceramic chip antenna using meander conductor lines," *Electron. Lett.*, Vol. 37, No. 15, 933–934, 2001.
3. Kuo, Y. L. and K. L. Wong, "Printed double-T monopole antenna for 2.4/5.2 GHz dual-band WLAN operations," *IEEE Trans. Antennas Propagat.*, Vol. 51, No. 9, 2187–2192, 2003.
4. Raj, R. K., M. Joseph, B. Paul, and P. Mohanan, "Compact planar multiband antenna for GPS, DCS, 2.5/5.8 GHz WLAN applications," *Electron. Lett.*, Vol. 41, No. 6, 290–291, 2005.
5. Liu, W. C., "Broadband dual-frequency cross-shaped slot cpw-fed monopole antenna for WLAN operation," *Microwave Opt. Technol. Lett.*, Vol. 46, No. 4, 353–355, 2005.
6. Liu, W. C., "Broadband dual-frequency meandered cpw-fed monopole antenna," *Electron. Lett.*, Vol. 40, No. 21, 1319–1320, 2004.

Design CPW Fed Slot Antenna for Wideband Applications

K. Nithisopa[1,2], **J. Nakasuwan**[1], **N. Songthanapitak**[1]
N. Anantrasirichai[3], **and T. Wakabayashi**[4]
[1]Department of Electronic and Telecommunication Engineering, Faculty of Engineering
Rajamangala University of Technology Thanyaburi (RMUTT)
Klong 6, Thanyaburi, Pathumthanee, Thailand
[2]Department of Electronic and Telecommunication Engineering, Faculty of Engineering
Rajamangala University of Technology Phra Nakhon (RMUTP)
North Bangkok, Bangkok, Thailand
[3]ReCCIT, Faculty of Engineering, King Mongkut's Institute of Technology Ladkrabang
Ladkrabang, Bangkok, 10520, Thailand
[4]School of Information Science and Engineering, Tokai University
Hiratsuka, Kanagawa, 256-1292, Japan

Abstract— In this paper, a CPW-fed slot antenna for wideband application was designed and simulated. In order to examine the performances of this antenna, a prototype was designed at frequency 2.4 GHz and simulated with various width of slot antenna in both sides for input impedances matching and simulated by IE3D software package of Zeland. The simulation result of bandwidth is 1.65 GHz (2.1–3.75 GHz) which covers the standard frequency of IEEE 802.11 b/g (2.4–2.4835 GHz) and Wimax (2.3–3.6 GHz). With these performances, the proposed antenna can be used in wideband applications.

1. INTRODUCTION

Microstrip antenna is one type of antennas which can be used for transmitting and receiving signals. Microstrip or printed antennas are low profile, small size, light weight and widely used in wireless and mobile communications, as well as radar applications. Microstrip antennas can be divided into two basic types by structure, namely microstrip patch antenna and microstrip slot antenna [1, 2]. The slot antennas can be fed by microstrip line, slot line and CPW [3, 4]. The CPW is the feeding which side-plane conductor is ground and center strip carries the signal. The advantage of CPW fed slot antenna is wideband antenna which many research introduce the several shape of slot antenna for use in WLAN applications. In this paper, we proposed the slot antenna fed by CPW at a designed frequency of 2.4 GHz and coverage frequency range from 2.1–3.75 GHz.

2. ANTENNA STRUCTURE

The CPW fed slot antenna is designed at 2.4 GHz with the symmetrical structure, as shown in Figure 1(a). This antenna is designed on RT/Duroid 5880, the substrate with thickness (h) of 1.575 mm and dielectric constant (ε_r) of 2.2. The coplanar waveguide (CPW) is designed to be 50 ohms in order to match the characteristic impedance of transmission line. The dimension of the slot antenna is referred to the guide wavelength (λ_g) which given by

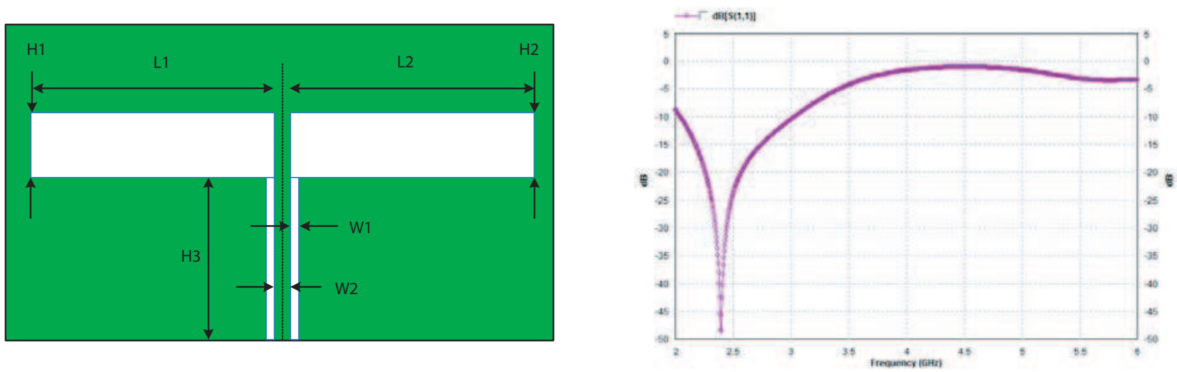
$$\lambda_g = \frac{c/f}{\sqrt{\varepsilon_{eff}}} \tag{1}$$

where ε_{eff} is an effective dielectric constant.

$$\varepsilon_{eff} \approx \frac{\varepsilon_r + 1}{2} \tag{2}$$

In this case, λ_g at frequency 2.4 GHz is 98.81 mm.

The total length of slot antenna ($L_1+L_2+W_2$) is $0.81\lambda_g$ (80.0 mm) and width of slot (H_1, H_2) is $0.1\lambda_g$ (10.5 mm). For match impedance with characteristic impedance of transmission line 50 ohms, the gap (W_1), width of the center strip (W_2) and length of CPW line (H_3) are 0.5 mm, 2.4 mm and 23 mm, respectively.



(a) (b)

Figure 1: CPW-fed slot antenna and S_{11} in case of same slot. (a) Configuration of CPW fed slot antenna. (b) Characteristic of return loss (S_{11}).

3. DESIGN PROCEDURE

In this paper, we proposed the concept of the designing CPW fed slot antenna for wideband which has three design procedures. This slot antenna composed of two small slots on the ground plane that are left and right slots. For every case of designing, we will fix the gap (W_1), width of the center strip (W_2) and length of CPW line (H_3) to 0.5 mm, 2.4 mm and 23 mm, respectively.

Design 1: Same length and same width of two slot (left slot and right slot).

The parameters of this structure are as following:

$$L_1 = L_2 = 38.8\,\text{mm} \qquad \text{(total length of slot antenna} = 80.0\,\text{mm)}$$
$$H_1 = H_2 = 10.5\,\text{mm}$$

Figure 1(b) show the characteristic of return loss S_{11} of Design 1. The simulation results of bandwidth and its return loss are shown in Table 1.

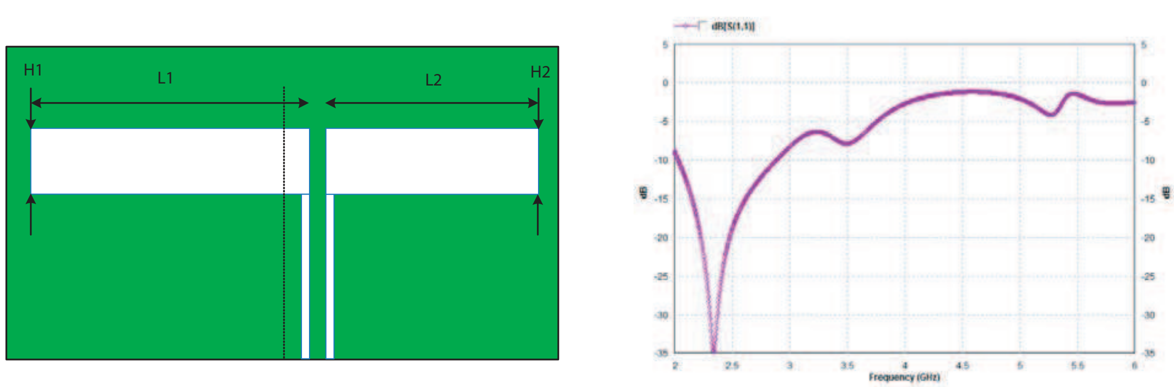
Design 2: Different length and same width of two slots.

The parameters of this structure are as following:

$$L_1 = 43.8\,\text{mm},\ L_2 = 33.8\,\text{mm} \qquad \text{(total length of slot antenna} = 80.0\,\text{mm)}$$
$$H_1 = H_2 = 9.5\,\text{mm}$$

The simulation results are shown in Table 1, and the return loss S_{11} is shown in Figure 2(b).

Design 3 : Different length and different width of two slots.

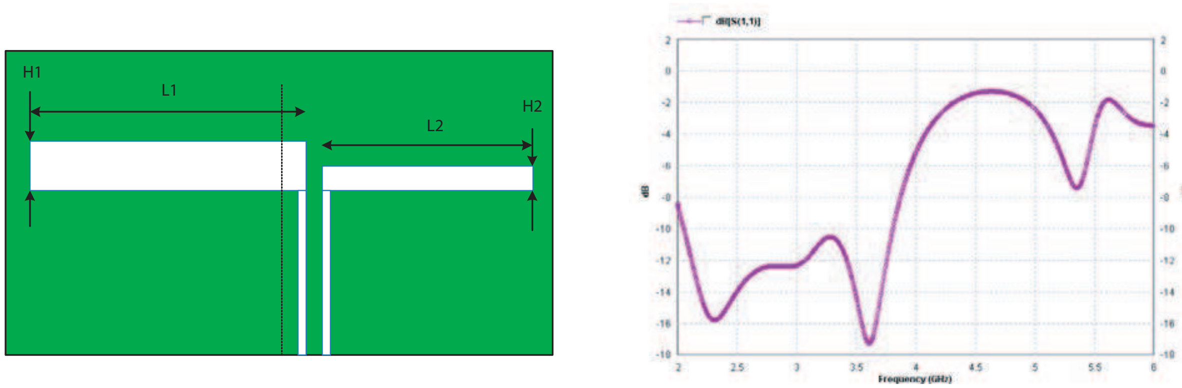


(a) (b)

Figure 2: CPW-fed slot antenna in case of difference length and same width of slot. (a) Configuration of CPW fed slot antenna. (b) Characteristic of return loss (S_{11}).

The parameters of this structure are as following:

$$L_1 = 43.8\,\text{mm},\ L_2 = 33.8\,\text{mm} \qquad \text{(total length of slot antenna} = 80.0\,\text{mm)}$$
$$H_1 = 7.8, \qquad H_2 = 4.1\,\text{mm}$$



(a) (b)

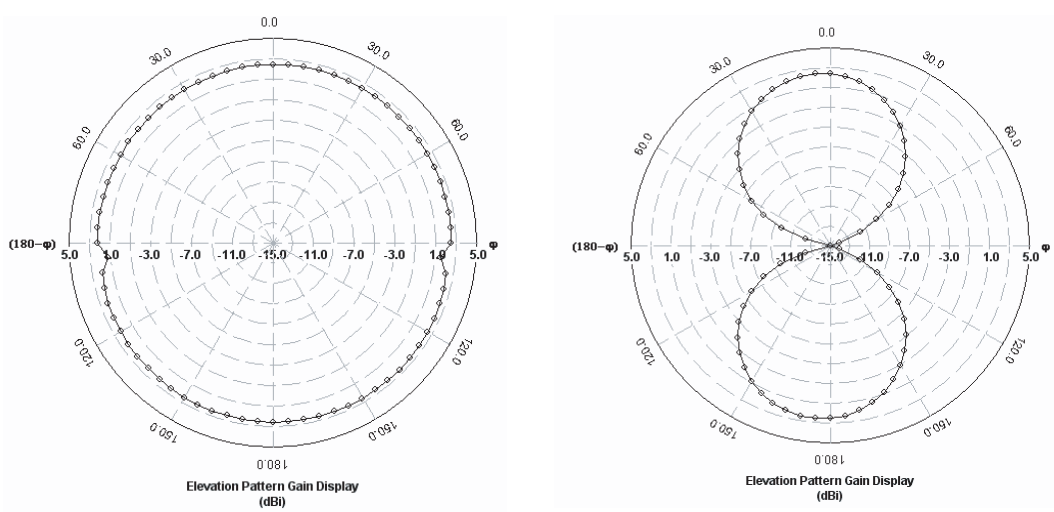
Figure 3: CPW-fed slot antenna in case of difference length and difference width of slot. (a) Configuration of CPW fed slot antenna. (b) Characteristic of return loss (S_{11}).

4. SIMULATE RESULTS

The basic slot antenna fed by CPW is shown in Figure 1(a). When varying the length of slot, it will affect on bandwidth and return loss as shown in Figure 2. When we increase width of slot, the bandwidth is increasing, as shown in Figure 3. The dimension and some parameters are listed in Table 1.

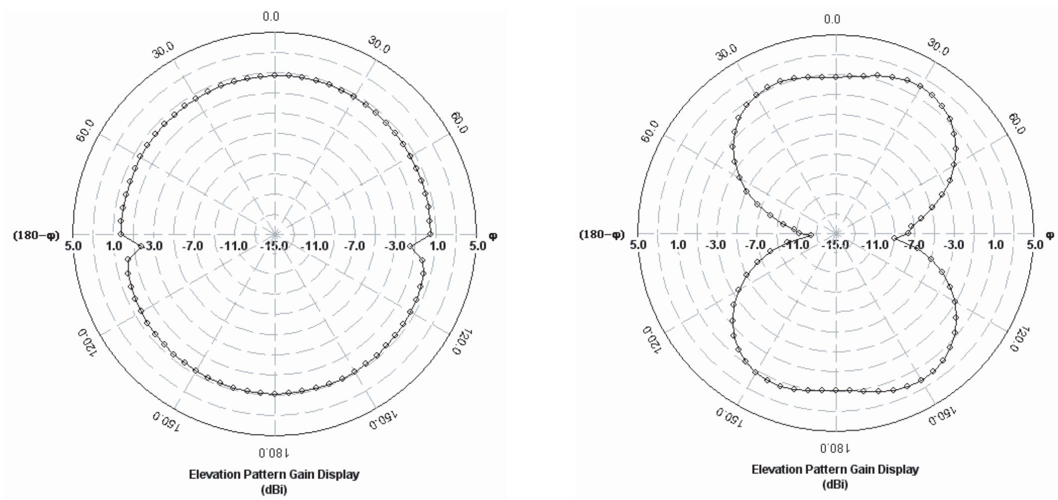
Table 1: The value of parameters and the characteristics of CPW-fed slot antenna.

L1 (mm)	L2 (mm)	H1 (mm)	H2 (mm)	Bandwidth (−10 dB) (GHz)	Return Loss (dB) ≅ 2.4 GHz
38.8	38.8	10.5	10.5	1.0	−48
43.8	33.8	9.5	9.5	0.8	−35
43.8	33.8	7.8	4.1	1.65	−15.5



(a) (b)

Figure 4: Radiation pattern at frequency 2.45 GHz. (a) yz-plane. (b) xz-plane.



(a) (b)

Figure 5: Radiation pattern at frequency 3.65 GHz. (a) yz-plane. (b) xz-plane.

The radiation pattern of y-z plane and x-z plane at frequency 2.45 GHz and 3.65 GHz are shown in Figure 4 and Figure 5.

5. CONCLUSION

The design of slot antenna fed by CPW is considered on the basic structure. It is proved by varying the length and the width of the slot for achieving the wideband for use in WLAN applications. This paper shows the maximum bandwidth of 1.65 GHz at design frequency of 2.4 GHz. The wideband is created with the different length and the different width of the slot antenna.

REFERENCES

1. Benson, F. A. and T. M. Benson, *Fields Waves and Transmission Lines*, Chaman & Hall, 1991.
2. Balanis, C. A., *Antenna Theory Analysis and Design*, John Wiley & Sons, Inc., 1997.
3. Giauffret, L., J.-M. Laheurte, and A. Papiernik, "Study of various shapes of the coupling slot in CPW-fed microstrip antenna," *IEEE Trans. Antennas Propagat*, Vol. 45, No. 4, 642–647, 1997.
4. Bhobe, A. U. and C. L. Holloway, "Wide-band slot antennas with CPW-feed line: hybride and log-periodic design," *IEEE Trans. Antennas Propagat*, Vol. 52, No. 10, 2545–2554, 2004.
5. Wang, C.-J., Member, IEEE, J.-J. Lee, and R.-B. Huang, Member, IEEE, "Experimental studies of a miniaturized CPW-fed slot antenna with the dual-frequency operation," *IEEE Antennas and Wireless Propagation Letters*, Vol. 2, 2003.

e-Shaped Slot Antenna for WLAN Applications

T. Archevapanich[1,2], J. Nakasuwan[1], N. Songthanapitak[1]
N. Anantrasirichai[3], and T. Wakabayashi[4]

[1]Department of Electronic and Telecommunication Engineering, Faculty of Engineering
Rajamangala University of Technology Thunyabury (RMUTT)
Pathumthanee, Thailand
[2]Department of Electronic and Telecommunication Engineering, Faculty of Engineering and Architecture
Rajamangala University of Technology Suvannabhumi, Thailand
[3]ReCCIT, Faculty of Engineering, King Mongkut's Institute of Technology Ladkrabang
Ladkrabang, Bangkok, Thailand
[4]School of Information Technology and Electronics, Tokai University, Hirasuka, Kanagawa 259-1292, Japan

Abstract— This paper present an e-shaped slot antenna for wireless communications. The antenna is designed for dual frequency band 2.4–2.52 GHz and 4.82–6.32 GHz, which support WLAN communications coverage IEEE 802.11b/g (2.4–2.4835 GHz), IEEE 802.11j (4.90–5.091), IEEE 802.11a (5.15–5.35 GHz), and IEEE 802.16d (5.7–5.9 GHz). The bandwidth at low resonant frequency and high resonant frequency are about 0.12 GHz and 1.5 GHz, respectively. The simulation results of e-shaped slot antenna are analyzed by using Method of Moment (MoM) from IE3D Software.

1. INTRODUCTION

The e-shaped slot antenna fed by microstrip line is one type of microstrip antenna which has advantages such as: low profile, lightweight and easy to fabrication [1]. The antenna was designed for two frequency bands and referred to the guided wavelength. The IE3D software as referred in [2] was used to analyze the proposed antenna.

In this paper, a microstrip fed e-shaped slot antenna is presented. The design objective is to satisfy Wireless Local Area Network (WLAN) of IEEE 802.11b/g/j/a and IEEE 802.16d. Method of Moment was applied to evaluate the characteristics of the proposed antenna. Although many researchers have studied the other shape of antenna, but this e-shaped slot antenna is the new shaped which we will purpose and controlled for dual frequency with matching resonant frequency was rarely investigated. Therefore, the effect of varying width of slot antenna was investigated in this paper by using simulation software. The simulation results show that this antenna can be applied to serve WLAN applications.

2. ANTENNA STRUCTURE

This antenna was designed on RT/Duroid 5880 with 1.575 mm of thickness, h, and 2.2 of dielectric constant, ε_r. The width of microstrip feed line (w) is designed to match impedance of characteristic impedance of transmission line 50 ohms which can be calculated by following:

$$\frac{w}{h} = \frac{2}{\pi}\left\{B - 1 - \ln(2B-1) + \frac{\varepsilon_r - 1}{2\varepsilon_r}[\ln(B-1)] + 0.39 - \frac{0.61}{\varepsilon_r}\right\} \tag{1}$$

where $B = \dfrac{60\pi^2}{Z_0\sqrt{\varepsilon_r}}$.

In this case, $w = 4.7$ mm.

The wave length (λ_g) in the substrate of this antenna can be calculated from following equations

$$\lambda_g = \frac{c/f}{\sqrt{\varepsilon_{eff}}} \tag{2}$$

where ε_{eff} is the effective dielectric constant:

$$\varepsilon_{eff} = \frac{\varepsilon_r + 1}{2} + \frac{\varepsilon_r - 1}{2}\left[1 + 12\frac{h}{W}\right]^{-1/2} \tag{3}$$

In this case, λ_g at frequency 2.4 GHz is 91.66 mm and ε_{eff} is 1.86.

The configuration of this antenna is shown in Figure 1 which has dimension of 27 mm × 17.3 mm. (A × B)

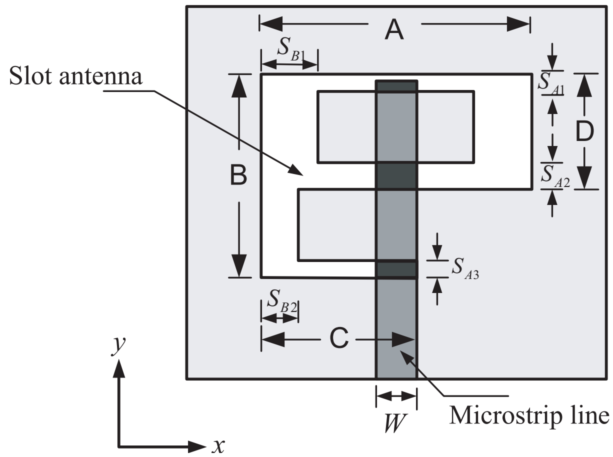


Figure 1: The configuration of e-Shaped Slot Antenna.

3. FINDING THE DUAL FREQUENCY FOR WLAN

The length/dimension of e-shaped slot antenna in each side are A $= 0.3\lambda_g$(27 mm), B $= 0.2\lambda_g$(17.3 mm), C $= 17\lambda_g$(15.85 mm), and D $= 0.1\lambda_g$(9.2 mm). The parameters in width of slot antenna are:

S_{B1} = width of upper slot in y-axis
S_{B2} = width of lower slot in y-axis
S_{A1} = width of upper slot in x-axis
S_{A2} = width of middle slot in x-axis
S_{A3} = width of lower slot in x-axis

In this case, we fixed the widths of three slots (S_{A1}, S_{A2}, S_{A3}) in x-axis to 1 mm. To find the dual frequency which match to the 50 ohms transmission line, we propose 2 steps for achieving the WLAN covering IEEE 802.11 a/b/g/j and IEEE 802.16d are as follows:

Step 1: $S_{B1} = S_{B2}$

Varying S_{B1} and S_{B2} to 1 mm, 2 mm, 3 mm, 4 mm, 5 mm. (The length of microstrip line is adjusted for match impedance of 50 ohms.)

The simulation result of return loss (S_{11}) is shown in Figure 2. Table 1 displays the results of return loss and bandwidth. The results show that when increasing the width of slot, the bandwidth and resonant frequency are increased, while low resonant frequency is slightly increased as shown in Figure 2. The simulation of low resonant frequency and high resonant frequency are shown in Figure 3, when varying S_{B1} and S_{B2}.

Step 2: $S_{B1} \neq S_{B2}$

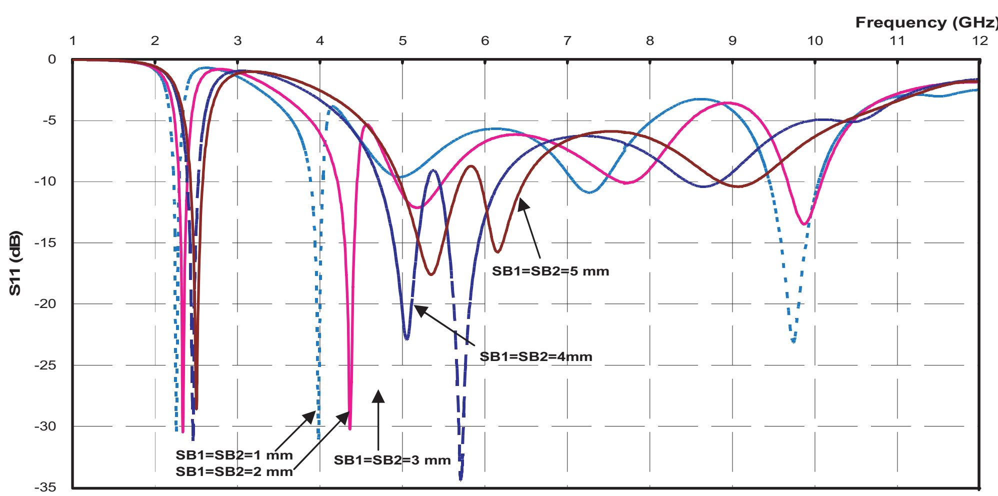


Figure 2: Characteristics of return loss when $S_{B1} = S_{B2}$ in the case of Step 1.

Table 1: The simulation results of S_{11} and bandwidth in the case of Step 1.

S_{B1}, S_B (mm)	$S_{11(L)}$ (dB)	Low Resonant Freq. (GHz)	Bandwidth (GHz)	$S_{11(H)}$ (dB)	High Resonant Freq. (GHz)	Bandwidth (GHz)
1	-30.41	2.26	0.04 (2.24-2.28)	-31.06	3.98	0.12 (3.90-4.02)
2	-30.45	2.34	0.08 (2.30-2.38)	-30.12	4.36	0.22 (4.22-4.44)
3	-28.67	2.40	0.10 (2.36-2.46)	-26.62	4.72	0.38 (4.48-4.86)
4	-31.20	2.46	0.12 (2.40-2.52)	-22.92	5.06	0.58 (4.72-5.30)
5	-28.59	2.50	0.12 (2.44-2.56)	-17.59	5.34	0.70 (4.98-5.68)

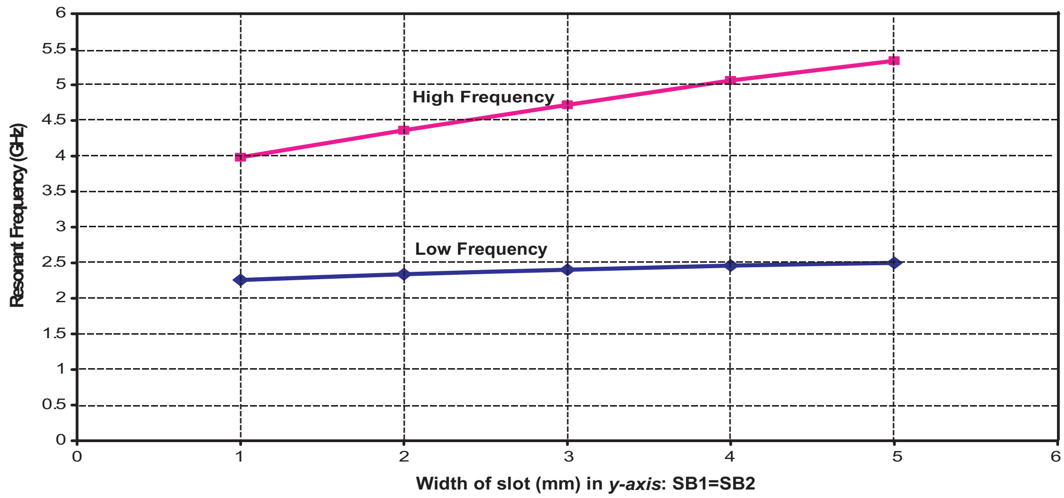


Figure 3: Effect of varying S_{B1}, S_{B2} in the case of Step 1.

Choosing the appropriate value of S_{B1} and S_{B2} from step 1 that are $S_{B1} = 4$ mm and $S_{B2} = 4$ mm. Adjusting S_{B1} to 5 mm for wideband at high frequency.

The simulation result of return loss (S_{11}) is shown in Figure 4 and Table 2. The bandwidth of low resonant frequency is same as step 1 that is 0.12 GHz (2.4 GHz–2.52 GHz), and bandwidth of high resonant frequency is 1.50 GHz (4.82 GHz–6.32 GHz), which is wideband frequency.

Finally, the bandwidth at $S_{11} = -10$ dB of lower resonant frequency is 4.88% and higher resonant frequency is 26.9%.

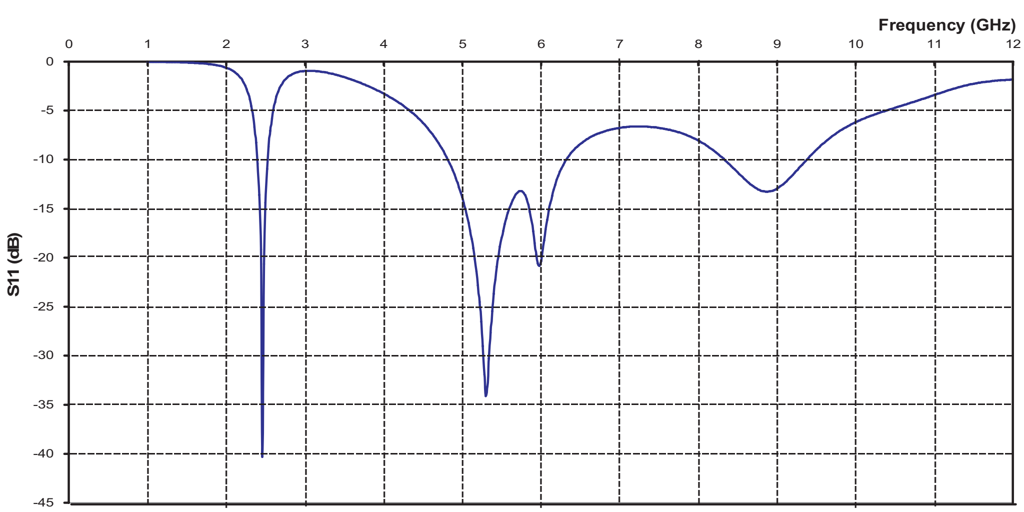


Figure 4: Characteristics of return loss in the case of Step 2.

Table 2: The simulation results of e-shaped slot antenna in the case of $S_{B1} = 5\,\text{mm}$ and $S_{B2} = 4\,\text{mm}$.

Low Resonant Freq. (GHz)	$S_{11(L)}$ (dB)	Bandwidth (GHz)	Gain (dBi)	High Resonant Freq. (GHz)	$S_{11(H)}$ (dB)	Bandwidth (GHz)	Gain (dBi)
2.46	-40.31	0.12 (2.4-2.52)	2.1	5.3	-34.10	1.5 (4.82-6.32)	4.7

4. RADIATION PATTERN

Figure 5(a) and 5(b) present the radiation pattern on y-z plane cut at phi = 90° at 2.46 GHz and 5.3 GHz, respectively. This antenna is linear polarization at low resonant frequency 2.46 GHz and is circular polarization at high resonant frequency around 5.8 GHz.

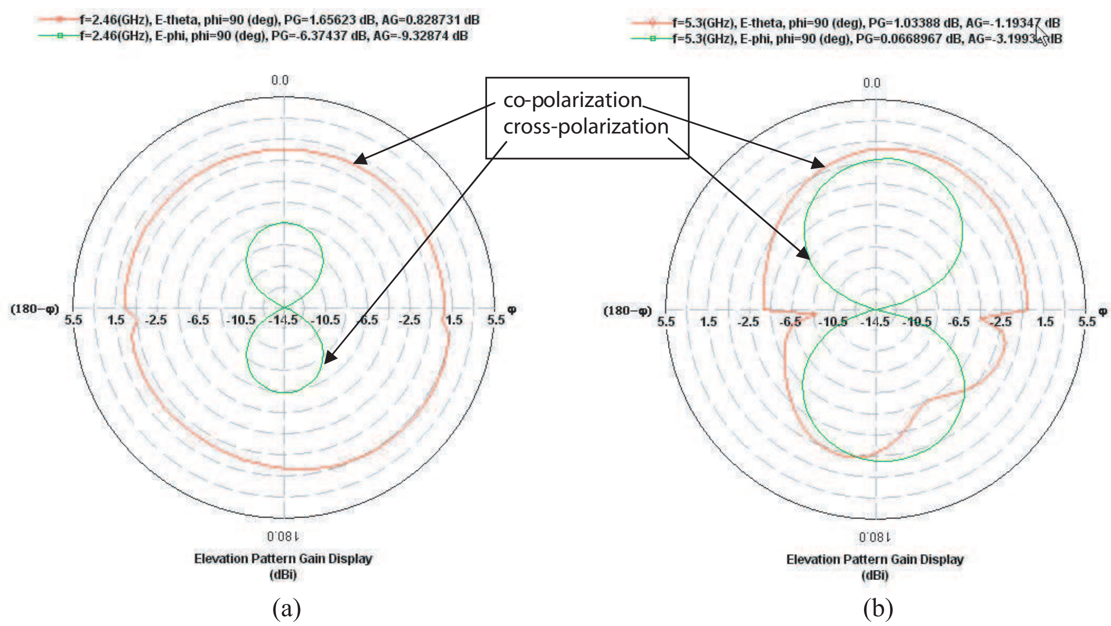


Figure 5: The simulation results of radiation pattern on y-z plane. (a) At resonant frequency 2.46 GHz. (b) At resonant frequency 5.3 GHz.

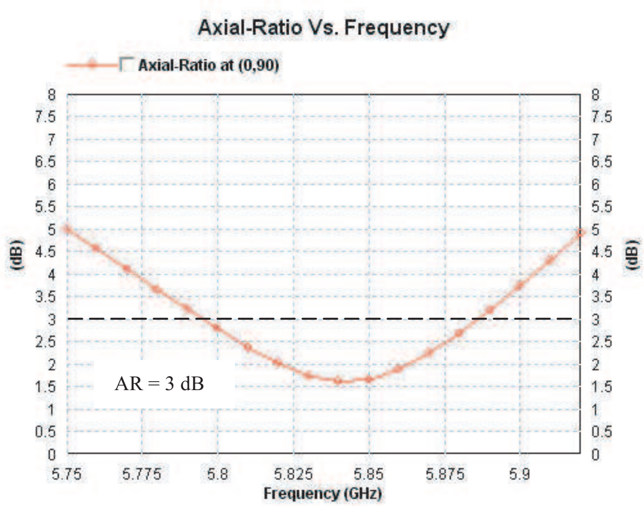


Figure 6: Characteristic of axial ratio represent the polarization of e-shaped slot antenna.

5. CONCLUSION

The e-Shaped slot antenna was designed to support WLAN communications at frequency band 2.4–2.52 GHz and 4.82–6.32 GHz for standards IEEE 802.11b/g/j/a and IEEE 802.16d. Varying the width of slot S_{B2} will affect on the match impedance at low frequency band and S_{B1} will affect on high frequency band.

REFERENCES

1. Balanis, C. A., *Antenna Theory Analysis and Design*, John Wiley & Sons, Inc., 1997.
2. *IE3D User's Manual Release 9*, Zeland software, Inc., U.S.A., 2002.
3. Benson, F. A. and T. M. Benson, *Fields Waves and Transmission Line*, Chapman & Hall, 1991.
4. Anantrasirichai, N., P. Rakluea, and T. Wakabayachi, "Slot antenna coupled by microstrip line for dual frequency," *ISITA/NOLTA 2002*, 635–638, October 7–11, 2002.