Analyzing of Coupling Region for CRLH/RH TL Coupler with Lumped-elements

Y. Wang[2], **Y. Zhang**[1, 2], **and F. Liu**[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 coupling region of the CRLH/RH TL coupler, which is composed of a composite right/left-handed transmission line (CRLH TL) with lumped-elements and a conventional right-handed transmission line (RH TL), is explicitly decided. Detailed formulas are given to define the frequency band edge of the coupling region. ADS S-parameter simulations are demonstrated to confirm our theoretical results.

1. INTRODUCTION

Over the last years, left-handed materials based on transmission line approach have become very popular due to their potential application in microwave community. This approach to realize metamaterials was first introduced at approximately the same time by three different groups [1–3]. Caloz and Itoh have referred to this type of left-handed materials as composite right-/left-handed transmission lines (CRLH TL). Recently, CRLH TL have led to the development of several new component and devices, such as a leaky-wave backward antennas [4], phase-shifting lines [5] and couplers [6–8].

The asymmetric CRLH/RH TL coupler was first introduced and studied in [7] and [8]. It was composed of a conventional right-handed transmission line (RH TL) and a composite right/left-handed transmission line (CRLH TL), which was realized by loading a conventional microstrip line with series capacitors and shunt inductors. The coupler showed superior performance such as broad bandwidth and tight coupling characteristics. The interesting features of such asymmetric CRLH/RH TL coupler were discussed using coupled-mode theory based on traveling waves. In Ref [9], it was shown that the formation of a stop-band and the excitation of complex modes occurred in the case of coupling between a forward wave and a backward-wave mode for a range of frequencies around the tuning frequency. Moreover, Ref [10] showed that by adjusting the spacing between the RH TL and CRLH TL, backward coupling operates in the left-handed region. In this paper the coupling region of the asymmetric CRLH/RH TL coupler is intensively studied. The conditions for tight coupling are presented and detailed formulas are given to define the edges of the coupling region.

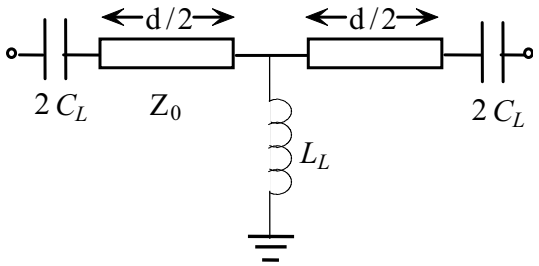


Figure 1: Circuit model for a unit cell of the CRLH TL, L_L and C_L are loaded lumped-elements.

2. CHARACTERISTICS OF A CRLH TL

A CRLH TL has been theoretically investigated in [4] and [6]. It can be implemented by loading a conventional transmission line with lumped element series capacitors (C_L) and shunt inductors (L_L) as depicted in Fig. 1. In this structure, the loading elements represent the left-handed attributes and the interconnecting transmission provides the right-handed contribution. Fig. 2 shows the dispersion diagram of the CRLH TL, which illustrates that it has a left-handed region and a right-handed region. The left-handed region has double-negative parameters, where both the permittivity and permeability are negative. Meanwhile, the right-handed region behaves as a conventional

transmission line with both positive permittivity and permeability. It is also shown that a CRLH TL whose series and shunt resonant frequencies are not equal ($f_{se} \neq f_{sh}$) displays a stopband.

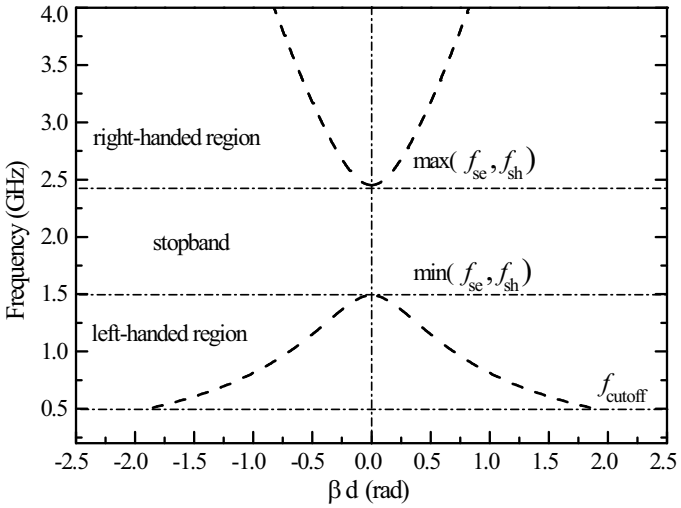


Figure 2: Dispersion diagram of the CRLH TL.

For a CRLH TL, the cutoff frequency of the left-handed region is given by

$$f_c = \frac{1}{4\pi\sqrt{L_L C_L}} \tag{1}$$

and the effective permittivity and permeability are given by the following approximate expressions

$$\varepsilon = \varepsilon(\omega) = \frac{1}{p}\left(C_R - \frac{1}{\omega^2 L_L d}\right) \tag{2}$$

$$\mu = \mu(\omega) = p\left(L_R - \frac{1}{\omega^2 C_L d}\right) \tag{3}$$

where, p is the structural constant, L_R (H/m) and C_R (F/m) are the distributed series inductance and shunt capacitance of the interconnecting transmission line. C_L (F) and L_L (H) are the parameter of lumped elements. The parameter d is the length of unit-cell.

From Eq. (2), it can be shown that for $\mu < 0$, the frequency satisfies

$$f < \frac{1}{2\pi\sqrt{L_R C_L} d} = f_{se} \tag{4}$$

Similarly, from Eq. (3), it can be shown that for $\varepsilon < 0$, the frequency satisfies

$$f < \frac{1}{2\pi\sqrt{L_L C_R} d} = f_{sh} \tag{5}$$

Therefore, the left-handed passband where $\mu < 0$ and $\varepsilon < 0$, is given by $f_c < f < \min(f_{se}, f_{sh})$.

3. DISCUSSION ABOUT THE COUPLING REGION

3.1. ENG CRLH TL/RH Coupler

A twenty-unit CRLH TL with circuit model in Fig. 1 is designed. The width of the transmission line is 2.945 mm with FR-4 substrate of $\varepsilon_r = 4.75$, of which the distributed series inductance L_R and shunt capacitance C_R equal to 319.72 nH/m and 127.89 pF/m respectively. The loading series capacitors are $C_L = 5.1$ pF and shunt inductors $L_L = 4.7$ nH. The unit cell dimension is $d = 7$ mm. The cutoff frequency defined by Eq. (1) is $f_c = 0.51$ GHz. The series and shunt resonant frequencies defined by Eqs. (4) and (5) can be obtained as $f_{se} = 1.49$ GHz and $f_{sh} = 2.45$GHz. Therefore, the left-handed passband of this CRLH TL is $f_c < f < f_{se}$. We refer this CRLH TL as ENG CRLH TL for the stopband of this CRLH TL is epsilon-negative.

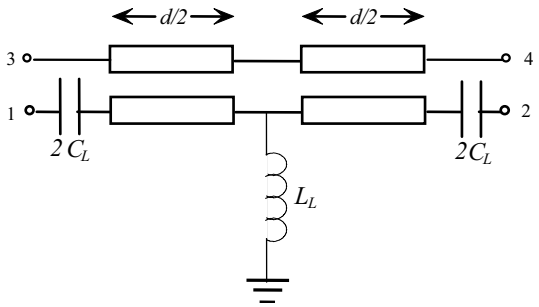


Figure 3: Circuit model for a CRLH/RH TL coupler, L_L and C_L are loaded lumped-elements.

When a conventional right-handed transmission line of the same length and width is near this CRLH TL, coupling will occur between the two lines. Then, the two lines constitute a CRLH TL/RH TL coupler with circuit model in Fig. 3. In fact, the coupling occurs only between the conventional transmission line and the part of interconnecting transmission line in the CRLH TL, but not between the transmission line and lumped elements in CRLH TL. As a result, the distributed series inductance L_R and shunt capacitance C_R of the interconnecting transmission line will be modified correspondingly, which will then lead to the changes of f_{se} and f_{sh}. However, the cutoff frequency f_c keeps constant because coupling won't change the values of C_L and L_L. When the space between the two lines is 0.2 mm, the distributed series inductance and shunt capacitance will change to $C'_R = 141.8\,\mathrm{pF/m}$ and $L'_R = 260.0\,\mathrm{nH/m}$ [11]. Then, the edges of the stopband become $f'_{se} = 1.62\,\mathrm{GHz}$ and $f'_{sh} = 2.33\,\mathrm{GHz}$. Therefore, the left-handed region becomes $f_c < f < f'_{se}$.

Figure 4 shows the change of the through coupling parameters S_{21} for the CRLH TL when the RH TL is put nearer and nearer. No coupling occurs if the original spacing between the two lines is infinity. The edges of the stopband are shown to be 1.49 GHz and 2.45 GHz. During the decreasing of the spacing, the bandgap edges move to 1.65 GHz and 2.4 GHz which confirm our theoretical prediction. What's more, a new gap appears on S_{21} during decreasing of the spacing between the two lines. This gap appears because of the coupling, so this frequency band corresponds to the coupling region. It demonstrates coupling region is just in the left-handed region of the CRLH TL. Therefore, the coupling region must satisfy $f_c < f < f'_{se}$.

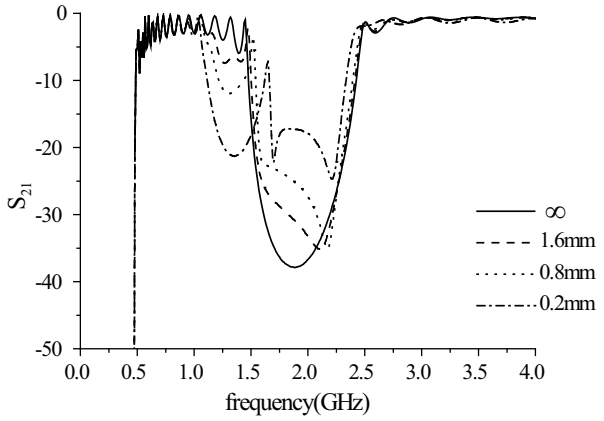


Figure 4: S_{21} of the ENG CRLH TL when the spacing between the lines is changing from 0.2 mm to infinity large.

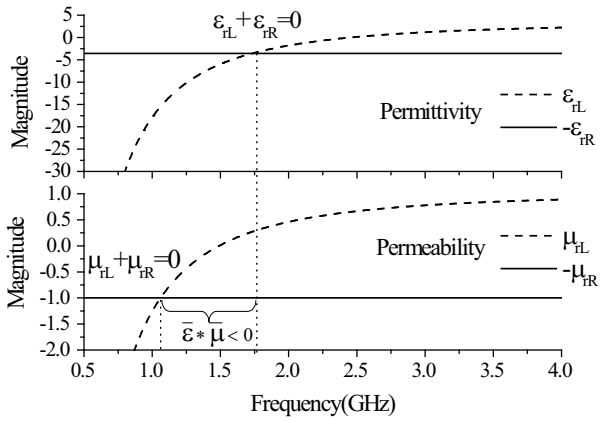


Figure 5: Relative premittivity ε_{rL} and permeability μ_{rL} of the ENG CRLH TL and relative premittivity ε_{rR} and permeability μ_{rR} of the RH TL.

Figure 5 shows the μ_{rL} and ε_{rL} of the CRLH TL and the negative μ_{rR} and ε_{rR} of the RH TL. The effective permeability μ_{rL} and permittivity ε_{rL} of the CRLH TL are defined by Eqs. (2) and (3). However, μ_{rR} and ε_{rL} of the RH TL are constant and equals to 1.00 and 3.56 respectively. It can be seen that $-\mu_{rR} = \mu_{rL}$ (i.e., $\bar{\mu} = 0$) at $f_{\bar{\mu}=0} = 1.06\,\mathrm{GHz}$, and $-\varepsilon_{rR} = \varepsilon_{rL}$ (i. e., $\bar{\varepsilon} = 0$) at $f_{\bar{\varepsilon}=0} = 1.74\,\mathrm{GHz}$. The frequency band with $\bar{\varepsilon} \cdot \bar{\mu} < 0$. Therefore, the Coupling occurs at the frequency band where the average permittivity and permeability satisfy $\bar{\varepsilon} \cdot \bar{\mu} < 0$. Therefore, the coupling region must satisfy $f_{\bar{\mu}=0} < f < f_{\bar{\varepsilon}=0}$ for ENG CRLH TL/RH TL coupler.

For an asymmetric ENG CRLH/RH TL coupler, the coupling region satisfies $\bar{\varepsilon} \cdot \bar{\mu} < 0$ and

$\mu_{rL} < 0$, $\varepsilon_{rL} < 0$ at the same time, therefore, the coupling region is $f_{\bar{\mu}=0} = 1.06\,\text{GHz} < f < f'_{se} = 1.62\,\text{GHz}$. Fig. 6 shows that coupling region is from 1.05 GHz to 1.65 GHz and confirms our theoretical prediction.

3.2. MNG CRLH TL/RH TL Coupler

If the stopband of the CRLH TL is mu-negative, the stopband is $f'_{sh} < f < f'_{se}$ and the left-handed region is $f_c < f < f'_{sh}$. For an asymmetric MNG CRLH TL/RH TL coupler, the coupling region satisfy $f_{\bar{\varepsilon}=0} < f < f'_{sh}$.

Another MNG CRLH TL is designed. The width of the interconnecting transmission line is still 2.945 mm. The length of the unit is 7 mm, The loading series capacitors are $C_L = 1\,\text{pF}$ and shunt inductors $L_L = 4.7\,\text{nH}$. The gap edges are derived to be $f_{se} = 3.36\,\text{GHz}$ and $f_{sh} = 2.45\,\text{GHz}$. The cutoff frequency is $f_c = 1.16\,\text{GHz}$.

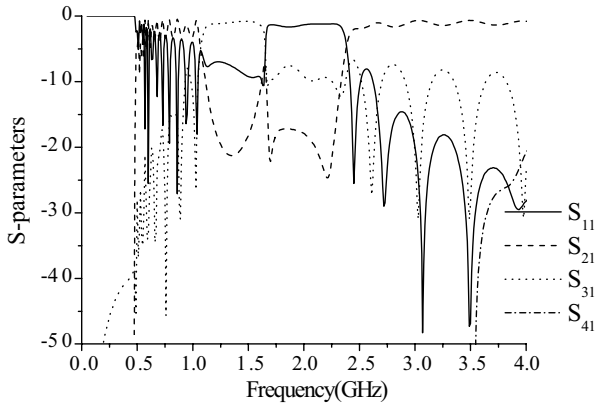


Figure 6: Simulating S-parameters for the CRLH/RH TL coupler.

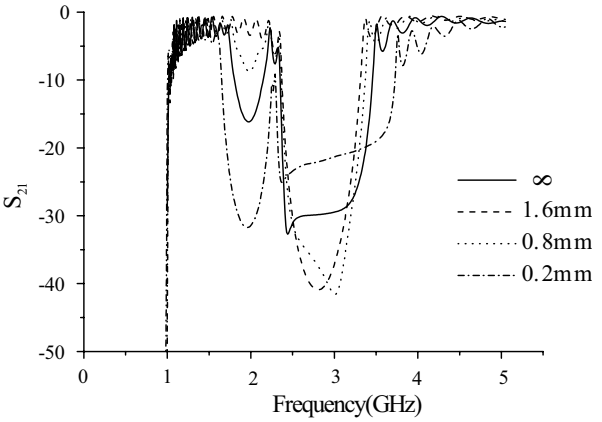


Figure 7: S_{21} of the MNG CRLH TL when the spacing between the lines is changing from 0.2 mm to infinity large.

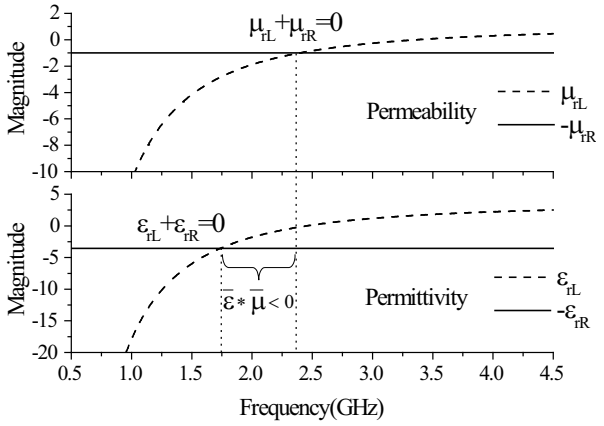


Figure 8: Relative premittivity ε_{rL} and permeability μ_{rL} of the MNG CRLH TL and relative premittivity ε_{rR} and permeability μ_{rR} of the RH TL.

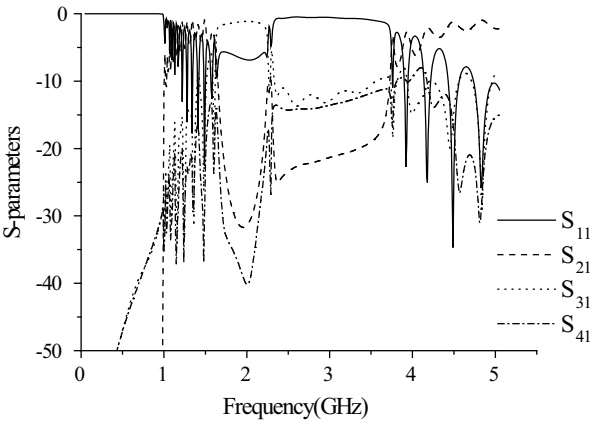


Figure 9: Simulating S-parameters for the MNG CRLH/RHTL coupler.

As same as the case of ENG, when a conventional right-handed transmission line is near this CRLH TL, coupling will influence the edges of the bandgap. As we have mentioned above, the distributed series inductance and shunt capacitance of the CRLH TL will change to $C'_R = 141.8\,\text{pF/m}$ and $L'_R = 260.0\,\text{nH/m}$ when the spacing between the two lines is 0.2 mm. Therefore, the gap edges

become

$$f'_{se} = \frac{1}{2\pi\sqrt{L'_R C_L d}} = 3.67\,\text{GHz}$$
$$f'_{sh} = \frac{1}{2\pi\sqrt{C'_R L_L d}} = 2.33\,\text{GHz}$$

Figure 7 shows f_{se} moves to 3.73 GHz and f_{sh} moves to 2.25 GHz. In this case, the left-handed passband is $f_c < f < f'_{sh}$ for MNG CRLH TL.

The permeability μ_{rL} and permittivity ε_{rL} of the CRLH TL and the negative μ_{rR} and ε_{rR} of the RH TL are shown in Fig. 8. It can be seen that $-\varepsilon_{rR} = \varepsilon_{rL}$ at $f_{\bar{\varepsilon}=0} = 1.67\,\text{GHz}$ and $-\mu_{rR} = \mu_{rL}$ at $f_{\bar{\mu}=0} = 2.38\,\text{GHz}$.

The coupling S-parameter of a MNG CRLH TL/RH TL coupler when the space is 0.2 mm is demonstrated in Fig. 9. It shows the coupling region from 1.67 GHz to 2.25 GHz, satisfies $f_{\bar{\varepsilon}=0} < f < f'_{sh}$.

4. CONCLUSION

The coupling region of the CRLH/RH TL coupler is studied. The coupling region of the CRLH/RH TL coupler satisfy $\bar{\varepsilon} \cdot \bar{\mu} < 0$ and $\mu_{rL} < 0$ and $\varepsilon_{rL} < 0$ at the same time. When the coupling occurs between an ENG CRLH TL and a RH TL, the lower edge of the coupling region defined by $\mu_{rL} + \mu_{rR} = 0$. The higher edge of the coupling region is f'_{se}. The stop band of the CRLH TL becomes $f'_{se} < f < f'_{sh}$. When the coupling occurs between a MNG CRLH TL and a RH TL, the lower edge of the coupling region defined by $\varepsilon_{rL} + \varepsilon_{rR} = 0$. The higher edge of the coupling region is $f'sh$. The stop band of the CRLH TL becomes $f'_{sh} < f < f'_{se}$.

ACKNOWLEDGMENT

This research was supported by National Basic program (973) of China (No. 2001CB610406 and No. 2004CB719802) and by the National Natural Science Foundation of China (No. 50477048 and No. 10474072).

REFERENCES

1. 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.
2. 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.
3. Oliner, A. A., "A periodic-structure negative-refractive-index medium without resonant elements," *IEEE AP-S/URSI Int. Symp. Dig.*, 41, 2002.
4. Grbic, A. and G. V. Eleftheriades, "Experimental verification of backward-wave radiation from a negative index metamaterial," *J. Appl. Phys.*, Vol. 92, No. 10, 5930–5934, Nov. 2002.
5. Antoniades, M. and G. V. Eleftheriades, "Compact, linear, lead/lag metamaterial phase shifters for broadband applications," *IEEE Antennas Wireless Propag. Lett.*, Vol. 2, No. 7, 103–106, 2003
6. Caloz, C., A. Sanada, and T. Itoh, "A novel composite right/left-handed coupled-line directional coupler with arbitrary coupling level and broad bandwidth," *IEEE Trans. Micro. Theo. Tech.*, Vol. 52, 980–992, 2004.
7. Caloz, C. and T. Itoh, "A novel mixed conventional microstrip and composite right/left-handed backward-wave directional coupler with broadband and tight coupling characteristics," *IEEE Micro. Wireless Comp. Lett.*, Vol. 14, No. 1, 31–33, 2004.
8. Islam, R. and G. V. Eleftheriades, "A planar metamaterial co-directional coupler that couples power backward," *IEEE-MTT Int. Symp. Dig.*, 321–324, June 2003.
9. Islam, R. and G. V. Eleftheriades, "Printed high-directivity metamaterial MS/NRI coupled-line coupler for signal monitoring applications," *IEEE Micro. Wireless Comp. Lett.*, Vol. 16, No. 4, 164–166, April 2006.
10. Wang, Y., Y. Zhang, L. He, F. Liu, H. Li, and H. Chen, "Discussion on coupling mechanism of asymmetric CRLH/RH TL coupler," *J. Zhejiang Univ. Science A*, 95–98, 2006.

11. Kirschning, M. and R. H. Jansen, "Accurate wide-range design equations for the frequency-dependent characteristic of parallel coupled microstrip lines," *IEEE Trans. Micro. Theo. Tech.*, Vol. MTT-32, 83–90, Jan. 1984.

Experimental Investigation of Relationship between the object- and Image Distance

Kun Ren[1], **Xiaobin Ren**[2], **Xiangdong Zhang**[2]
Zhifang Feng[3], **Zhiyuan Li**[1], **and Daozhong Zhang**[1]

[1]Beijing National Laboratory for Condensed Matter Physics, Institute of Physics
Chinese Academy of Sciences, China
[2]Department of Physics, Beijing Normal University, China
[3]School of Science, Beijing University of Chemical Technology, China

Abstract— We experimentally investigate the focusing properties of a triangular two-dimensional photonic crystals of which all the cylinders are coated. The relationship between the location of object and image point is analyzed. It's demonstrated that at a certain frequency with relative refractive index of -1, the position of image point obey the geometric-optics analysis and non-near-field imaging can be achieved.

1. INTRODUCTION

Recently a lot of interest has been focused on negative refraction of electromagnetic (EM) wave in the left-hand-materials (LHMs) [1–3] or photonic crystals (PCs) [4–14]. A promising application is that a parallel-sided slab made of NIR materials can focus electromagnetic (EM) waves just like a lens [1]. This flat lens possesses some advantages over conventional lenses. For example, it can partly overcome the diffraction limit and achieve subwavelength resolution [6–8].

Negative refraction in the PCs can be analyzed by the equifrequency surface (EFS) of the band structures [4–13]. The dispersion properties of the PC are determined by the structure and various parameters of the system such as dielectric constant, size of the scatterers and so on. Due to the anisotropy of the dispersion in 2D PC, the images in many cases only appear in near-field region, and the image distance has little dependence on the source distance [7–9]. Recently, non-near-field images, which obey fairly well the image distance relationship characteristic of an $n = -1$ material, have been theoretically obtained in 2D PCs [11–14]. In this paper, we experimentally investigate negative refraction of a triangular two-dimensional photonic crystals consisting of coated cylinders. The relationship of the location of object and image point is analyzed at a certain frequency with relative refractive index of -1.

2. MEASUREMENT OF REFRACTIVE ANGLE

The sample used in our experiments consists of a number of coated cylinders immersed in a Styrofoam template. The coated cylinders have metallic cores coated with a dielectric coating, which has been investigated theoretically in Ref. 13. The radii of metallic core and coated cylinder are $0.25a$ and $0.45a$, respectively, where a ($= 12\,\text{mm}$) is the lattice constant. The dielectric constant of dielectric coating is 7.0. We use copper as the metallic core. In the experiments, we only consider the transverse magnetic (TM) modes (S wave), where the electric field is kept parallel to the extension axis of the metallic rods.

The refraction experiments are performed in a semicircular cavity. Figure 1 shows the experimental setup. The wedge shaped samples with different wedge angles are used. When incident wave impinges on the wedge interface the refraction wave travels either on the right side (positive refraction) or the left side (negative refraction) of the surface normal. To detect the refraction beam, a dipole antenna is mounted on a goniometer that can be rotated around the outer edge of the semicircular cavity. By examining the variation of the refraction angle θ with the incident angle θ_0, we can obtain the refractive index. It's found that at $12.7\,\text{GHz}$ θ_0 is linearly proportional to θ_0 in the whole angle region. We can conclude that at this frequency the relative refractive index is -1. The direction of the refracted wave inside the PC can be estimated from the EFS. Our experimental result agrees well with the theoretical estimation.

3. MEASUMENT OF IMAGE DISTANCE

We examine a series of slab samples with different thickness. A $400\,\text{mm}$ wide and $75\,\text{mm}$ thick slab sample are taken as the first example. A monochromatic point source radiating a frequency

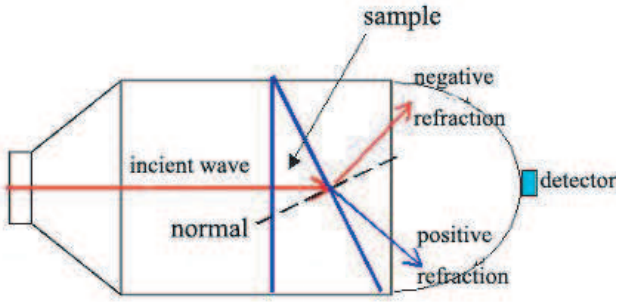


Figure 1: Setup to measure the refraction angle. A wedge shaped sample is put in the parallel aluminum plate.

of 12.7 GHz is placed at a distance 37 mm from the left surface of the slab. Its noticed that the image is formed at a distance about half thickness of slab from the right surface slab. When the 115 mm thick sample is taken the point source is also placed at the distances of half thickness of the samples from the slab. We find that the image-slab distance (image distance or V) is still equal to the object-slab distance (object distance or U), which means that the summation of them is about the thickness of the slab (L).

The above results are for the cases in which the point sources are placed at a distances of half thickness of the samples. We further change the position of source and examine the relationship between the object distance and the image distance. The experiments were carried out for three different source distances, namely 32, 58 and 83 mm. Figure 2 shows the experiment and simulation results for the same sample under different source distance. It can be seen that the image distance varies with the object distance, and our further data analysis shows the summation of image- and object distance keeps the same. The imaging obeys fairly well the formula $U + V = L$.

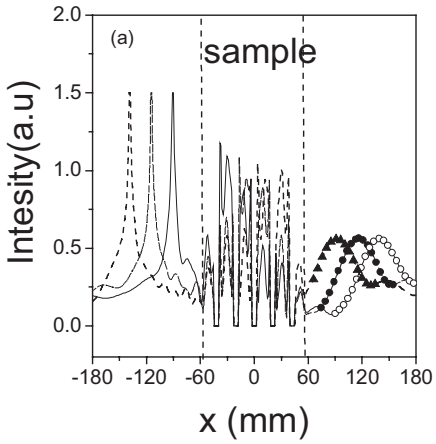


Figure 2: Intensity distribution along the light propagation direction and across the source for the same sample under different source distance. The distances of point source are 32 mm (solid lines), 58 mm (short dots lines) and 82 mm (dash lines), respectively. The experimental results are corresponding represented by triangles, dots and circles, respectively.

4. CONCLUSIONS

In conclusion, we experimentally investigate the imaging property of PCs that consist of the coated cylinders with the triangular lattice. This PC can act as an isotropic effective metamedium with relative refractive index $n = -1$. It's demonstrated that the image distance plus the object distance is equal to the thickness of PC slab. The location of image point follows the well-known wave-beam negative refraction law.

ACKNOWLEDGMENT

This work was supported by the National Key Basic Research Special Foundation of China (Grant Nos. 2001CB610402 and 2004CB719804). The support from the Supercomputing Centre, CNIC, CAS is acknowledged.

REFERENCES

1. Veselago, V. G., "The electrodynamics of substances with simultaneously negative values of permittivity and permeability," *Sov. Phys. Usp. 10*, Vol. 509, No. 10, 509–514, 1968.
2. Shelby, R. A., D. R. Smith, and S. Schultz, "Experimental veritcation of a negative index of refraction," *Science*, Vol. 292, 77–79, 2001.
3. Pendry, J. B., "Negative refraction makes a perfect lens," *Phys. Rev. Lett.*, Vol. 85, No. 18, 3966–3969, 2000.
4. Notomi, M., "Theory of light propagation in strongly modulated photonic crystals: Refractionlike behavior in the vicinity of the photonic band gap," *Phys. Rev. B*, Vol. 62, No. 16, 10696–10705, 2000.
5. Luo, C., S. G. Johnson, J. D. Joannopoulos, and J. B. Pendry, "All-angle negative refraction without negative effective index," *Phys. Rev. B*, Vol. 65, No. 20, 201104, 2002.
6. Parimi, P. V., W. T. Lu, P. Vodo, and S. Sridhar, "Imaging by flat lens using negative refraction," *Nature*, Vol. 426, 404, 2003.
7. Luo, C., S. G. Johnson, J. D. Joannopoulos, and J. B. Pendry, "Subwavelength imaging in photonic crystals," *Phys. Rev. B*, Vol. 045115, No. 15, 68, 2003.
8. Li, Z. Y. and L. L. Lin, "Evaluation of lensing in photonic crystal slabs exhibiting negative refraction," *Phys. Rev. B*, Vol. 68, No. 24, 245110, 2003.
9. Ren, K., S. Feng, Z. F. Feng, Y. Seng, Z. Y. Li, B. Y. Cheng, and D. Z. Zhang, "Imaging properties of triangular lattice photonic crystal at the lowest band," *Phys. Lett. A*, Vol. 348, No. 3–6, 405–409, 2006.
10. Ren, K., Z. Y. Li, X. B. Ren, S. Feng, B. Y. Cheng, and D. Z. Zhang, "Three-dimensional light focusing in inverse opal photonic crystals," *Phys. Rev. B*, Vol. 75, No. 11, 115108, 2007.
11. Wang, X., Z. F. Ren, and K. Kempa, *Optics Express*, Vol. 12, 2919, 2004.
12. Hu, X. and C. T. Chan, "Photonic crystals with silver nanowires as a near-infrared superlens," *Appl. Phys. Lett.*, Vol. 85, No. 9, 1520, 2004.
13. Zhang, X., "Image resolution depending on slab thickness and object distance in a two-dimensional photonic-crystal-based superlens," *Phys. Rev. B*, Vol. 70, No. 19, 195110, 2004.
14. Feng, Z., X. Zhang, Y. Q. Wang, Z. Y. Li, B. Y. Cheng, and D. Z. Zhang, *Phys. Rev. Lett.*, Vol. 94, No. 24, 247402, 2005.

Relativistic Energy Loss and Induced Photon Emission in the Interaction of a Left-handed Sphere with an External Electron Beam

Xiangdong Zhang and Jinying Xu

Department of Physics, Beijing Normal University, Beijing 100875, China

Abstract— Relativistic energy loss and photon emission in the interaction of a left-hand sphere with an external electron beam are studied based on the classical electrodynamics. Comparing the energy loss spectra and the photon emission of the left-hand sphere with the right-hand one, we find that they can exhibit some new features. The features are directly related to the unusual properties of the left-handed materials. That is to say, our results show that some unusual properties of the left-hand material can be explored by scanning transmission electron microscopy (STEM).

1. INTRODUCTION

Scanning transmission electron microscopy (STEM) has proved to be a powerful technique for determining different microstructures of nanometer scale [1]. The interaction between the electron and the microstructures gives rise to the emission and excitation of the radiation. Thus, the electron energy loss spectroscopy in the STEM is also a useful tool to investigate both surface and bulk excitations of the samples [2–9]. In the last decades a remarkable progress has been made on this subject [2–9]. Various geometrical objects have been investigated and many interesting results have been obtained [2–9]. However, the above investigations all focus on the dielectric or metal systems.

Recently, left-handed materials (LHMs) have attracted a great deal of attention from both the theoretical and experimental sides [10–13]. These materials are characterized by simultaneous negative permittivity and permeability. Properties of such materials were analyzed theoretically by Veselago over 30 years ago [10], but only recently they were demonstrated experimentally [11]. As was shown by Vesselago, the left-handed materials possess a number of unusual electromagnetic effects including negative refraction, inverse Snell's law, reversed Doppler shift, and reversed Cerenkov radiation [10, 12, 13]. New surface polarized modes around left-handed slab, cylinder and sphere have also been demonstrated theoretically [14, 15]. Then, whether these unusual properties and new surface modes can be explored by the STEM? The investigation on such a problem is the center issue of the present paper.

2. THEORY

We consider a fast electron moving along a straight-line trajectory with constant velocity v and passing near a homogeneous sphere located in vacuum. The relative permittivity $\varepsilon_1(\omega)$ and the relative permeability $\mu_1(\omega)$ are given by [11, 14, 15]

$$\varepsilon_1(\omega) = 1 - \frac{\omega_p^2}{\omega(\omega + i\gamma)} \tag{1}$$

and

$$\mu_1(\omega) = 1 - \frac{F\omega^2}{\omega^2 - \omega_0^2 + i\omega\Gamma}, \tag{2}$$

where ω_p is plasma frequency, γ and Γ are the respective electric and magnetic loss terms. ω_0 is the magnetic resonance frequency.

Such a electron approaching the sphere may engender collective excitations which act back upon the electron, which leads to the energy loss of the moving electron. The energy loss can be related to the force exerted by the induced electric field E^{ind} acting on it as

$$\Delta E^{\text{loss}} = \int dt\vec{v} \cdot \vec{E}^{ind}(\vec{r}_t, t) = \int_0^\infty \omega d\omega \Gamma^{\text{loss}}(\omega), \tag{3}$$

where

$$\Gamma^{\rm loss}(\omega) = \frac{1}{\pi\omega}\int dt {\rm Re}\left\{e^{-i\omega t}\vec{v}\cdot\vec{E}^{ind}\left(\vec{r}_t,\omega\right)\right\} \tag{4}$$

is the so-called loss probability, which can be divided into the contributions of magnetic and electric modes,

$$\Gamma^{\rm loss} = \Gamma^{M,{\rm loss}} + \Gamma^{E,{\rm loss}}. \tag{5}$$

Here

$$\Gamma^{M,{\rm loss}}(\omega) = \sum_{l=1}^{\infty}\sum_{m=-1}^{l}\frac{mv}{\pi\omega^2}K_m\left(\frac{\omega b}{v\gamma}\right)\times{\rm Re}\left\{\left(A_{lm}^{+}\right)^* e^{im\varphi_0+i\omega z_0/v_i-l}\psi_{lm}^{M,ind}\right\} \tag{6}$$

and

$$\Gamma^{E,{\rm loss}}(\omega) = \sum_{l=1}^{\infty}\sum_{m=-l}^{l}\frac{c}{2\pi\omega^2\gamma}K_m\left(\frac{\omega b}{v\gamma}\right)\times{\rm Re}\left\{B_{lm}^{*}e^{im\varphi_0+i\omega z_0/v_i-l}\psi_{lm}^{E,ind}\right\}. \tag{7}$$

where $\psi_{lm}^{M,ind}$ and $\psi_{lm}^{E,ind}$ represent the induced scattered fields, which can be expressed as

$$\psi_{lm}^{M,ind} = t_l^M\psi_{lm}^{M,ext} \tag{8}$$

and

$$\psi_{lm}^{E,ind} = t_l^E\psi_{lm}^{E,ext}. \tag{9}$$

Here t_l^M and t_l^E are the scattering-matrix elements [9]. $\psi_{lm}^{M,ext}$ and $\psi_{lm}^{E,ext}$ can be obtained from the bare electric field produced by the moving electron in vacuum as

$$\psi_{lm}^{M,ext} = \frac{-4\pi i^{1-l}kv}{c^2}\frac{mA_{lm}^{+}}{l(l+1)}K_m\left(\frac{\omega b}{v\gamma}\right)e^{-im\varphi_0-i\omega z_0/v} \tag{10}$$

and

$$\psi_{lm}^{E,ext} = \frac{-2\pi i^{1-l}k}{c\gamma}\frac{B_{lm}}{l(l+1)}K_m\left(\frac{\omega b}{v\gamma}\right)e^{-im\varphi_0-i\omega z_0/v}. \tag{11}$$

where

$$B_{lm} = A_{l,m+1}^{+}\sqrt{(l+m+1)(l-m)} - A_{l,m-1}^{+}\sqrt{(l-m+1)(l+m)}, \tag{12}$$

$$A_{lm}^{\pm} = \frac{1}{\beta^{l+1}}\sum_{j=m}^{l}\frac{C_j^{lm,\pm}}{\gamma^j}, \tag{13}$$

$$C_j^{lm,\pm} = \frac{(\pm i)^{l-j}\alpha_{lm}(2l+1)!!}{2^j(l-j)![(j-m)/2]![(j+m)/2]!}I_{j,l-j}^{lm} \tag{14}$$

and

$$I_{i_1,i_2}^{lm} = (-1)^m\int_{-1}^{1}d\mu\left(1-\mu^2\right)^{i_1/2}\mu^{i_2}P_l^m(\mu). \tag{15}$$

The coupling of the electron with radiation modes of the sphere gives rise to radiation emission

$$\Delta E^{\rm rad} = \frac{c}{4\pi}\int dt\int d\Omega r^2\left[\vec{E}\left(\vec{r},t\right)\times\vec{H}\left(\vec{r},t\right)\right]\cdot\vec{r}, \tag{16}$$

where $\vec{r}$ points to the surface of the large sphere and the integral over the time has been included. Expressing the fields in terms of their frequency components, one finds

$$\Delta E^{\rm rad} = \int_0^{\infty}\omega d\omega\int d\Omega\Gamma^{\rm rad}(\omega,\Omega), \tag{17}$$

where

$$\Gamma^{\rm rad}(\omega,\Omega) = \frac{r^2}{4\pi^2 k}{\rm Re}\left\{\left[\vec{E}(\omega)\times\vec{H}(-\omega)\right]\cdot\vec{r}\right\} \tag{18}$$

is the probability of emitting a photon of energy ω per unit energy range and unit solid angle around the direction Ω_r. In the $\vec{r} \to \infty$ limit, only the induced fields contribute to the radiation. Integrating over angles, one finds

$$\Gamma^{\mathrm{rad}}(\omega) = \frac{1}{c\omega} \sum_{l=1}^{\infty} \sum_{m=-l}^{l} K_m^2 \left(\frac{\omega b}{v\gamma} \right) \left[C_{lm}^M \left| t_l^M \right|^2 + C_{lm}^E \left| t_l^E \right|^2 \right]. \tag{19}$$

3. NUMERICAL RESULTS AND DISCUSSION

In this section, we present numerical results for the spectroscopy of the relativistic energy loss and the photon emission of the left-handed sphere based on Eqs. (5) and (19). In the numerical calculations, we use the value $\omega_p = 10\,\mathrm{GHz}$ for the plasma frequency, and $\omega_0 = 4\,\mathrm{GHz}$ for the magnetic resonance frequency [14]. For the damping parameters, the values $\gamma = 0.03\omega_p$ and $\gamma = 0.03\omega_0$ are assumed and $F = 0.56$ [14]. From these parameters, we find that the sphere is left-handed (negative n) between 4 and 6 GHZ.

In order to understand well the properties of the relativistic energy loss (Γ^{loss}) and induced photon emission (Γ^{rad}) in the interaction of the left-handed sphere with an external electron beam, we first consider the case of the dielectric dispersion ($\mu = 1$ and permittivity given by Eq. (1)). Dotted lines in Figs. 1(a) and (b) represent such a case with the sphere radius $\omega_p a/c = 4$ and $\omega_p a/c = 12$, respectively. Here $b = a$ and $v/c = 0.9$ are taken. In general, the loss probability is related to the different combinations of electron velocity, impact parameters, sphere radius and dielectric functions. These relations have been disclosed in the previous investigations [1–9]. Our results are agreement with them. However, the situations can be changed when both the dielectric and the magnetic dispersions are considered. Solid lines in Fig. 1 describes such a case. It is clearly seen that new exciting mode (marked by arrow in Fig. 1(a)) appears around 4.7 GHZ (in negative n region). For small particle size, the loss spectroscopy of the left-handed sphere is similar to that of the right-handed one except these new exciting modes. With the increase of the particle size, the differences appear. At $\omega_p a/c = 12$ (Fig. 1(b)), large difference can be observed. The loss of the left-handed sphere is smaller than that of the right-handed one due to the properties of the left-handed material which leads to "invisible" or "transparent" to an outside observer. These indicate that the STEM can disclose the unusual properties of the left-hand material.

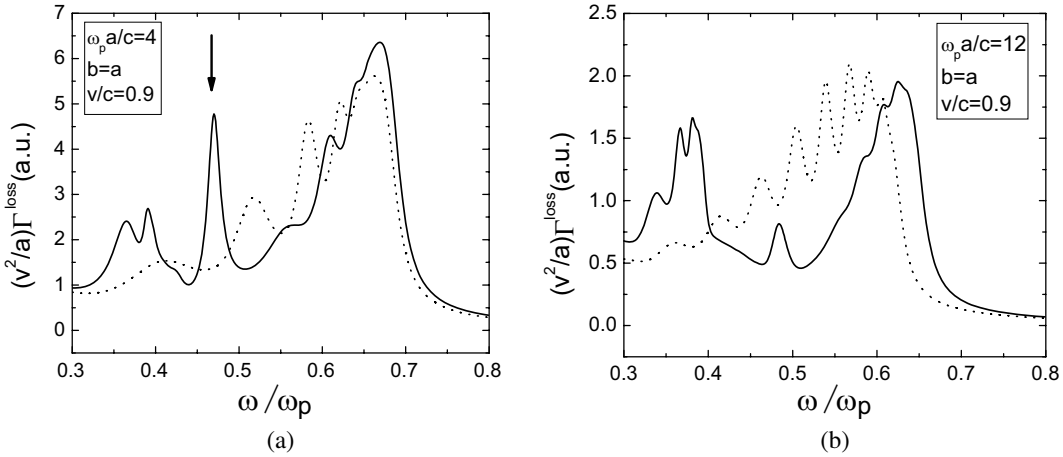


Figure 1: Energy-loss spectra as a function of energy range ω for electrons passing near a sphere with radius $\omega_p a/c = 4$ (a) and $\omega_p a/c = 12$ (b). Solid lines correspond to the left-handed case and dotted lines to right-handed one.

Analogous to the energy loss, the radiation emission has also been calculated. The corresponding results are plotted in Figs. 2(a) and (b). Radiation emission Γ^{rad} is always smaller than Γ^{loss} due to the presence of the absorption. It is interesting that some exciting modes that are present in the loss probability are nearly absent in the photon emission probability indicating the different character of the different modes.

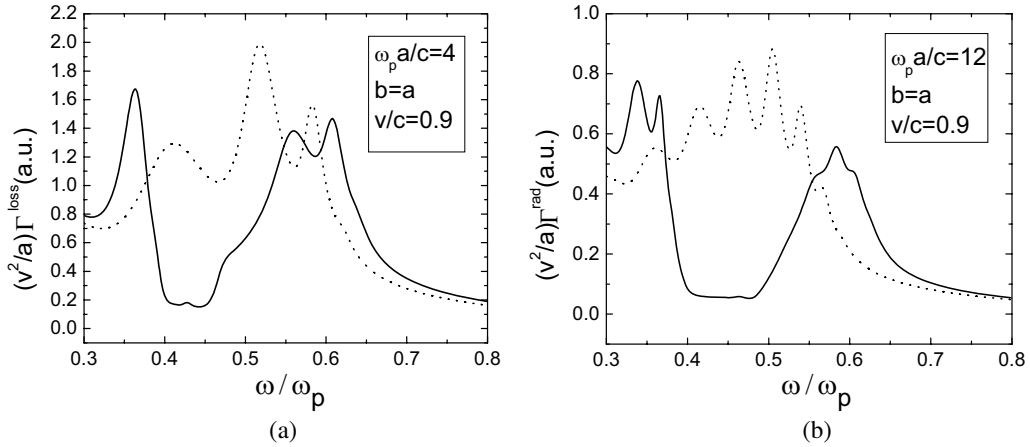


Figure 2: Radiation emission probability as a function of energy range ω for electrons passing near a sphere with radius $\omega_p a/c = 4$ (a) and $\omega_p a/c = 12$ (b). Solid lines correspond to the left-handed case and dotted lines to right-handed one.

REFERENCES

1. Nellist, P. D. and S. J. Pennycook, *Phys. Rev. Lett.*, Vol. 81, 4156, 1998; *Phys.Rev.B*, Vol. 96, 575, 1995.
2. Ritchie, R. H., *Phys. Rev.*, Vol. 106, 874, 1957.
3. Echenique, P. M. and J. B. Pendry, *J. Phys. C*, Vol. 8, 2936, 1975.
4. Zabala, N., A. Rivacoba, and P. M. Echenique, *Surf. Sci.*, Vol. 209, 465, 1989.
5. Garcia-Molina, R., A. Gras-Marti, and R. H. Ritchie, *Phys. Rev. B*, Vol. 31, 121, 1985.
6. Ferrell, T. L. and P. M. Echenique, *Phys. Rev. Lett.*, Vol. 55, 1526, 1985.
7. Ferrell, T. L., R. J. Warmack, V. E. Anderson, and P. M. Echenique, *Phys. Rev. B*, Vol. 35, 7365, 1987.
8. Pendry, J. B. and L. Martin-Moreno, *Phys. Rev. B*, Vol. 50, 5062, 1994.
9. Garcia de Abajo, F. J. and A. Howie, *Phys. Rev. Lett.*, Vol. 80, 5180, 1998.
10. Veselago, V. G., *Sov. Phys. Usp.*, Vol. 10, 509, 1968.
11. Smith, D. R., W. J. Padilla, D. C. Vier, S. C. Nemat-Nasser, and S. Schultz, *Phys. Rev. Lett.*, Vol. 84, 4184, 2000; Shelby, R. A., D. R. Smith, and S. Schultz, *Science*, Vol. 292, 77, 2001.
12. Pendry, J. B., *Phys. Rev. Lett.*, Vol. 85, 3966, 2000.
13. Pacheco Jr., J., T. M. Grzegorczyk, T. B. I. Wu, Y. Zhang, and J. A. Kong, *Phys. Rev. Lett.*, Vol. 89, 257401, 2002.
14. Ruppin, R., *J. Phys. C*, Vol. 13, 1811, 2001; *Phys. Lett. A*, Vol. 277, 61, 2000.
15. Ruppin, R., *J. Phys. C*, Vol. 16, 5991, 2004; *Solid State Communications*, Vol. 116, 411, 2000.

Principal Component Analysis (PCA) in the Context of Radar Polarimetry

W.-M. Boerner[1], **E. Lüneburg**[2], **and A. Danklmayer**[3]

[1]University of Illinois at Chicago, USA
[2]EML Consultants, Germany
[3]German Aerospace Center, Germany

Abstract— Statistical and computational techniques for revealing the internal structure that underlies the set of random correlated data exists in a great variety at present; and target decomposition theorems, either in the coherent or incoherent formulation, are well established. In spite of this fact a rather innovative and new concept is presented in this contribution. In turn the Principal Component Analysis (PCA) is considered to possibly add value to existing approaches, and it allows for an interpretation of polarimetric synthetic aperture radar measurements using variables obtained via linear transformation. Starting with the Sinclair backscatter matrix S which will be further transformed into the so called target feature vector by stacking column elements of S and generating the covariance matrices averaged over a certain pixel array, we show, how the Sinclair backscatter matrix is decomposed into the sum of a maximum of four 2×2 elementary point scatter matrices which are weighted by the principal components, whereas the variances of these components agree with the eigenvalues of the covariance matrix. This mathematical development defines a decomposition which expresses scattering mechanism from distributed targets in terms of scattering matrices via an incoherent step.

1. INTRODUCTION

In radar and optical polarimetry there exists essentially two different methods to characterize polarimetric scattering properties of plane electromagnetic waves scattered by randomly distributed targets using second order multivariate statistics [1–3]. The first is the Kennaugh matrix which is used for finding solutions for maximal and minimal power transfer between transmitting and receiving antennas. The second is the covariance matrix of which the analysis is used for entropy and variance considerations and for the generation of uncorrelated random variables.

The basic problem of principal component analysis [4, 5, 8] is how to find a suitable representation of multivariate data in order to make the essential structure more visible and to identify any distinct feature. The main purpose of PCA is to convert a set of possibly correlated random variables into new uncorrelated variables.

In radar polarimetry and interferometry, the principal component analysis is known as target decomposition theory, and relevant contributions are due to [1, 2, 6, 9–13] and many others. Only covariance matrices rather than correlation matrices are used. Considerable emphasis is paid to the interpretation of the coefficients and the eigenvectors of the covariance matrix. Confusion still prevails with respect to the correct interpretation of the principal components (the new uncorrelated orthonormal random variables) and the only superficially related concepts of the classical Sinclair/Kennaugh and the Jones/Mueller matrices. The principal component analysis (PCA) relies on two basic requirements of a random vector: (i) to find linear functions of the components of the random vector having maximum variance in subsequent steps; and (ii) that these new random variables should be orthonormal and *uncorrelated*. These requirements imply that the first few components carry most of the total variation of the original random vector and are representative for multivariate random processes.

2. PRINCIPLE COMPONENT ANALYSIS (PCA)

With principal component analysis one tries to find new variables which are uncorrelated but not necessary independent. As a starting point the scattering matrix represented by the 2×2 S-matrix is considered, describing completely the polarization transforming properties of a target at a single frequency in the reference direction.

$$S(t) = \begin{pmatrix} S_{HH}(t) & S_{HV}(t) \\ S_{VH}(t) & S_{VV}(t) \end{pmatrix} \tag{1}$$

The indices of the matrix elements represent the transmit and receive polarization of the plane electromagnetic wave and t stands for time or ensemble values. The *vec* operator [7] is used to arrive at the target feature vector k given by

$$k_4(t) = \mathrm{vec}\, S(t) = \begin{bmatrix} S_{HH}(t) \\ S_{VH}(t) \\ S_{HV}(t) \\ S_{VV}(t) \end{bmatrix} \tag{2}$$

where the "vec" operator can be considered as a simple stacking of the columns of the scattering matrix $S(t)$. This target feature vector is used to calculate the 4×4 covariance matrix explicitly according to

$$C_4 = < k_4(t) k_4^{\dagger}(t) > . \tag{3}$$

where the dagger symbol represents complex conjugation and transposition (Hermition conjugate) and the angle brackets denote the averaging of the outer product of $k_4(t)$.

Being Hermitian positive semidefinite, the covariance matrix C_4 can be unitarily diagonalized by a unitary matrix with non-negative eigenvalues.

$$U^{-1} C_4 U = \Lambda \equiv \mathrm{diag}[\lambda_1, \lambda_2, \lambda_3, \lambda_4] \text{ with } U = [\hat{x}_1, \hat{x}_2, \hat{x}_3, \hat{x}_4]. \tag{4}$$

$\lambda_i \, (i = 1, \ldots, 4)$ are denoting the eigenvalues of the covariance matrix, and $x_i \, (i = 1, \ldots, 4)$ are the eigenvectors, respectively. Now we introduce the new target feature vector $Z(t)$ by a linear transformation

$$Z(t) = \begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \end{bmatrix} = U^{\dagger} k(t) = \begin{bmatrix} x_1^{\dagger} k(t) \\ x_2^{\dagger} k(t) \\ x_3^{\dagger} k(t) \\ x_4^{\dagger} k(t) \end{bmatrix} \text{ or } k(t) = U \cdot Z(t) \tag{5}$$

The components $z_i \, (i = 1, \ldots, 4)$ are called the principal components (the new uncorrelated orthonormal random variables). They are uncorrelated new random variables and their variances are equal to the corresponding eigenvalues of C_4. Furthermore the spectral decomposition of the Covariance matrix is given by

$$C_4 = U \Lambda U^{\dagger} = \sum_{i=1}^{4} \lambda_i x_i x_i^{\dagger} = \sum_{i=1}^{4} \lambda_i C_{4,i} \tag{6}$$

where C_{4i} are 4×4 covariance matrices of point targets with rank 1. The reverse *vec* operator may be applied to $\boldsymbol{x}_i$ and the results can be interpreted as 2×2 elementary deterministic point targets S_i with $\mathrm{span}(S_i) = 1$.

$$x_i = \mathrm{vec} S_i = \begin{bmatrix} x_{i,1} \\ x_{i,2} \\ x_{i,3} \\ x_{i,4} \end{bmatrix} \Leftrightarrow S_i = \begin{bmatrix} x_{i,1} & x_{i,3} \\ x_{i,2} & x_{i,4} \end{bmatrix} \; (i = 1, \ldots, 4) \tag{7}$$

and hence using the relation $k(t) = UZ(t)$

$$S(t) = \begin{bmatrix} S_{xx} & S_{xy} \\ S_{yx} & S_{yy} \end{bmatrix} = \sum_{i=1}^{4} z_i(t) \begin{bmatrix} x_{i1} & x_{i3} \\ x_{i2} & x_{i4} \end{bmatrix} = \sum_{i=1}^{4} z_i(t) S_i \tag{8}$$

This is an expansion of the scattering matrix into four uncorrelated point targets formed by the loading vectors or loading matrices with random coefficients $z_i(t)$. For reciprocal targets, the matrix $S(t)$ is symmetric; then the elementary point target matrices $S_i(t)$ are also symmetric, i.e., $x_{i2} = x_{i3}$ for all i.

The basic targets are orthonormal in the sense that

$$(\mathrm{vec} S_i)^{\dagger} \mathrm{vec} S_j = x_i^{\dagger} x_j = \sum_{k=1}^{4} x_{ik}^{*} x_j k = \sum_{k=1}^{4} |x_{ik}|^2 \delta_{ij} = \delta_{ij} \tag{9}$$

for all $i, j = 1, \ldots, 4$.

3. APPLICATION OF PCA TO POLARIMETRIC SAR DATA

Calculations have been carried out following the mathematical developments presented in the previous chapter. A polarimetric SAR data set (L-band) was used and Fig. 1 shows a RGB-image of the testsite Oberpfaffenhofen, Germany. The different channels of the image are red for HH-, green for HV-, and blue color for VV-polarization. The window size for the applied averaging of the covariance matrix was chosen 5×5 pixels. At the end of the calculations the result is the extension of the scattering matrix into the sum of four matrices (cf. right hand side of (8)). These matrices represent the product of the principal components $z_i(t)$ with the elementary point scatter matrix and the elements of these matrices were used for visualization of the results.

4. RESULTS

The image which is presented in Fig. 2 refers to the results that were obtained by the ICA. It was chosen to concentrate on the results of the principal components with the highest variance. One matrix of the four possible matrices has been used in order to generate the images.

In Fig. 2, the RGB image ($R = z_4x_{41}$, $G = z_4x_{44}$, $B = z_4x_{42} = z_4x_{43}$) for the elements of the matrices obtained from the principal component with the highest variance multiplied with the corresponding point scatter matrix is presented. It has been observed that the RGB image in Fig. 2 has differences in comparison with the original data in Fig. 1. Grassland for instance appears brighter in the RGB image of Fig. 2 and has a better contrast to neighbouring regions. For a comparison see the lower right of Figs. 1 and 2, respectively.

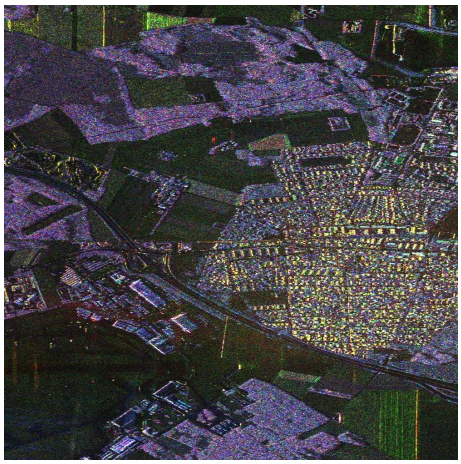


Figure 1: RGB image of the original data.

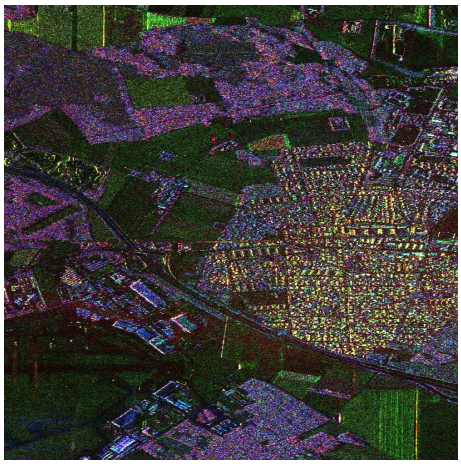


Figure 2: RGB image ($R = z_4x_{41}$, $G = z_4x_{44}$, $B = z_4x_{42} = z_4x_{43}$).

5. CONCLUSIONS

Having a series of coherent observations given by the scattering matrix $S(t)$ the principal component analysis provides a link to the incoherent method of interpretation (covariance matrix analysis) and furthermore an equivalent coherent representation containing a maximum of 4 possible uncorrelated features. The target description can be formed using z_i values and S_i point scatter matrices. The proposed technique is one of a variety of statistical and computational techniques to reveal the internal structure that underlies the set of random correlated data. Further work has to be devoted to elaborate a quantitative analysis and comparison to existing methods. Moreover a statistical analysis of the calculated quantities is necessary.

REFERENCES

1. Boerner, W.-M., H. Mott, E. Lüneburg, B. Brisco, R. Brown, J. S. Paterson (authors); S. R. Cloude, E. Krogager, J. S. Lee, D. L. Schuler, J. J. van Zyl, D. Randall, and P. Budkewitsch (contributing authors), "Polarimetry in remote sensing: Basic and applied concepts," Chapter 5 in R. A. Reyerson, ed. *The Manual of Remote Sensing*, 3rd Edition, *American Society of Photogrammetry and Remote Sensing*, ASPRS Publishing, Bethesda, MD, 1998.

2. Mott, H., "Antennas for radar and communications," *A Polarimetric Approach*, John Wiley and Sons, New York, 1992. {Please, see his updated book *Remote Sensing with Polarimetric Radar*, Published with IEEE Press and Wiley-Interscience., to be released by November 2006}.
3. Lüneburg, E., *IEICE Transactions on Electronics (Special Issue on Electromagnetic Theory)*, E78-C 10, 1339, 1995.
4. Jolliffe, I. T., *Principal Component Analysis*, 2nd edn., Springer, New York, 2002.
5. Danklmayer, A., M. Chandra, and E. Lueneburg, "Principal component analysis in radar polarimetry," URSI, Kleinheubacher Berichte 2004, *Advances in Radio Science*, Vol. 3, 2005, in press.
6. Cloude, S. R. and E. Pottier, "A review of target decomposition theorems in radar polarimetry," *IEEE Trans. GRS*, Vol. 34, No. 2, 498–518, Mar. 1996.
7. Horn, R. A. and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, New York, 1985.
8. Lüneburg, E. and W.-M. Boerner, "Statistical aspects of radar polarimetry," 43–52 in: P. Russer, M. Mongiardo (Editors), *Fields, Networks, Computational Methods, and Systems in Modern Electrodynamics* A Tribute to Leopold B. Felsen (*the Proceedings of the Symposium on Fields, Networks and Computation: A Modern View of Engineering Electrodynamics*, 2004 June 01–02, Technical University of Munich (in dedication to Leopold B. Felsen), Springer-Verlag Berlin-Heidelberg-New York, 2004.
9. Cloude, S. R., "Radar target decomposition theorems," *Electronics Letters*, Vol. 21, No. 1, 22–24, 1985.
10. Cloude, S. R., "Uniqueness of target decomposition theorems in radar polarimetry," in Boerner, W.-M. et al. (eds), Proceedings of the NATO Advances Research Workshop on *Direct And Inverse Methods in Radar Polarimetry*, Bad Windsheim, Germany, September 18–24, 1988; Kluwer Academic Publishers, Dordrecht 1992; *NATO ASI Series C*, Vol. 350, 267–296, 1988.
11. Cloude, S. R., "Polarimetric optimization based on the target covariance matrix," *Electronics Letters*, Vol. 26, No. 20, 1990.
12. Swartz, A. A., H. A. Yueh, J. A. Kong, L. M. Novak, and R. T. Shin, "Optimal polarizations for archiving maximum contrast in radar images," *J. Geophys. Research, Solid Earth and Planets*, Vol. 93, 15, 252–15, 260, 1988.
13. Cloude, S. R. and E. Pottier, "An entropy based classification scheme for land applications of polarimetric SAR," *IEEE Trans. Geosci. Remote Sensing*, Vol. 35, No. 1, 68–78, 1997.

On the Geršgorin Theorem Applied to Radar Polarimetry

E. Lüneburg[1], **A. Danklmayer**[2], **and W.-M. Boerner**[3]

[1]EML Consultants, Wessling, Germany
[2]German Aerospace Center, Germany
[3]University of Illinois at Chicago, USA

Abstract— This contribution is concerned with the mathematical formulation and theoretical background of the Geršgorin theory in the context of Radar Polarimetry. Named after its founder Semian *A. Geršgorin* the Geršgorin theorem basically states that there are certain regions in the complex plane that can be derived from any $n \times n$ complex matrix by rather simple arithmetic operations. These regions are containing more information, specifically its eigenvalues lying within or at the boundaries of circles, where the radii are obtained by the deleted absolute row and/or column sums of the respective $n \times n$ complex matrices.

1. INTRODUCTION

We consider strict radar backscattering (the monostatic case), characterized by the random Sinclair matrix $S(t)$ in the common linear $\{x, y\}$-basis

$$S(t) = \begin{bmatrix} S_{xx}(t) & S_{xy}(t) \\ S_{yx}(t) & S_{yy}(t) \end{bmatrix}, \tag{1}$$

In the case of reciprocal backscattering the Sinclair matrix is symmetric $S_{xy} = S_{yx}$ for a determinitstic or point target and $S_{xy}(t) = S_{yx}(t)$ for any instant of time or space for a reciprocal random target. A change of the orthonormal polarization basis induces a unitary consimilarity transformation for $S(t)$.

$$S(t) \to S'(t) = U^T S(t) U, \tag{2}$$

This implies that the Sinclair matrix $S(t)$ due to its symmetry can be condiagonalized for any instant of time by unitary consimilarity with the unitary matrix $U(t)$. This follows from Takagi's theorem. There is, however, a unique unitary matrix only for point targets with a delta-type probability density function. We consider the backscatter case and omit the subscript. The standard target feature vector in the general case are given by

$$\vec{k}_4(t) = \operatorname{vec} S(t) = \begin{bmatrix} S_{xx}(t) \\ S_{yx}(t) \\ S_{xy}(t) \\ S_{yy}(t) \end{bmatrix} \tag{3}$$

The corresponding covariance matrices are given by

$$C_4 = < \vec{k}(t) \vec{k}_4^{\dagger} > \tag{4}$$

The covariance matrices are Hermitian positive semidefinite and can be diagonalized by general unitary similarity transformations with a certain 4×4 unitary matrix V

$$\begin{aligned} &V^{-1} C_4 V = \Lambda_4 = \begin{bmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & \lambda_3 & 0 \\ 0 & 0 & 0 & \lambda_4 \end{bmatrix}, \\ &C_4 = V \Lambda_4 V^{-1} \quad \text{with} \quad 0 \le \lambda_4 \le \lambda_3 \le \lambda_2 \le \lambda_1. \end{aligned} \tag{5}$$

With $V = [\hat{x}_1, \hat{x}_2, \hat{x}_3, \hat{x}_4]$ we obtain the eigenvalue/eigenvector equations

$$\begin{aligned} &C_4 \hat{x}_i = \lambda_i \hat{x}_i \quad \text{with} \quad < \hat{x}_i, \hat{x}_j > = \hat{x}_i^{\dagger} \hat{x}_j = \lambda_i \delta_{ij} \\ &(i = 1, 2, 3, 4). \end{aligned} \tag{6}$$

All the eigenvectors can be multiplied by arbitrary phase factors $\hat{x}_i \to \exp(j\phi_i)\hat{x}_i$. If all four eigenvalues are different there are four one-dimensional C_4 invariant subspaces: Span $(\hat{x}_i)$, $i =$

$1, \ldots, 4$. The total number of invariant subspaces (including the zero subspace and the entire space C^4) is $2^4 = 8$. These subspaces assume a particularly simple form if the unitary similarity to the diagonal form Λ_4 is used. Then

$$\hat{x}_i = \hat{e}_i \quad \rightarrow \quad \Lambda_4 \hat{e}_i = \lambda_i \hat{e}_i \quad (i = 1, \ldots, 4). \tag{7}$$

For backscattering, the space C^4 containing the general vectors $\vec{k}_4(t)$ is restricted to the subspace C_s^4 spanned by the vectors $\vec{k}_4^{(s)}(t)$ with $S_{xy}(t) = S_{yx}(t)$. For the covariance matrix this can be expressed in the form

$$C_s^4 = PC^4 = \mathrm{Im}P \quad \text{with} \quad P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1/2 & 1/2 & 0 \\ 0 & 1/2 & 1/2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{8}$$

where P is a projector $P^2 = P$. The projector P can be expressed in the following way:

$$P = B^+ B \quad \text{with} \quad B = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad B^+ = B^T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/\sqrt{2} & 0 \\ 0 & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{9}$$

B is a 3×4 matrix and hence has no inverse in the ordinary sense, The matrix B^+ is the so-called Moore-Penrose inverse of B and is characterized as the solution of the following equations

$$BB^+B = B \quad \text{and} \quad B^+BB^+ = B^+. \tag{10}$$

Note that $BB^+ = I_3$, the 3×3 unit matrix.

The operator B is a transformation from $\mathrm{C}^4 \rightarrow \mathrm{Im}\, P$ with the properties

$$\mathrm{Im}B = \begin{cases} \mathrm{Im}P & \text{if } \vec{x} \in \mathrm{Im}\, P \\ 0 & \text{if } \vec{x} \in \mathrm{Ker}\, P \end{cases}. \tag{11}$$

From the general bi-static scattering matrix C_4 we obtain for strict backscattering the singular matrix

$$C_{4b} = \begin{bmatrix} < |S_{xx}(t)|^2 > & < S_{xx}(t)S_{xy}^*(t) > & < S_{xx}(t)S_{xy}^*(t) > & < S_{xx}(t)S_{yy}^*(t) > \\ < S_{xy}(t)S_{xx}^*(t) > & < |S_{xy}(t)|^2 > & < |S_{xy}(t)|^2 > & < S_{xy}(t)S_{yy}^*(t) > \\ < S_{xy}(t)S_{xx}^*(t) > & < |S_{xy}(t)|^2 > & < |S_{xy}(t)|^2 > & < S_{xy}(t)S_{yy}^*(t) > \\ < S_{yy}(t)S_{xx}^*(t) > & < S_{yy}(t)S_{xy}^*(t) > & < S_{yy}(t)S_{xy}^*(t) > & < |S_{yy}(t)|^2 > \end{bmatrix}. \tag{12}$$

This matrix can be decomposed as

$$C_{4,b} = \mathrm{Re}\, C_{4,b} + j\, \mathrm{Im}\, C_{4,b} \tag{13}$$

$$\mathrm{Re}C_{4b} = \begin{bmatrix} < |S_{xx}(t)|^2 > & \mathrm{Re}< S_{xx}(t)S_{xy}^*(t) > & \mathrm{Re}< S_{xx}(t)S_{xy}^*(t) > & \mathrm{Re}< S_{xx}(t)S_{yy}^*(t) > \\ \mathrm{Re}< S_{xx}(t)S_{xy}^*(t) > & < |S_{xy}(t)|^2 > & < |S_{xy}(t)|^2 > & \mathrm{Re}< S_{xy}(t)S_{yy}^*(t) > \\ \mathrm{Re}< S_{xx}(t)S_{xy}^*(t) > & < |S_{xy}(t)|^2 > & < |S_{xy}(t)|^2 > & \mathrm{Re}< S_{xy}(t)S_{yy}^*(t) > \\ \mathrm{Re}< S_{xx}(t)S_{yy}^*(t) > & \mathrm{Re}< S_{xy}(t)S_{yy}^*(t) > & \mathrm{Re}< S_{xy}(t)S_{yy}^*(t) > & < |S_{yy}(t)|^2 > \end{bmatrix} \tag{14}$$

$$\mathrm{Im}C_{4b} = \begin{bmatrix} 0 & \mathrm{Im}< S_{xx}(t)S_{xy}^*(t) > & \mathrm{Im}< S_{xx}(t)S_{xy}^*(t) > & \mathrm{Im}< S_{xx}(t)S_{yy}^*(t) > \\ -\mathrm{Im}< S_{xx}(t)S_{xy}^*(t) > & 0 & 0 & \mathrm{Im}< S_{xy}(t)S_{yy}^*(t) > \\ -\mathrm{Im}< S_{xx}(t)S_{xy}^*(t) > & 0 & 0 & \mathrm{Im}< S_{xy}(t)S_{yy}^*(t) > \\ -\mathrm{Im}< S_{xx}(t)S_{yy}^*(t) > & -\mathrm{Im}< S_{xy}(t)S_{yy}^*(t) > & -\mathrm{Im}< S_{xy}(t)S_{yy}^*(t) > & 0 \end{bmatrix} \tag{15}$$

$\mathrm{Re}C_{4,b}$ is symmetric and $\mathrm{Im}C_{4,b}$ skew-symmetric.

This matrix operator acts in the restricted space C_s^4 which is invariant with respect to the projector P. Hence we can write

$$\begin{aligned} &C_{4b} = PC_{4b}P = B^+BC_{4b}B^+B =: B^+C_3B \\ &\text{with} \quad C_3 = BC_{4b}B^+ \end{aligned} \tag{16}$$

or explicitly

$$C_3 = \begin{bmatrix} < |S_{xx}(t)|^2 > & \sqrt{2} < S_{xx}(t)S_{xy}^*(t) > & < S_{xx}(t)S_{yy}^*(t) > \\ \sqrt{2} < S_{xy}(t)S_{xx}^*(t) > & 2 < |S_{xy}(t)|^2 > & \sqrt{2} < S_{xy}(t)S_{yy}^*(t) > \\ < S_{yy}(t)S_{xx}^*(t) > & \sqrt{2} < S_{yy}(t)S_{xy}^*(t) > & < |S_{yy}(t)|^2 > \end{bmatrix}. \tag{17}$$

Being a similarity transformation the matrices C_{4b} and $\tilde{C}_{4b}$ have the same eigenvalues and the matrix $\tilde{C}_{4b}$ h is also Hermitian positive semidefinite. Deflation can be performed in any basis of the target feature vector.

The 3×3 covariance matrix C_3 can thus be generated directly from the feature vector

$$\vec{k}_3(t) = B\vec{k}_{3,b}(t) = \begin{bmatrix} S_{xx}(t) \\ \sqrt{2}\, S_o(t) \\ S_{yy}(t) \end{bmatrix} \quad \text{with} \quad S_o(t) = S_{xy}(t) = S_{yx}(t). \tag{18}$$

by the standard definition

$$C_3 =< B\vec{k}_{4,b}(t)\vec{k}_{4,b}^{\dagger}(t)B^T >=< \vec{k}_3(t)\vec{k}_3^{\dagger}(t) > . \tag{19}$$

The unitary matrix $U^T \otimes U^T$ has the form

$$W^{\dagger} = (U \otimes U)^T = U^T \otimes U^T = \begin{bmatrix} u_{11}u_{11} & u_{11}u_{21} & u_{21}u_{11} & u_{21}u_{21} \\ u_{11}u_{12} & u_{11}u_{22} & u_{21}u_{12} & u_{21}u_{22} \\ u_{12}u_{11} & u_{12}u_{21} & u_{22}u_{11} & u_{22}u_{21} \\ u_{12}u_{12} & u_{12}u_{22} & u_{22}u_{12} & u_{22}u_{22} \end{bmatrix} \tag{20}$$

and if applied to a vector

$$\vec{x} = \begin{bmatrix} a \\ b \\ b \\ c \end{bmatrix} \in C_s^4 = PC^4 = ImP \quad \Rightarrow \quad \vec{x}' = W^{\dagger}\vec{x} = U^T \otimes U^T\vec{x} = \begin{bmatrix} a' \\ b' \\ b' \\ c' \end{bmatrix} \in C_s^4, \tag{21}$$

This in particular applies to the standard target feature vector $\vec{k}_{4,b}$, i.e., the subspace C_s^4 is invariant under the unitary transformation $U^T \otimes U^T$.

In general the unitary transformations that diagonalize the covariance matrices are not of the form of a polarimetric basis transformation, i.e., in general

$$U(C_S) \neq (U(S) \otimes U(S))^T \text{ and } U(C_J) \neq (U(C_J) \otimes U^*(C_J))^T. \tag{22}$$

In the following we refer to some results contained in Horn and Johnson [1] and Varga [2]

2. GERŠGORIN THEOREM

Let $[a_{ij}] \in M_n$, and

$$R_i'(A) \equiv \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}|, \ 1 \leq i \leq n \tag{23}$$

denote the *deleted absolute row sums* of A. Then all the eigenvalues of A are located in the union of n discs

$$U_{i=1}^n \{z \in C : |z - a_{ii}| \leq R_i'(A)\} \equiv G(A). \tag{24}$$

Furthermore, if a union of k of these n forms a connected region that is disjoint from all the remaining $n - k$ discs then there are precisely k eigenvalues of A in this region.

The region $G(A)$ if often called the Geršgorin *region* (for rows) of A; the individual discs in $G(A)$ are called Geršgorin *discs*, and the boundaries of these discs are called Geršgorin *circles*. Since the matrices A and A^T have the same eigenvalues, one can obtain a Geršgorin disc theorem for columns by applying the Geršgorin disc theorem to A^T to obtain a region that contains the eigenvalues of A and is specified in terms of deleted absolute column sums

$$C_j'(A) \equiv \sum_{\substack{i=1 \\ i \neq j}}^{n} |a_{ij}|, \quad 1 \leq j \leq n. \tag{25}$$

3. CONCLUSIONS

The Geršgorin disc theorem is presented and adopted to the covariance matrices used in radar polarimetry, where the theorem shows potential to allow for target identification and classification which has to be further investigated in a follow up contribition.

REFERENCES

1. Horn, R. A. and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, New York, 1985.
2. Varga, R. S., *Geršgorin and his Circles*, Springer, 2004.
3. Lüneburg, E. and W.-M. Boerner, "Statistical aspects of radar polarimetry," 43–52 in: P. Russer, M. Mongiardo (Editors), *Fields, Networks, Computational Methods, and Systems in Modern Electrodynamics*, A Tribute to Leopold B. Felsen (*the Proceedings of the Symposium on Fields, Networks and Computation: A Modern View of Engineering Electrodynamics*, 2004 June 01–02, Technical University of Munich, Springer-Verlag Berlin-Heidelberg-New York, 2004.

Multitemporal C-Band Radar Measurement on Rice Fields

Ka-Sing Lim[1], **Chue-Poh Tan**[1], **Jun-Yi Koay**[1]
Voon-Chet Koo[1], **Hong-Tat Ewe**[1], **Yew-Chiong Lo**[1], **and Azmah Ali**[2]

[1]Multimedia University, Persiaran Multimedia, 63100 Cyberjaya, Selangor, Malaysia
[2]Malaysia Center for Remote Sensing (MACRES), Ministry of Science, Technology and Innovation
50480, Bukit Tunku Kuala Lumpur, Malaysia

Abstract— This paper investigates the relationship between C-Band backscatter measurement on the physical structure of rice fields and its growth stages. The study is based on a ground-based scatterometer experiment conducted on rice fields at Sungai Burung site in Malaysia for the year 2005 growing season. Seven C-band scatterometer acquisitions at full polarization, namely Horizontal-Horizontal, Horizontal Vertical, Vertical-Horizontal, Vertical-Vertical polarization with incidence angle from 0° to 60°, were measured. At the same time, ground truth data for an entire rice growing season were obtained at 12-day intervals from September to December 2005. The dates were chosen so as to coincide with RADARSAT-1 image acquisitions. The paper describes the experiment and investigates the radar sensitivity to the physical structures of rice at different polarization and incident angles for different rice growth stages. Based on the result, a close agreement of backscattering coefficient between the scatterometer and RADARSAT-1 was obtained.

1. INTRODUCTION

Recently, the need for the management of the food supply makes crops growth monitoring an important field in remote sensing. Among the major crop types in the world, rice constitutes the basic food for billions of people in Asia, as well as Europe and America [1]. In spite of its important as a world of food resources, studies on its microwave backscattering characteristic have been rather limited in the past years.

Three noticeable investigations on radar backscatter from rice crops can be found in literature. T. Le Toan in France conducted an experiment using a airborne X-Band Synthetic Aperture Radar (SAR), VARAN-S, Operate in HH and VV at 55°–60° [2] and T. Kurosu conducted an experiment on entire life cycles of rice (VV at 23°) using ERS-1 C-band SAR [3, 4]. Kim SooBum conducted a measurement over the whole period of rice growth in Korea using X-band ground based scatterometer [5]. In general the radar backscattering coefficient ∂° varies with respect to different plant height different microwave remote sensing system. The difference may have been due to the difference in the frequencies and the incident angles, but the difference in [3, 4] were not complete enough to make detailed comparison of the temporal variation of ∂° with [4].

Since rice is grown mainly on watered fields in Malaysia through direct sowing into the soil, its radar response will be rather homogenous. One of the motivations of this study is to examine the temporal behavior of the radar backscatter from rice crops during the whole growth period at C-Band. As a result, comparison of radar backscattering coefficient with RADARSAT-1 obtained a close agreement to each other.

The paper is outlined as follows. Section 2 describes the ground based scatterometer specification and ground truth measurement procedure. Section 3 presents the RADARSAT-1 backscattering coefficient acquisition schedules. The backscattering coefficient from Ground based scatterometer and RADARSAT-1 is used to predict and analyze rice growth stages and the discussion is presented in section 4. In Section 5, we summarize and conclude this paper.

2. PROCEDURE AND GROUND TRUTHS

A C-band scatterometer system named C-band polarimetric scatteromter (C-SCATT) was designed, with a center frequency of 6 GHz and is installed on a mobile telescopic tuck platform. This is inexpensive FM-CW radar that is efficiently constructed from a combination of commercially available components and in-house fabricated circuitry. A dual polarized parabola antenna permits the system to conduct full polarization of HH, HV, VH and VV measurement. An antenna positioning system is designed and installed to control the incident angle from 0–70°. The system has full polarimetric capability for determining the complete backscattering matrix of a natural target.

Table 1: The system specification of C-SCATT.

System Parameter	Specification
System Configuration	
Operating Frequency	6 GHz (C band)
Operating wavelength	5 cm
Sweep Bandwidth	400 MHz
Modulation Frequency	60 Hz
Polarization	HH, HV, VH, VV
Polarization Isolation	35 dB
Antenna Gain	35 dB
Antenna 3 dB Bandwidth	3°
Best Possible Range Resolution	0.375 m
Platform	Boom Truck
Platform Height	25 m (vertical)
Measurement Capability	
Transmit Power	10 dBm
Received Power	−15 dBm to −92 dBm
∂°Dynamic range	+20 dB to −40 dB
Measurement Range	20 to 100 m
Incident Angle Coverage θ	0° to 70°
Minimum signal to noise ratio SNR	10 dB
Effective range resolution	∼1.8 m at $\theta = 45^\circ$ 4.5 m at $\theta = 60^\circ$

It will be used to conduct in-situ backscatter measurements on Earth terrain, such as vegetation fields, forests, and soil surfaces [6]. Table 1 gives the system specification of C-SCATT.

In this studies, measurement were made at HH, VV, HV and VH polarizations and the incidence angles of 0° to 60°. The test site covers rice fields at Sungai Burung, Selangor, on the outskirts of Kuala Lumpur, Malaysia, with the size of approximate 50 m by 100 m for each plot of the field. The duration of whole cycle of the rice growth takes about 100 days. After sowing, the field is covered by water for land preparation for the next planting season. There are no distinct row and column effects because the farmers practice random broadcasting method. About 2 weeks later, the tillering stage starts and lasts for about one month. In this stage, the plants grow additional stems. Next phase is the reproductive stage where the panicles are formed and the plants flower. It takes about a month later for the vertical growth of the rice crop to stop and the leaves bend more horizontally. Finally, the ripening stage occurs when the moisture content of the plants decreases and the color of the plants changes from green to yellow. The supply of water is stopped about two weeks before harvesting. This growth cycle repeats starting with the land preparation of the rice fields.

3. RADARSAT-1 BACKSCATTERING COEFFICIENT ACQUISITION

The RADARSAT satellite operates at a frequency of 5.3 GHz and uses HH (radar with horizontally polarized transmission and horizontally polarized reception) polarized waves. RADARSAT works in different modes, each with a different image resolution. Fine Mode 2 is chosen for this research, and the satellite is at an angle of about 40° from the earth's surface (normally called 'inclination angle') when the images are captured. The images obtained have a resolution of approximately 7.6 m. The satellite passes by the same exact spot every 24 days, so the images were obtained on 4 separate occasions with 24 day intervals; 15 September, 09 October, 02 November 2005 and 26

November 2005, which coincided with 4 field trips.

4. EXPERIMENTAL RESULT

Figure 1 shows the angular response of ∂° on selected days of crop growth. The general angular trend shows typical volume scattering dominants behaviors. In general, rice field has dominant vertical structure, which will produce strong scattering of the VV backscatter compared to the HH backscatter in the tillering and reproductive stages. On the other hand, during the maturation period, HH backscatter will be comparable to VV. Hence, multi-polarization and multi-angular data will definitely bring additional information for rice growth monitoring throughout the season.

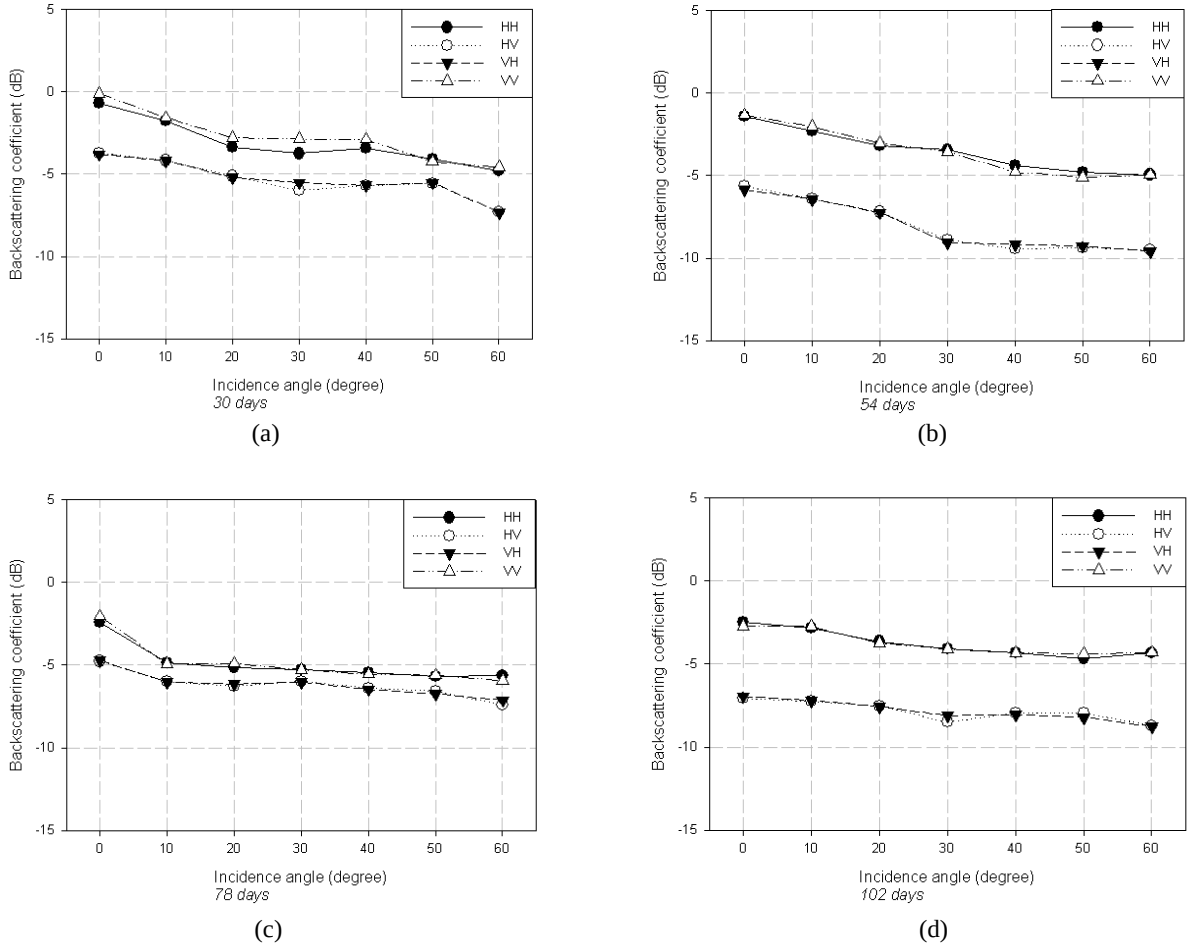


Figure 1: Angular response of ∂° on selected days: (a) 30 days, (b) 54 days, (c) 78 days, and (d) 102 days of seed broadcasting.

Figure 2 shows the relationship of rice growth stages between radar backscattering using C-SCATT and RADARSAT-1. The solid line dashed line indicates the radar backscattering coefficient ∂° from C-SCATT and RADARSAT-1 respectively. During the ground truth measurement period, three rice growth stages can be observed. It is shown that the peak value of ∂° for first planting schedule falls on the 09 October 2005, while peak value of ∂° for second planting schedule and third planting schedule reach on the 02 November 2005 and 26 November 2005, respectively. Based on the result, a close agreement of backscattering coefficient between the scatterometer and RADARSAT-1 is obtained.

5. SUMMARY AND CONCLUSION REMARKS

The radar backscattering characteristic of multi-temporal rice crop has been investigated using two different remote sensing systems, namely ground based scatterometer and RADARSAT-1. The measurement has been carried out at C band, with full polarization and at incident angle of 0°–60°. The complete monostatic angular behaviors and the multi-polarization data at C-band

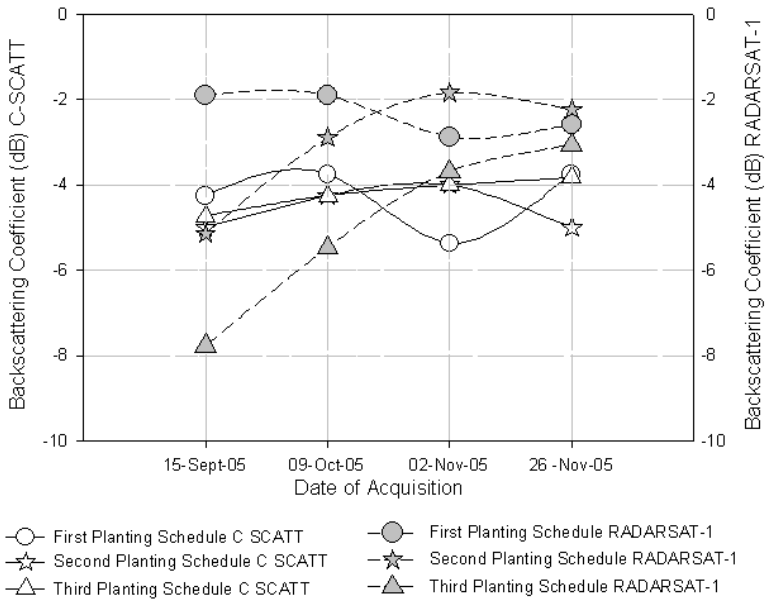


Figure 2: Comparison of rice growth stages using C-SCATT and RADARSAT-1.

throughout the growth stages of rice field are demonstrated successfully. These experiment results are expected to be helpful in better understanding the backscattering behavior of rice field, and therefore in monitoring in rice field using operational SAR systems.

ACKNOWLEDGMENT

The authors would like to acknowledge the support of MACRES for the national collaboration with Multimedia University. They would also like to thank Prof. J. A. Kong and Prof. Le Toan for providing valuable information and knowledge.

REFERENCES

1. Madsen, S. and N. Skou, "Problems to be addressed by SAR systems in the 21st century: Where are we and where are we heading?," *IEEE Geosci. Remote Sensing Soc. News.*, Mar. 1998.
2. Toan, T. L., H. Laur, E. Mougin, and A. Lopes, "Multitemporal and dual-polarization observations of agricultural vegetation covers by X-band SAR images," *IEEE Trans. Geosci. Remote Sensing*, Vol. 27, 709–717, Nov. 1989.
3. Kurosu, T., M. Fujita, and K. Chiba, "Monitoring of rice crop growth from space using the ERS-1 C-band SAR," *IEEE. Trans. Geosci. Remote Sensing*, Vol. 33, 1092–1096, July 1995.
4. Toan, T. L., F. Ribbes, L. F. Wang, N. Floury, K. H. Ding, J. A. Kong, M. Fujita, and T. Kurosu, "Rice crop mapping and monitoring using ERS-1 data based on experiment and modeling results," *IEEE Trans. Geosci. Remote Sensing*, Vol. 35, 41–56, Jan. 1997.
5. Kim, S. K., B. W. Kim, Y. K. Kong, and Y.-S. Kim, "Radar backscattering measurement of rice crop using X-band scatterometer," *IEEE Transaction on Geosci Remote Sensing*, Vol. 38, No. 3, May 2000.
6. Chet, K. V., C. B. Kuan, and C. H. Teik, "Development of a ground-based radar for scattering measurements," *IEEE Antennas and Propagation Magazine*, Vol. 45, Issue 2, 36–42, April 2003.

A Hybrid Entropy Decomposition and Support Vector Machine Method for Agricultural Crop Type Classification

Chue-Poh Tan[1], **Hong-Tat Ewe**[2], **and Hean-Teik Chuah**[1]

[1]Faculty of Engineering, Multimedia University
Persiaran Multimedia, 63100 Cyberjaya, Selangor, Malaysia
[2]Faculty of Information Technology, Multimedia University
Persiaran Multimedia, 63100 Cyberjaya, Selangor, Malaysia

Abstract— This paper presents the development of Synthetic Aperture Radar (SAR) image classifier based on the hybrid method of "Entropy Decomposition and Support Vector Machine" (EDSVM) for agricultural crop type classification. The Support Vector Machine (SVM) is successfully applied to the key parameters extracted from Entropy Decomposition to obtain good image classifications. In this paper, this novel classifier has been applied on a multi-crop region of Flevoland, Netherlands with multi-polarization data for crop type classification. Validation of the classifiers has been carried out by comparing the classified image obtained from EDSVM classifier and SVM. The EDSVM classifier demonstrates the advantages of the valuable decomposed parameters and statistical machine learning theory in performing better results compared with the SVM classifier. The final outcome of this research clearly indicates that EDSVM has the ability in improving the classification accuracy for agricultural crop type classification.

1. INTRODUCTION

Over the past few years, research in image processing for images acquired by polarimetric synthetic aperture radar (SAR) systems has raised much interest in the remote sensing community. Compared to conventional SAR systems, more detailed information on the target is expected from fully polarimetric SAR system. Using such a system, the characterization and classification of crop type may be significantly improved.

The potential of SAR in discriminating among different agricultural crop type has been demonstrated in several studies [1–3]. Previously, Van Zyl [4] proposed an unsupervised classification to classify terrain types by identifying the scattering process as odd bounce, even bounce or diffuse scattering. The ocean surface and flat ground basically have the characteristics of Bragg scattering (odd bounce) while the city blocks, buildings, and hard targets have the characteristics of double bounce scattering (even bounce) and forest and heavy vegetation have the characteristics of volume scattering (diffuse scattering). It is interesting to find that this classification algorithm provides information for terrain type identification. For a refined classification into more classes, Cloude and Pottier [5] proposed an unsupervised classification algorithm based on their target decomposition theory that utilized two parameters: the entropy, H, and alpha, $\overline{\alpha}$. The entropy, H measures the randomness of scattering mechanisms, and the angle $\overline{\alpha}$ characterizes the scattering mechanism. However, the main disadvantage of this algorithm is the arbitrary location of decision boundaries in the $H - \overline{\alpha}$ feature space [6]. To surmount this insufficiency, we propose a combined method of Entropy Decomposition and an advance machine learning method based on Support Vector Machine (EDSVM) to perform the agricultural crop type classification. This method offers an efficient classification with unique properties of SVM in providing the optimal discrimination between the agricultural crop type.

The paper is divided into five sections. Section II presents the background on the proposed technique for the agricultural crop type classification. In this section, the basic concept of the Entropy Decomposition and Support Vector Machine are discussed. Section III briefly describes the site used for the classification. The results and accuracy assessment with the comparisons between SVM and EDSVM are discussed in Section IV. In Section V, we conclude this paper with the discussion on future work.

2. PROPOSED TECHNIQUE

EDSVM is a method that combines Entropy Decomposition (ED) and Support Vector Machine (SVM). In ED, the most important parameter obtained from the output of radar systems is the

3×3 coherency matrix $[T]$. The coherency matrix $[T]$ is obtained from an ensemble of scattering matrix samples $[S_i]$ by forming the Pauli scattering vectors [7].

$$[S_i] = \begin{bmatrix} S_{HH} & S_{HV} \\ S_{VH} & S_{VV} \end{bmatrix} \rightarrow k = \frac{1}{\sqrt{2}}[S_{HH} + S_{VV} \quad S_{HH} - S_{VV} \quad 2S_{HV}]. \tag{1}$$

Averaging the outer product of them over the given samples yields

$$[T] = \langle k\, k^{+} \rangle = \begin{bmatrix} \langle |S_{HH}+S_{VV}|^2 \rangle & \langle (S_{HH}+S_{VV})(S_{HH}-S_{VV})^* \rangle & 2\langle (S_{HH}+S_{VV})S_{HV}^* \rangle \\ \langle (S_{HH}-S_{VV})(S_{HH}+S_{VV})^* \rangle & \langle |S_{HH}-S_{VV}|^2 \rangle & 2\langle (S_{HH}-S_{VV})S_{HV}^* \rangle \\ 2\langle (S_{HH}+S_{VV})^* S_{HV} \rangle & 2\langle (S_{HH}-S_{VV})^* S_{HV} \rangle & 4\langle |S_{HV}|^2 \rangle \end{bmatrix} \tag{2}$$

where k^{+} refers to the conjugate transpose of k and $(S_i)^*$ is the conjugate element of (S_i). The coherency matrix $[T]$ is Hermitian positive semi-definite, therefore it can always be digitalized by a unitary similarity transformation of the form

$$[T] = [U][\Lambda][U]^{-1} \tag{3}$$

where $[U]^{-1}$ represents the inverse matrix of $[U]$ and

$$[\Lambda] = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}, \tag{4}$$

$$[U] = [u_1 \quad u_2 \quad u_3], \tag{5}$$

$$[u_i] = [\cos\alpha_i \quad \cos\alpha_i \sin\beta_i e^{j\delta_i} \quad \sin\alpha_i \sin\beta_i e^{j\gamma_i}]. \tag{6}$$

The angle $\overline{\alpha}$ characterizes the scattering mechanisms as it indicates surface scattering when $\overline{\alpha} = 0^\circ$, volume scattering when $\overline{\alpha} = 45^\circ$ and multiple scattering when $\overline{\alpha} = 90^\circ$, respectively [8]. In order to obtain the averaged scattering mechanisms from surface scattering to multiple scattering, the average alpha is defined as

$$\overline{\alpha} = P_1\alpha_1 + P_2\alpha_2 + P_3\alpha_3, \tag{7}$$

$$P_i = \frac{\lambda_i}{\sum_{k=1}^{3} \lambda_k} \tag{8}$$

where P_i are the probabilities obtained from the eigen values of $[T]$. The angle β is twice the polarization orientation angle. The angle δ is the phase difference between decomposed $(S_{HH}+S_{VV})$ and $(S_{HH}-S_{VV})$ terms and the angle γ is the phase difference between the decomposed $(S_{HH}+S_{VV})$ and S_{HV} terms [9]. To introduce the degree of statistical disorder of each target, the entropy H is defined in the Von Neumann sense from the logarithmic sum of eigen values of $[T]$ as

$$H = -\sum_{i=1}^{3} P_i \log_3(P_i). \tag{9}$$

However, the main disadvantage of the entropy decomposition is that the location of the decision boundary is arbitrary [6]. Thus, to overcome this problem, we propose a combined method of Entropy Decomposition and an advance machine learning method based on Support Vector Machine (EDSVM).

The SVM algorithm is a machine learning technique based on statistical theory [10] that can be used for classification purposes. The aim of Support Vector Machine classifier is to find an ideal separating hyperplane in a higher dimensional feature space. For a given training sample belonging to two different classes, SVM derives a hyperplane, which is at a maximum distance from the closest points belonging to both the classes. To find the optimal separating hyperplane, assume that the two classes to be distinguished are linearly separable, and denote the input spaceXwith input vectors, $\vec{x}$ and the training set $T_r = \{(x_1, y_1), \ldots, (x_N, y_N)\}$, where $x_i \in X$ and $y_i \in Y$, $Y = \{1, -1\}$. In practice, it will often be the case where the data cannot be separated linearly

by means of a hyperplane. One of the basic ideas behind SVM is to have a mapping Φ from the original input space X into a high-dimensional feature space F.

The SVM method solves for

$$\min ||w||^2 \tag{10}$$

with

$$y_i(\langle \Phi(\vec{x}_i), w\rangle + b) \geq 1 \quad \text{for} \quad i = 1, \ldots, N \tag{11}$$

where $\vec{w}$ is a vector perpendicular to the hyperplane while b determines the displacement of the hyperplane along the normal vector $\vec{w}$ [11]. To solve the constrained minimization problem, the Lagrangian dual problem method is introduced as

$$\text{maximize} \quad W(l) = \sum_{i=1}^{N} l_i - \frac{1}{2_i} \sum_{i,j=1}^{N} l_i l_j y_i y_j \langle \Phi(\vec{x}_i), \Phi(\vec{x}_j)\rangle \tag{12}$$

$$\text{subject to} \quad l_i \geq 0, \;\; i = 1, \ldots N, \quad \text{and} \quad \sum_{i=1}^{m} l_i y_i = 0 \tag{13}$$

with Langrange multipliers $l_i \geq 0$. After solving this dual problem, the decision function implemented by the classifier for any test vectors x is expressed by

$$f(x) = \text{sgn}\left(\sum_{i=1}^{N} l_i y_i \langle \Phi(\vec{x}), \Phi\vec{x}_i)\rangle + b\right). \tag{14}$$

3. SITE DESCRIPTION

The JPL L band polarimetric SAR image of Flevoland (Netherlands) is used for the crop classification purpose. The sample image was downloaded from ESA's Earth Observation and it was originally processed with 4-look average in Stokes matrix with pixel size of 12 m. The L-band multipolarized image for the study area was taken on 30 May 1990 with 1024 samples and 750 lines. Figure 1 blue. The topography is almost perfectly flat, and the general altitude is three meters below sea level. This site is composed of various crop plantations, bare soil, forest and waterbody. Eight crop classes are identified, comprising stem bean, potato, lucerne, wheat, beet, rape seed, peas and grass.

Figure 1: Flevoland study area HH assigned in red, HV assigned in green and VV assigned in blue.

4. RESULTS AND DISCUSSION

Using multi-polarization polarimetric AIRSAR data of Flevoland, the overall accuracy of the aforementioned classification techniques for selected testing areas as presented in [12] with different window size are shown in Figure 2. The optimum window size is 9 and the accuracy of 76.28% for SVM classifier and 98.39% for the proposed EDSVM classifier, respectively.

From these results, the classified images with the optimum window size for SVM and EDSVM are shown in Figure 3(a) and Figure 3(b), respectively.

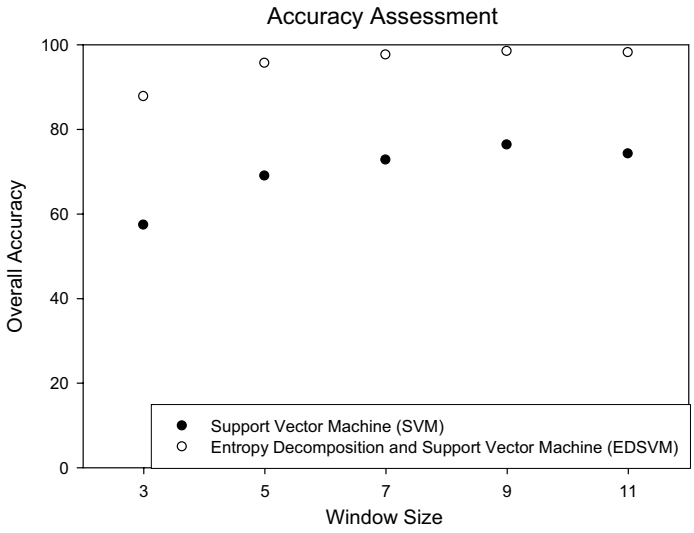


Figure 2: Accuracy Assessment for the study area of Flevoland.

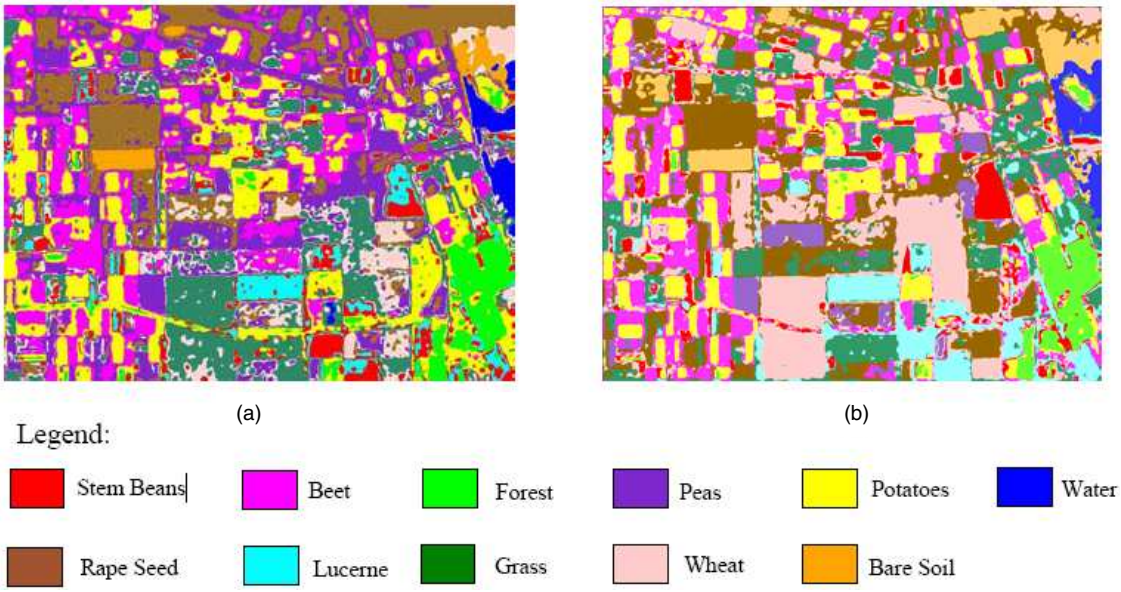


Figure 3: Classification Techniques (a) SVM (b) EDSVM.

5. CONCLUSION

The project is aimed to investigate the hybrid method of EDSVM in the classification of agricultural crop type using the multi-polarized AIRSAR image. The potential of the proposed classifier has been demonstrated and it has shown a desirable result compared to SVM classifier. Our future work shall extend this to multi-frequency and multi-temporal data.

REFERENCES

1. Ulaby, F. T., R. K. Moore, and A. K. Fung, *Microwave Remote Sensing:Active and Passive*, Norwood, MA, Artech House, 1986.
2. Bouman, B. A. M. and D. H. Hoekman, "Multi-temporal multi-frequency radar measurements of agricultural crops during the agriscatt-88 campaign in the netherlands," *Int. J. Remote Sens.*, Vol. 14, 1595–1614, 1993.
3. T. N. Tran, R. Wehrens, D. H. Hoekman, and L. M. C. Buydens, "Initialization of markov random field clustering of large remote sensing images," *IEEE Trans. Geosci. RemoteSensing*, Vol. 43, No. 8, 1912–1919, 2005.

4. Van Zyl, J. J., "Unsupervised classification of scattering behavior using radar polarimetry data," *IEEE Trans. Geosci. Remote Sensing*, Vol. 27, No. 1, 36–45, Jan. 1989.
5. Cloude, S. R. and E. Pottier, "An entropy based classification scheme for land applications of polarimetric SAR," *IEEE Trans. Geosci. Remote Sensing*, Vol. 35, No.1, 68–78, Jan. 1997.
6. Yahia, M. and Z. Belhadj "Unsupervised classification of polarimetric sar images using neural networks," *IEEE International Conference Information and Communication Technologies 2004*, 335–337, 2004.
7. Lee, J. S., M. R. Grunes, and R. Kwok, "Classification of multi-look polarimetric SAR imagery based on complex wishart distribution," *Int. J. Remote Sensing*, Vol. 15, No. 11, 2299–2311, 1994.
8. Cloude, S. R. and E. Pottier, "A review of target decomposition theorems in radar polarimetry," *IEEE Trans. Geoscience. Remote Sensing*, Vol. 34, No. 2, 498–518, 1996.
9. Lee, J. S., M. R. Grunes, and T. L. Ainsworth, "Unsupervised classification using polarimetric decomposition and the complex wishart distribution," *IEEE Transactions Geoscience. And Remote Sensing*, Vol. 37, No. 5, 2249–2259, 1999.
10. Vapnik, V., "The nature of statistical learning theory," *Springer Verlag*, 1999.
11. Schökopf, B. and A. J. Smola, *Learning With Kernels*, The MIT Press, Cambridge, Massachusetts, London, England., 2002.
12. Pottier, E., J. S. Lee, and L. Ferro-Famil, "Advanced concepts in polarimetric sar image analysis: a tutorial review," EuSAR 2006 Tutorial, EuSAR 2006 Conference, Dresden, Germany, May 2006.

Iteration Based Polarimetric SAR Image Classification

Jian Yang, Xiaoli She, and Tao Xiong
Department of Electronic Engineering, Tsinghua University, Beijing 100084, China

Abstract— In this paper, an iteration method is proposed for supervised polarimetric synthetic aperture radar (SAR) image classification. In this iterative approach, the optimization of polarimetric contrast enhancement (OPCE) is employed for enlarging the distance between the mean values of two kinds of targets and the Fisher method is employed for reducing the variances of two distributions. Using the proposed approach, polarimetric SAR image can be classified only after a few iterations. For comparison, the authors also use the maximum likelihood (ML) classifier for classification, based on the complex Wishart distribution. The classification results of a NASA/JPL AIRSAR L-band image over San Francisco demonstrate the effectiveness of the proposed approach.

1. INTRODUCTION

Classification of Earth terrain is one important application of polarimetric synthetic aperture radar (SAR) remote sensing. Many methods have been employed for classification of polarimetric SAR data, based on the maximum likelihood (ML) [1], artificial neural network (NN) [2, 3], support vector machines (SVMs) [4], or other approaches [5–15, 19]. Among these methods, the ML classifier [1] can be employed for obtaining accurate classification results, but it is based on the assumption of the complex Wishart distribution of the covariance matrix. If the covariance matrix of the polarimetric SAR data does not conform to the Wishart distribution, one may obtain unsatisfactory classification results. It is hence important to propose a new approach to classification which is independent of any assumption of a statistical distribution. More recently, the authors proposed a method for polarimetric SAR image classification [19], based on the generalized optimization of polarimetric contrast enhancement (GOPCE) [17]. However, only the distance between the function mean values of two kinds of targets were considered in the method [19], the variances of the function distributions of two kinds of targets were not considered.

In this paper, an iteration method is proposed for supervised polarimetric SAR image classification, based on the optimization of polarimetric contrast enhancement (OPCE) and the Fisher method. In this iterative approach, the OPCE is employed for enlarging the distance between the mean values of two kinds of targets and the Fisher method is employed for reducing the variances of two distributions. Using the proposed approach, polarimetric SAR image can be classified only after a few iterations. For comparison, the maximum likelihood classification is also used. The classification results of a NASA/JPL AIRSAR L-band image over San Francisco demonstrate the effectiveness of the proposed approach.

2. OPCE AND FISHER'S METHOD

2.1. OPCE

Let TA and TB denote two kinds of targets, and let $\left[\bar{K}\left(TA\right)\right]$ and $\left[\bar{K}\left(TB\right)\right]$ be the average Kennaugh matrices of TA and TB, respectively. For the OPCE, we need to find the optimal polarization states $\mathbf{g}$ and $\mathbf{h}$ such that the power ratio of the two kinds of targets is maximal:

$$\begin{aligned} \text{maximize} \quad & \frac{\mathbf{h}^t \left[\bar{K}\left(TA\right)\right]\mathbf{g}}{\mathbf{h}^t \left[\bar{K}\left(TB\right)\right]\mathbf{g}}, \\ \text{subject to :} \quad & g_1^2+g_2^2+g_3^2=1 \\ & h_1^2+h_2^2+h_3^2=1. \end{aligned} \tag{1}$$

Using an iteration method [18], the optimal polarization states $\mathbf{g}_m$ and $\mathbf{h}_m$ are easily obtained.

2.2. Fisher's Method

For a kind of targets, any parameter, such as the received power, the polarization entropy, has a distribution.The OPCE can be used to enlarge the distance between the mean values of two kinds of targets. To classify two kinds of targets, however, we not only need to enlarge the distance between two mean values, but also need to reduce the two variances of both kinds of targets. On

the other hand, for different targets, the entropy and the similarity parameters should be different. For this reason, we consider the following expression:

$$FP = \left[\sum_{i=1}^{3} x_i r_i\right] \times \mathbf{h}_m^t [K] \mathbf{g}_m, \tag{2}$$

where r_1 and r_2 are similarity parameters defined in [16], respectively. $r_3 = H$ denotes the polarization entropy. Now we apply the Fisher method for obtaining the unknown coefficients x_i by the following optimization:

$$\text{minimize } \frac{\text{Var}(FP(KA)) + \text{Var}(FP(KB))}{[\text{mean}(FP(KA)) - \text{mean}(FP(KB))]^2}, \tag{3}$$

where FP is defined by (2). Please note that (2) is linear for variables x_i. (3) is hence a Rayleigh ratio and x_i can easily be obtained by solving an eigenvalue equation.

3. PROPOSED APPROACH

Based on the OPCE and the Fisher method, we propose a supervised classification method. The basic idea of the method is to divide the mixed targets of two classes into three parts: two parts have been classified clearly, belonging to two different classes; and the other part is still mixed, it will also be re-divided into three parts by the OPCE and the Fisher method in the next iteration. The proposed classification scheme consists of the following steps:

Step 1 Obtaining some results from the training sites

(1.1) After selecting the training sites of a polarimetric SAR image, denote all the classes as C_1, C_2, ..., C_k by the order of the average span of every class, where the average span of C_j is less than that of C_{j+1}.

(1.2) For two neighbor classes (denoted as C_j and C_{j+1}) of training sites, remove all the pixels which their spans are larger than the average span of C_{j+1} or less than the average span of C_j. Denoting the remainders as C_j^1 and C_{j+1}^1, and letting the remainders be the TA and TB in $(2, 3)$, respectively, then obtain the optimal coefficients and optimal polarizations, denoted as $\mathbf{x}_m^1(j)$, $\mathbf{g}_m^1(j)$ and $\mathbf{h}_m^1(j)$.

(1.3) Calculating FP by Eq. (2), remove all the pixels for which their FPs are larger than the average FP of C_{j+1}^1 or less than the average FP of C_j^1. Denoting the remainders as C_j^2 and C_{j+1}^2 and letting the remainders be the TA and TB in $(1, 3)$, respectively, then obtain the optimal coefficients and optimal polarizations, denoted as $\mathbf{x}_m^2(j)$, $\mathbf{g}_m^2(j)$ and $\mathbf{h}_m^2(j)$, respectively.

(1.4) Repeat Step 1.3 until the number of the final remainders is less than one percent of the total number of C_j and C_{j+1}. Denote the optimal coefficients and optimal polarizations of the final OPCE as $\mathbf{x}_m^n(j)$, $\mathbf{g}_m^n(j)$ and $\mathbf{h}_m^n(j)$, respectively. In addition, denote the final remainders as C_j^{n+1} and C_{j+1}^{n+1}.

Step 2 Classification of a polarimetric SAR image

(2.1) An arbitrary pixel of a polarimetric SAR image is divided into one of three classes, i.e., C_1 (if the span is less than the average span of the first training site), C_k (if the span is larger than the average span of the k-th training site), and a mixed class of C_j and C_{j+1}.

(2.2) If a pixel is classified to the mixed class of C_j and C_{j+1}, then calculate FP by Eq. (2) by using $\mathbf{x}_m^1(j)$, $\mathbf{g}_m^1(j)$ and $\mathbf{h}_m^1(j)$. If the corresponding FP is larger than the average FP of C_{j+1}^1 (defined by Step 1.2), than the pixel is classified to C_{j+1}; if the corresponding FP is less than the average FP of C_j^1 (defined by Step 1.2), then the pixel is classified to C_j. If the pixel is not classified to C_j, neither it is to C_{j+1}, then repeat the above processing by using $\mathbf{x}_m^i(j)$, $\mathbf{g}_m^i(j)$ and $\mathbf{h}_m^i(j)$, $i = 2, 3, \ldots, n$. If the pixel is not classified to C_j or C_{j+1} after n times, then divide the pixel to C_j or C_{j+1} by using the distance between the corresponding FP associated with $\mathbf{x}_m^n(j)$, $\mathbf{g}_m^n(j)$ and $\mathbf{h}_m^n(j)$ and the average FP of C_j^{n+1} or C_{j+1}^{n+1}.

A Note: For some special case, if the span distribution of one kind of targets has a very large variance, it may be necessary to classify not only two neighbor kinds of targets, but also arbitrary two kinds of targets. Then we decide what class a pixel belongs to according to weighted classification results.

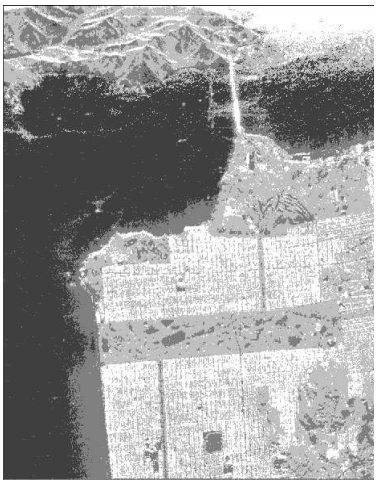


Figure 1. Span image of San Francisco. **Figure 2.** Classification results by ML. **Figure 3.** Classification results by the proposed method.

Table 1: Classification results by ML and the proposed method.

(a)

Proposed Method	Sea area	Quasi-natural surface	Woods area	Urban area
Sea area	**0.997**	0.003	0	0
Quasi-natural surface	0.001	**0.987**	0.011	0
Woods area	0	0.017	**0.974**	0.008
Urban area	0	0.007	0.084	**0.909**

(b)

ML	Sea area	Quasi-natural surface	Woods area	Urban area
Sea area	**0.944**	0.056	0	0
Quasi-natural surface	0.001	**0.964**	0.034	0.001
Woods area	0	0.002	**0.994**	0.004
Urban area	0	0.001	0.160	**0.839**

4. CLASSIFICATION RESULTS AND CONCLUSION

The NASA/JPL AIRSAR L-band data of San Francisco is used to illustrate the effectiveness of the proposed approach to classification. Fig. 1 shows the span image. Fig. 2 illustrates the classification results by using the proposed approach. Table 1(a) shows the classification accuracy by the proposed method. For comparison, we also use the maximum likelihood classifier based on the complex Wishart distribution for the polarimetric covariance matrix, the corresponding classification results and classification accuracy are shown in Fig. 3 and Table 1(b), respectively. From Table 1(a) and (b), we conclude that the classification accuracy (on average) by the proposed method is a little better than that by the maximum likelihood classifier. Please note that any assumption on probability distribution is not necessary in the proposed approach, whereas the ML is based on the assumption of the complex Wishart distribution for the polarimetric covariance matrix.

ACKNOWLEDGMENT

This work was supported in part by the National Natural Science Foundation of China (40571099), by the Research Fund for the Doctoral Program of Higher Education of China, and by the Fundamental Research Foundation of Tsinghua University.

REFERENCES

1. Lee, J.-S., M. R. Grunes, and R. Kwok, "Classification of multi-look polarimetric SAR imagery based on complex Wishart distribution," *Int. J. Remote Sensing*, Vol. 15, No. 11, 2299–2311, 1994.

2. Chen, K. S., W. P. Huang, et al., "Classification of multifrequency polarimtric SAR imagery using a dynamic learning neural network," *IEEE Trans. Geosci. Remote Sensing*, Vol. 34, No. 3, 814–820, 1996.
3. Ito, Y. and S. Omatu, "Polarimetric SAR data classification using competitive neural networks," *International Journal of Remote Sensing*, Vol. 19, No. 14, 2665–2684, 1998.
4. Fukuda, S., R. Kataqiri, and H. Hirosawa, "Unsupervised approach for polarimetric SAR image classification using support vector machines," *International Geoscience and Remote Sensing Symposium (IGARSS)*, Vol. 5, 2599–2601, 2002.
5. Chen, C.-T., K.-S. Chen, and J.-S. Lee, "The use of fully polarimetric information for the fuzzy neural classification of SAR images," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 41, No. 9, PART I, 2089–2100, 2003.
6. Du, L. and J.-S. Lee, "Fuzzy classification of earth terrain covers using complex polarimetric SAR data," *International Journal of Remote Sensing*, Vol. 17, No. 4, 809, 1996.
7. Van Zyl, J. J., "Unsupervised classification of scattering behavior using radar polarimetry data," *IEEE Trans. Geosci. Remote Sensing*, Vol. 27, No. 1, 36–45, Jan. 1989.
8. Kong, J. A., A. A. Swartz, H. A. Yueh, et al., "Identification of terrain cover using the optimal terrain classifier," *J. Electronmagn. Waves Applicat.*, Vol. 2, 171–194, 1988.
9. Yueh, H. A., A. A. Swartz, J. A. Kong, et al., "Optimal classification of terrain cover using normalized polarimetric data," *J. Geophys. Res.*, 15261–15267.
10. Van Zyl, J. J. and C. F. Burnette, "Bayesian classification of polarimetric SAR images using adaptive a priori probabilities," *Int. J. Remote Sensing*, Vol. 13, No. 5, 835–840, 1992.
11. Cloude, S. R. and E. Pottier, "An entropy based classification scheme for land applications of polarimetric SAR," *IEEE Trans. Geosci. Remote Sensing*, Vol. 35, No. 1, 68–78, Jan. 1997.
12. Lee, J.-S., M. R. Grunes, et al., "Unsupervised classification using polarimetric decomposition and the complex Wishart classifier," *IEEE Trans. Geosci. Remote Sensing*, Vol. 37, No. 5, 2249–2257, Sep. 1999.
13. Lee, J.-S. and K. Hoppel, "Principal component transformation of multifrequency polarimetric SAR imagery," *IEEE Trans. Geosci. Remote Sensing*, Vol. 30, No. 4, 686–696, July 1992.
14. Lee, J.-S., M. R. Grunes, and G. F. de Grandi, "Polarimetric SAR speckle filtering and its implication for classification," *IEEE Trans. Geosci. Remote Sensing*, Vol. 37, No. 5, 2363–2373, Sep. 1999.
15. Alberqa, V., "Comparison of polarimetric methods in image classification and SAR interferometry applications," Ph.D. Thesis, DLR, Deutschen Forschungsanstalt fur Luft-und Raumfahrt, Germany, 2004.
16. Yang, J., Y. N. Peng, and S. M. Lin, "Similarity between two scattering matrices," *Electron. Letters*, Vol. 37, No. 3, 193–194, Feb. 2001.
17. Yang, J., G. W. Dong, Y. N. Peng, Y. Yamaguchi, and Y. Yamada, "Generalized optimization of polarimetric contrast enhancement," *IEEE Geosci. Remote Sensing Lett.*, Vol. 1, No. 3, 171–174, 2004.
18. Yang, J. Y. Yamaguchi, W.-M. Boerner, and S. M. Lin, "Numerical methods for solving the optimal problem of contrast," *IEEE Trans. Geosci. Remote Sensing*, Vol. 38, No. 2, 965–971, 2000.
19. Yang, J., T. Xiong, Y. Peng, et al., "Polarimetric SAR image classification by using generalized optimization of polarimetric contrast enhancement," *International Journal of Remote Sensing*, Vol. 27, No. 16, 3413–3424, 2006.

Coherence Enhancement for Polarimetric SAR Interferometry

Tao Xiong[1], **Jian Yang**[1], **Weijie Zhang**[1], **Yoshio Yamaguchi**[2], **and Hiroyoshi Yamada**[2]

[1]Department of Electronic Engineering, Tsinghua University, Beijing, China
[2]Department of Information Engineering, Niigata University, Niigata, Japan

Abstract— A new method for coherence enhancement is proposed based on the mathematical model on the relationship between the amplitude and the phase of the complex signals. An optimal projection direction of scattering vectors is derived for maximizing the amplitudes of the master and slave complex images. The experimental results demonstrate the effectiveness of the proposed approach.

1. INTRODUCTION

Since each scattering cell of the polarimetric interferometric SAR (PolInSAR) data has two scattering matrixes or scattering vectors corresponding to the master and the slave antennas, coherence enhancement between the master and slave images becomes possible. In recent years, several algorithms have been proposed, such as the optimal vector interferometry in [1], the coherent decompositions method in [2] and so on. In this paper, a quite different method is proposed, based on the optimization of the amplitudes of the complex signals. For each pixel, an optimal projection direction is obtained for the master and slave scattering vectors. The coherence between two projections is enhanced remarkable and the interferometric phase is improved obviously.

2. THE OPTIMAL PROJECTION DIRECTION MODEL

In the single channel interformetry, e.g., HH, the master and slave images are denoted as M and S, respectively. The coherence between them is defined as:

$$q = \left|\left\langle S_{\mathrm{M}}^{*} S_{\mathrm{S}}\right\rangle\right| / \sqrt{\left\langle S_{\mathrm{M}}^{*} S_{\mathrm{M}}\right\rangle \left\langle S_{\mathrm{S}}^{*} S_{\mathrm{S}}\right\rangle} \tag{1}$$

where $\langle . \rangle$ represents the neighborhood average, S_{M} and S_{S} are signals in M and S, $*$ denotes complex conjugate. In the complex images, the amplitudes of the signals vary with the terrain. For some pixels, the amplitudes of the received signals are very weak. In this case, the phases are unreliable. For example, if there are two signals $S_1 = 0.00001$ and $S_2 = -0.00001$, both of them are very close to zero, i.e., $|S_1| \approx |S_2| \approx 0$. However, the difference between the phases is π. In addition, weak signals imply in general that the corresponding SNR is low. So the phases of very weak signals are not reliable.

In SAR interferometry, the interferometric phase is defined as the difference between the master and slave phases. If any one of two phases is unreliable, the interferometric phase is unreliable. Therefore, we hope that both the master and slave amplitudes would be as large as possible.

For polarimetric SAR interferometry the master and slave scattering vectors of each pixel are denoted as:

$$\boldsymbol{k}_i = \left[S_{i,\,\mathrm{HH}} + S_{i,\,\mathrm{VV}},\ S_{i,\,\mathrm{HH}} - S_{i,\,\mathrm{VV}},\ 2S_{i,\,\mathrm{HV}}\right]^{\mathrm{T}} / \sqrt{2},\ i = \mathrm{M},\ \mathrm{S} \tag{2}$$

where T means transpose.

In order to enhance the reliability (coherence) of the interferometric phase, a projection direction (a unitary vector $\boldsymbol{w}$) should be derived. We define the projections as $\mu_{\mathrm{M}} = \boldsymbol{w}^{\mathrm{H}}\boldsymbol{k}_{\mathrm{M}}$ and $\mu_{\mathrm{S}} = \boldsymbol{w}^{\mathrm{H}}\boldsymbol{k}_{\mathrm{S}}$, then we hope that both the amplitudes of μ_{M} and μ_{S} are as large as possible synchronously. So the optimal projection direction model is given as follows:

$$\begin{aligned} &\max_{w} \left(\min\left(\left|\boldsymbol{w}^{\mathrm{H}}\boldsymbol{k}_{\mathrm{M}}\right|,\ \left|\boldsymbol{w}^{\mathrm{H}}\boldsymbol{k}_{\mathrm{S}}\right|\right)\right) \\ &\text{s.t.} \qquad \|\boldsymbol{w}\| = 1 \end{aligned} \tag{3}$$

where H means complex conjugate transpose.

3. THE SOLUTION OF THE MODEL

In order to obtain the analytic solution of the above problem, we transform it into an equivalent problem as follows:

$$\begin{aligned} & \max_{w}(a,\, b) \\ \text{s.t.}a &= \max_{w} |\boldsymbol{w}^{\mathrm{H}}\boldsymbol{k}_{\mathrm{M}}|^2,\, |\boldsymbol{w}^{\mathrm{H}}\boldsymbol{k}_{\mathrm{M}}|^2 \le |\boldsymbol{w}^{\mathrm{H}}\boldsymbol{k}_{\mathrm{S}}|^2 \\ b &= \max_{w} |\boldsymbol{w}^{\mathrm{H}}\boldsymbol{k}_{\mathrm{S}}|^2,\, |\boldsymbol{w}^{\mathrm{H}}\boldsymbol{k}_{\mathrm{M}}|^2 \ge |\boldsymbol{w}^{\mathrm{H}}\boldsymbol{k}_{\mathrm{S}}|^2 \\ \|\boldsymbol{w}\| &= 1 \end{aligned} \tag{4}$$

From the quadratic programming theory, the solution has two following situations.

In the first situation, the non-restricted maximum of $|\boldsymbol{w}^{\mathrm{H}}\boldsymbol{k}_{\mathrm{M}}|^2$ and $|\boldsymbol{w}^{\mathrm{H}}\boldsymbol{k}_{\mathrm{S}}|^2$ are both not located in the restricted range. At this time, a and b must be located on the boundary. So when $\boldsymbol{w}$ is the optimal projection direction, the both projections must have the same amplitudes:

$$|\boldsymbol{w}_1^{\mathrm{H}}\boldsymbol{k}_{\mathrm{M}}| = |\boldsymbol{w}_1^{\mathrm{H}}\boldsymbol{k}_{\mathrm{S}}| \tag{5}$$

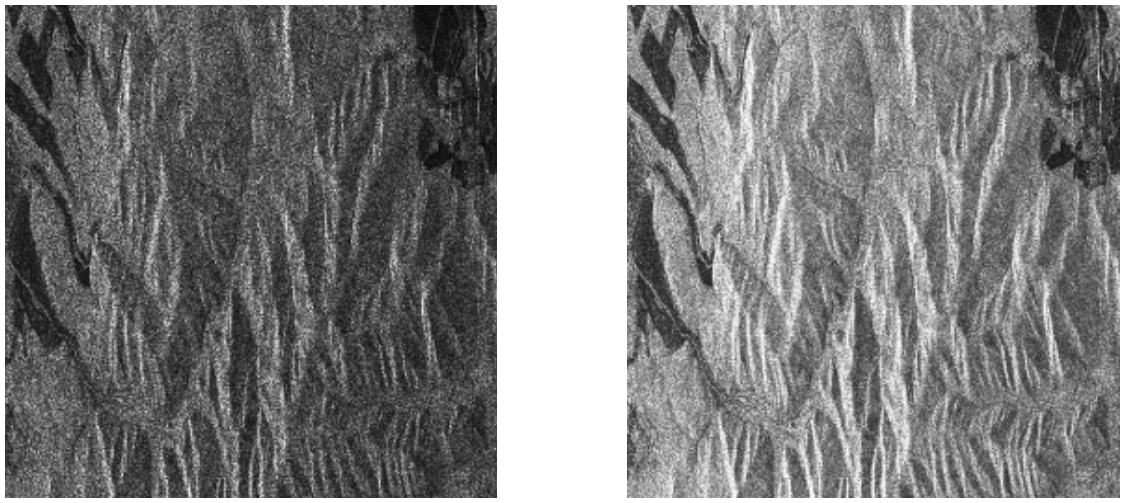
In this situation, an eigen-decomposition method can be used to obtain the analytic solution. From (5), it can be derived that $\boldsymbol{w}^{\mathrm{H}}(\boldsymbol{k}_{\mathrm{M}}\boldsymbol{k}_{\mathrm{M}}^{\mathrm{H}} - \boldsymbol{k}_{\mathrm{S}}\boldsymbol{k}_{\mathrm{S}}^{\mathrm{H}})\boldsymbol{w} = 0$. If $\boldsymbol{k}_{\mathrm{M}} = c\boldsymbol{k}_{\mathrm{S}}$ (c is an arbitrary complex number except zero), it can be contained in the second situation. If $\boldsymbol{k}_{\mathrm{M}} \neq c\boldsymbol{k}_{\mathrm{S}}$, the matrix $\boldsymbol{k}_{\mathrm{M}}\boldsymbol{k}_{\mathrm{M}}^{\mathrm{H}} - \boldsymbol{k}_{\mathrm{S}}\boldsymbol{k}_{\mathrm{S}}^{\mathrm{H}}$ has three eigenvalues $\lambda_1 > 0$, $\lambda_2 < 0$, $\lambda_3 = 0$ and the corresponding eigenvecters are $\boldsymbol{v}_1$, $\boldsymbol{v}_2$, $\boldsymbol{v}_3$, respectively. We finally obtain:

$$\boldsymbol{w}_1 = \frac{\alpha\boldsymbol{v}_1 + \boldsymbol{v}_2}{\|\alpha\boldsymbol{v}_1 + \boldsymbol{v}_2\|},\ \alpha = \sqrt{-\frac{\lambda_2}{\lambda_1}}\frac{\|\boldsymbol{v}_2\|_2^2}{\|\boldsymbol{v}_1\|_2^2}\exp\left\{-j\arg\left(\boldsymbol{v}_1^{\mathrm{H}}\boldsymbol{k}_{\mathrm{M}}\boldsymbol{k}_{\mathrm{M}}^{\mathrm{H}}\boldsymbol{v}_2\right)\right\} \tag{6}$$

In the second situation, the non-restricted maximum of $|\boldsymbol{w}^{\mathrm{H}}\boldsymbol{k}_{\mathrm{M}}|^2$ or $|\boldsymbol{w}^{\mathrm{H}}\boldsymbol{k}_{\mathrm{S}}|^2$ is located in the restricted range. So the optimal $\boldsymbol{w}$ has the same direction as $\boldsymbol{k}_{\mathrm{M}}$ or $\boldsymbol{k}_{\mathrm{S}}$ which has a less amplitude:

$$\boldsymbol{w}_2 = \begin{cases} \boldsymbol{k}_{\mathrm{M}}/\|\boldsymbol{k}_{\mathrm{M}}\|, & \text{if} \quad \|\boldsymbol{k}_{\mathrm{M}}\| \le \|\boldsymbol{k}_{\mathrm{S}}\| \\ \boldsymbol{k}_{\mathrm{S}}/\|\boldsymbol{k}_{\mathrm{S}}\|, & \text{if} \quad \|\boldsymbol{k}_{\mathrm{M}}\| > \|\boldsymbol{k}_{\mathrm{S}}\| \end{cases} \tag{7}$$

For each pixel, both the situations are considered. Comparing the values of the objective function in different situations, we can make a decision.



(a) HH channel (b) Polarimetric fusion

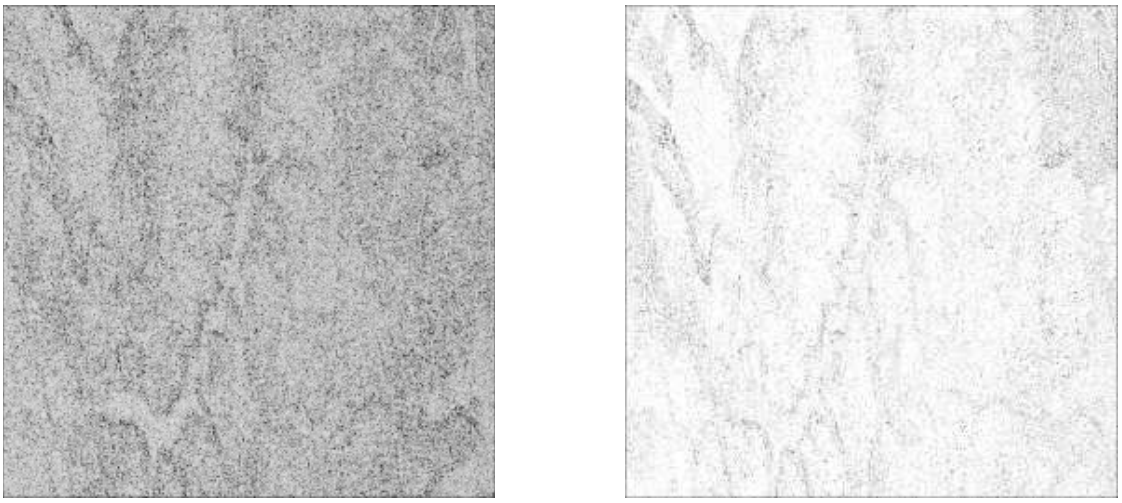
Figure 1: The minimal amplitude of the master and slave images of HH channel and polarimetric fusion.

4. EXPERIMENTAL RESULT

The experimental data used are L-band scattering matrix single-look complex image pairs of the Tien Shan test site acquired by the SIR-C/X-SAR radar system on Oct. 8 and 9, 1994.

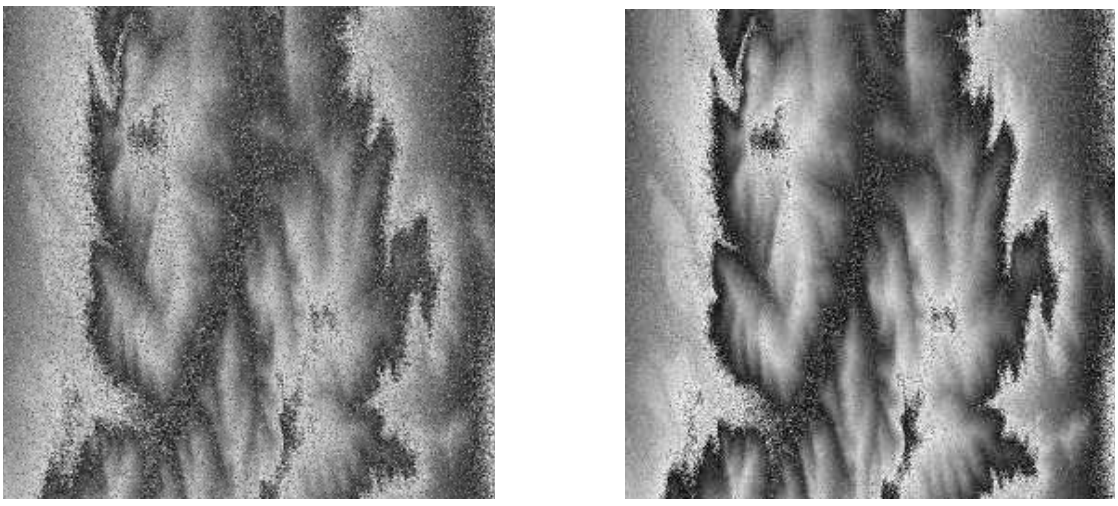
In Fig. 1, most of the terrain is hilly country. In the shadow, the coherence may be low. The minimal amplitude of the pair of images in HH channel is shown in Fig. 1(a). The average is 0.3250. After the coherence enhancement, the average of the fused image increases to 0.5414 and the image is shown in Fig. 1(b).

The coherence maps of HH channel and the fused images are shown in Fig. 2. The value is obviously larger after fusion. The mean coherence increases from 0.7632 to 0.8868. The enhanced coherence is much better than the one of HH channel. So the corresponding interferometric phase map becomes better after fusion, which is shown in Fig. 3. The residue number reduces from 94923 to 22528.



(a) HH channel (b) Polarimetric fusion

Figure 2: The coherence of HH channel and Polarimetric fusion.



(a) HH channel (b) Polarimetric fusion

Figure 3: The interferometric phase of HH channel and Polarimetric fusion.

5. CONCLUSION

Considering the infection of the signal amplitude to the phase, a novel polarimetric fusion method has been proposed to enhance the coherence. After projecting the two scattering vectors to the

same optimal direction, the less amplitude of a pair of projections is maximized. So the reliability (coherence) of the interferometric phase is enhanced effectively and the phase noise is reduced obviously. This will be helpful for phase unwrapping and DEM generation.

ACKNOWLEDGMENT

This work was supported in part by the National Natural Science Foundation of China (40571099), by Program for New Century Excellent Talents in University and by the Research Fund for the Doctoral Program of Higher Education of China.

REFERENCES

1. Cloude, S. R. and K. P. Papathanassiou, "Polarimetric SAR interferometry," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 36, 1551–1565, 1998.
2. Alberga, V., E. Krogager, M. Chandra, and G. Wanielik, "Potential of coherent decompositions in SAR polarimetry and interferometry," *IGARSS'04*, Vol. 3, 1792–1795, 2004.

Initial Polarimetric Calibration Results of ALOS PALSAR

T. Moriyama, M. Shimada and M. Watanabe

Earth Observation Research Center (EORC), Japan Aerospace and Exploration Agency
Tokyo, Japan

Abstract— This paper discusses the polarimetric calibration of Advanced Land Observing Satellite (ALOS) Phased Array L-band Synthetic Aperture (PALSAR). PALSAR is the first spaceborne L-band polarimetric synthetic aperture radar. Since the spaceborne SAR at low frequencies is affected by Faraday rotation effect which rotates the polarization plane of radar signal in the ionosphere, the scattering matrix which consists of the four polarimetric data (HH, HV, VH and VV) is distorted. Thus, in addition to the estimation of channel imbalance and cross-talk, to estimate Faraday rotation angle and to remove the contribution of Faraday rotation in the scattering matrix are an important problem in the polarimetric calibration of PALSAR. In this paper, we examine the polarimetric calibration of PALSAR. Freeman method is used to estimate the polarimetric calibration parameters including Faraday rotation angle.

1. INTRODUCTION

Advanced Land Observing Satellite (ALOS) was launched on January 24, 2006. ALOS has the three sensors which are the Panchromatic Remote-sensing Instrument for Stereo Mapping (PRISM) for digital elevation mapping, the Advanced Visible and Near Infrared Radiometer type 2 (AVNIR-2) for precise land coverage observation and the Phased Array type L-band Synthetic Aperture Radar (PALSAR) for day-and-night and all weather land observation. PALSAR has a unique observation capability in these three sensors and is the first spaceborne L-band polarimetric SAR. PLASAR data is expected to be used for many applications such as agriculture, forestry, ocean, natural disasters, etc. However, since the measured polarimetric data include the distortions, which are caused by the transmitting and receiving hardware [1], during polarimetric channels, the accuracy of inversion and classification algorithm depends on the polarimetric calibration. Moreover, Faraday rotation influences the measured polarimetric data. Since the radar wave travels through the ionosphere, its polarization plane is rotated. This mean that the cross-polarized echoes occur from trihedral and sphere targets which generate only the like-polarized echoes in theory. Faraday rotation effect is related to the solar activity which follows a cycle of approximately 11 years and the one-way maximum Faraday rotation angle at L-band is expected to be $\leqq 40$ degrees [2]. Fortunately, the next minimum solar activity is expected around 2007. However, it is important to investigate the influence of Faraday rotation on measurements of the polarimetric scattering matrix by PALSAR. Freeman introduced a method to calibrate polarimetric SAR data subject to Faraday rotation [3]. We estimate the channel imbalances and Faraday rotation angles based on his method and show some calibration results.

2. POLARIMETRIC CALIBRATION MODEL

The polarimetric measurements made by POLSAR can be modelled as follows [1]:

$$\boldsymbol{M} = A\exp(j\phi)\,\boldsymbol{RST} + \boldsymbol{n} \tag{1}$$

where A and ϕ are residual amplitude and phase calibration factors, and $\boldsymbol{M}$ and $\boldsymbol{S}$ are measured and true scattering matrices. It is assumed that the HV and VH backscatters are reciprocal ($S_{HV} = S_{VH}$). $\boldsymbol{R}$ and $\boldsymbol{T}$ are matrices representing the distortions on receiving and transmitting systems, and are expressed as,

$$\boldsymbol{R} = \begin{pmatrix} 1 & \delta_1 \\ \delta_2 & f_1 \end{pmatrix} \text{ and } \boldsymbol{T} = \begin{pmatrix} 1 & \delta_3 \\ \delta_3 & f_2 \end{pmatrix} \tag{2}$$

where the diagonal terms f_1 and f_2 are channel imbalance, and off diagonal terms δ are cross-talk. n is system noise. In calibrating the polarimetric data acquired from PALSAR, Faraday rotation becomes significant problem due to spaceborne L-band SAR system. If Faraday rotation influences the SAR signal, equation (1) is modefied as [3],

$$\boldsymbol{M} = A\exp(j\phi)\boldsymbol{RFSFT} + \boldsymbol{n}, \quad \boldsymbol{F} = \begin{pmatrix} \cos\Omega & \sin\Omega \\ -\sin\Omega & \cos\Omega \end{pmatrix} \tag{3}$$

Table 1: Characteristics of used data.

No.	Site	Observation date	Descending/ Ascending	Off-nadir angle[deg.]
1	Tomakomai, Japan	May 19, 2006	D	21.5
2	Alaska, USA	April 27, 2006	A	21.5
3	Alaska, USA	May 31, 2006	A	21.5
4	Alaska, USA	May 31, 2006	D	21.5
5	Alaska, USA	June 12, 2006	A	21.5
6	Alaska, USA	June 16, 2006	D	21.5

where $\boldsymbol{F}$ is the Faraday rotation matrix and Ω is the one-way Faraday rotation angle. Faraday rotation means the rotation of polarization plane as the radar signal travels through the ionized atmosphere. The contribution of Faraday rotation to true scattering matrix is written as follows:

$$\begin{array}{ll} M_{HH} = S_{HH}\cos^2\Omega - S_{VV}\sin^2\Omega, & M_{HV} = S_{HV} + (S_{HH} + S_{VV})\sin\Omega\cos\Omega \\ M_{VH} = S_{HV} - (S_{HH} + S_{VV})\sin\Omega\cos\Omega, & M_{VV} = S_{VV}\cos^2\Omega - S_{HH}\sin^2\Omega \end{array} \tag{4}$$

It can be seen that S_{HH} and S_{VV} appear in other polarization components. The approximated one-way Faraday rotation angle is given by [2]

$$\Omega = \frac{k}{f^2} \times B\cos\psi\sec\theta_0 \times TEC \quad \text{[radians]} \tag{5}$$

where k is a constant of value 2.365×10^4, B is the magnetic flux density, f is the frequency, and ψ and θ_0 are angle between earth's magnetic field and radar wave, and incident angle, respectively. TEC is the total electron content and depends on time of day, season, solar activity, geographical location, etc.

In this paper, Freeman method is used to evaluate the polarimetric calibration parameters and one-way Faraday rotation angle. His method is applied to the natural distributed areas which satisfy the azimuthal symmetry. This property means that the co- and cross-polarized responses are uncorrelated as $< S_{HH}\cdot S_{HV}^* >=< S_{HV}\cdot S_{VV}^* >= 0$ where $<\cdot>$ is an ensemble averaging processing. Then, the channel imbalance ratio (f_1/f_2) can be estimated from these areas. Moreover, if a trihedral corner reflector, which has the polarimetric characteristics $S_{HH} = S_{VV}$, is deployed in calibration site, the channel imbalances f_1 and f_2 can be calculated from the ratio of trihedral corner reflector echoes between HH and VV $(f_1 f_2)$ and the channel imbalance ratio (f_1/f_2). On the other hand, Freeman method assumes that a SAR system can be built with negligible levels of cross-talk ($-30\,$dB). Then the cross-talk removal can not be considered in his method. Faraday rotation angle is derived as follows:

$$\Omega = -\frac{1}{4}Arg\left(Z_{12}Z_{21}^*\right), \quad \begin{bmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{bmatrix} = \begin{bmatrix} 1 & j \\ j & 1 \end{bmatrix} \begin{bmatrix} M_{HH} & M_{HV} \\ M_{VH} & M_{VV} \end{bmatrix} \begin{bmatrix} 1 & j \\ j & 1 \end{bmatrix} \tag{6}$$

where M is the element of measured scattering matrix as Eq. (3). Therefore, the polarimetric calibration parameters such as channel imbalance and Faraday rotation angle can be given by Freeman method.

3. PALSAR POLARIMETRIC CALIBRATION RESULTS

In calibration phase, PALSAR observes many calibration sites where the corner reflectors are deployed and radiometric and polarimetric calibrations and data quality check are examined using these data. The calibration sites are distributed in the world. In order to evaluate the polarimetric calibration parameters and the one-way Faraday rotation angle, we used Tomakomai data, Japan and Alaska data, USA. The observation date, off-nadir angle, descending path and ascending path are listed in Table 1.

The typical latitude and Longitude of Tomakomai and Alaska sites are as follows:

$$\text{Tomakomai} : \text{N}42^\circ44'18.28'',\ \text{E}141^\circ31'26.18''$$

Table 2: Summary of Farafay rotation angle and TEC.

No.	Site	Absolute Faraday rotation angle [deg.]	TEC [TECU]
1	Tomakomai, Japan	1.77	10.8
2	Alaska, USA	0.09	6.0
3	Alaska, USA	0.95	8.5
4	Alaska, USA	2.93	11.5
5	Alaska, USA	2.12	8.8
6	Alaska, USA	2.14	9.1

Alaska : N63°48′00″, W145°48′10.8″.

Alaska is located in high latitudes as compared with Tomakomai. Fig. 1 shows the time variation of TEC in Tomakomai and Alaska on May 19, 2006. These data are acquired from the ionosphere products in AIUB anonymous ftp server site (ftp://ftp.unibe.ch/aiub/CODE/). The time spacing of these data is 2 hours. The peak of TEC appears around midday in local time of each area. It is expected that the descending path data is mainly influenced by Faraday rotation. It seems that TEC of Tomakomai is higher than that of Alaska slightly due to their latitudes. Moreover, these two sites include the area where the azimuth symmetry is satisfied and the trihedral corner reflector is deployed. We evaluated the channel imbalances f_1 and f_2, and the one-way Faraday rotation angle by Freeman method. Fig. 2 shows the amplitude and phase of f_1 and f_2. f_1 is not close to 1.0 and is smaller than f_2. This means that H channel does not balance with V channel on receiving system. The dispersion of f_1 and f_2 was seen to be less than ± 0.04 (peak to peak) in amplitude and ± 2 degrees in phase. These results indicate that the radar system is stable regardless of observation region, off-nadir angle, descending and ascending paths. Absolute Faraday rotation angle and TEC are shown in Table 2. TEC values are derived by the interpolation of the ionosphere products. Since there are not sufficient number of data, Faraday rotation angle can be related to the magnitude of TEC in the case of Alaska data. When TEC became 11.5 TECU in descending path of May 31, 2006, Faraday rotation angle 2.96 degrees. When TEC became 6.0 TECU in ascending path of April 27, 2006, the angle was estimated to be 0.09 degrees.

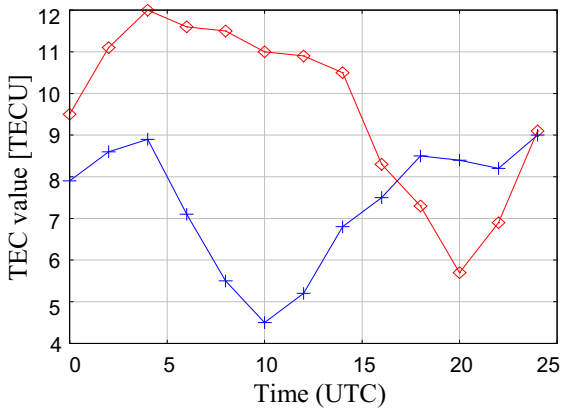


Figure 1: Time variation of TEC (red: Tomakomai, blue: Alaska).

Finally, we introduce some polarimetric analysis results before and after the polarimetric calibration. We use two analysis techniques such as the polarimetric signature and the entropy-alpha technique. The polarimetric signature is a way for visualising the target's scattering properties. For each of the possible incident polarizations, the strength of the backscatter can be computed for the same polarization on transmitting and receiving (Co-pol) and for orthogonal polarization on transmitting and receiving (X-pol). The polarimetric entropy and alpha can be obtained by eigenvalues and eigenvectors of a coherency matrix which is made from the scattering matrix. The

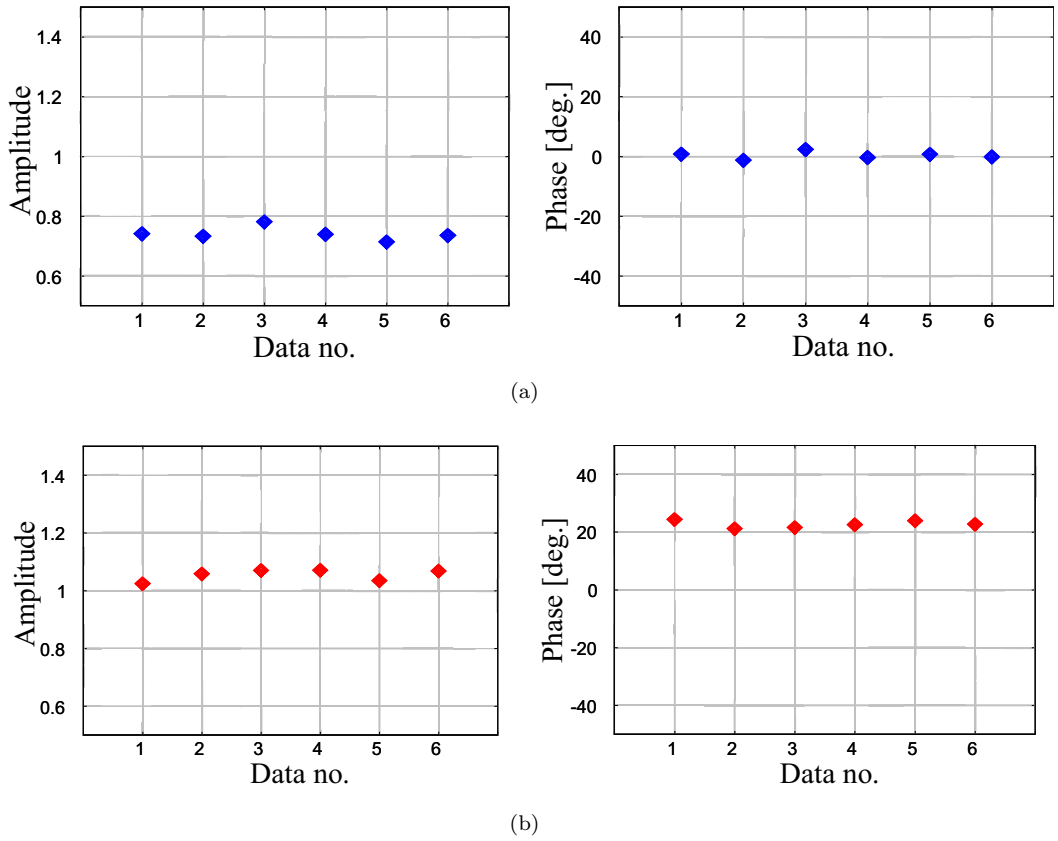


Figure 2: Amplitude and Phase of channel imbalances $f1$ and $f2$. (a) $f1$ and (b) $f2$.

polarimetric entropy shows randomness of scattering medium, and alpha represents average scattering mechanism. A target used for polarimetric analysis is a trihedral corner reflector. Fig. 3 indicates the measured co-polarized polarimetric signature of the trihedral corner reflector which was deployed in Tomakomai calibration site before and after the polarimetric calibration. It is confirmed that the distortion appears in un-calibrated data, while it is seen that the polarimetric signature of calibrated data is close to that of ideal target. On the other hand, entropy-alpha technique is applied to the same trihedral corner reflector. The entropy and alpha of the ideal trihedral corner reflector become zero. An area including CR echo is used for this evaluation. These averaged parameters of measured data before and after the polarimetric calibration are as follows:

$$\text{Before} : H = 0.12,\ \alpha = 16.56[\text{deg.}]$$

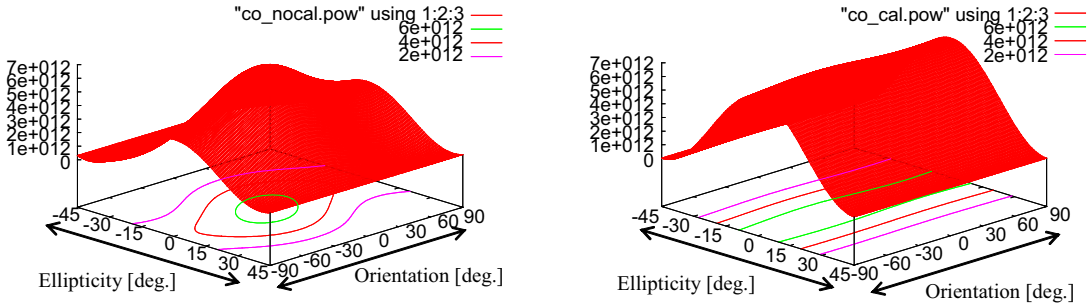


Figure 3: Co-polarized polarimetric signature of measured corner reflector before and after polarimetric calibration, (a) before calibration, (b) after calibration.

$$\text{After} : H = 0.13,\ \alpha = 3.32[\text{deg.}]$$

Therefore, these results show that the calibrated data has the expected polarimetric characteristic.

4. CONCLUSIONS

We examined the polarimetric calibration of ALOS PALSAR. In order to estimate the polarimetric calibration parameters, we used Freeman calibration method, because PALSAR data is affected by Faraday rotation and his method can derive this rotation angle. The analysis data were Tomakomai and Alaska calibration sites. Alaska is located in high latitude area as compared with Tomakomai. The channel imbalance $f1$ and $f2$ became stable with respect to amplitude and phase. Faraday rotation angle fluctuated between 2.93 and 0.09 with the change of TEC in the case of Alaska. However, it seems that this angle is very small.

REFERENCES

1. Freeman, A., "SAR calibration: An overview," *IEEE Trans. Geosci. Remote Sensing*, Vol. 30, No. 6, 1107–1121, 1992.
2. Wright, A. P., et al., "Faraday rotation effects on L-band spaceborne SAR data," *IEEE Trans. Geosci. Remote Sensing*, Vol. 41, No. 12, 2735–2744, 2003.
3. Freeman, A., "Calibration of linearly polarized polarimetric sar data subject to faraday rotation," *IEEE Trans. Geosci. Remote Sensing*, Vol. 42, No. 8, 1617–1624, 2004.

Application of the Array Scanning Method to 1D-periodic Microstrip Lines

R. Rodríguez-Berral[1], **F. Mesa**[1], **P. Baccarelli**[2], **and P. Burghignoli**[2]

[1]Department of Applied Physics 1, University of Seville, Av. Reina Mercedes s/n, 41012 Seville, Spain
[2]Department of Electronic Engineering, "La Sapienza" University of Rome
Via Eudossiana 18, 00184 Rome, Italy

Abstract— The Array Scanning Method (ASM) is employed to analyze the excitation of 1D-periodic microstrip lines by a *nonperiodic* source. It should be noted that the nonperiodic nature of the source precludes the direct application of Floquet's theorem. However, a combination of the ASM with the Method of Moments allows for the computation of the current density launched on the line. Some results are presented showing the current excited by a delta-gap voltage source at frequencies corresponding to passband as well as to stopband regimes of the periodic microstrip line.

Planar printed structures with periodicity along one or two dimensions (1D- or 2D-periodic structures, respectively) are of great interest in their application to the design of passive microwave circuit components such as filters, EBG's, and leaky-wave antennas with scanning capabilities. Most of the previous studies devoted to these structures have been restricted to the computation of the propagation characteristics of their modal solutions (Bloch waves or Floquet modes) to find, for example, the different frequency passbands and stopbands and the radiative regimes.

This work proposes an efficient approach to deal with the excitation of 1D-periodic microstrip lines by a single nonperiodic source. The presence of a nonperiodic source breaks the periodicity of the problem, thus causing the usual techniques based on the direct application of Floquet's theorem not to be applicable. This fact notably increases the degree of difficulty of the problem under consideration with respect to the modal analysis of the periodic structure that is being excited. As it will be explained, the formulation presented here to overcome this difficulty relies on the fact that the current excited by the nonperiodic source can be expressed as a continuous superposition of Floquet-periodic functions that are solutions to auxiliary Floquet-periodic problems. For simplicity, a simple canonic structure will be considered, which consists of a uniform microstrip line periodically loaded with gaps (or, equivalently, an infinite 1D array of rectangular metallic patches) and excited by a delta-gap voltage source. In order to solve the above-mentioned auxiliary problems, the method of moments is employed.

Our goal is the computation of the current excited on a periodic microstrip line when it is excited by a delta-gap voltage source. Specifically, the structure under study (see Fig. 1) is an array of identical and perfectly conducting rectangular patches that extends indefinitely along the $\hat{\mathbf{z}}$ direction; namely, a 1D-periodic structure whose unit cell has length p. The patches are printed on a lossless isotropic dielectric substrate of thickness h. This structure is excited by a delta-gap

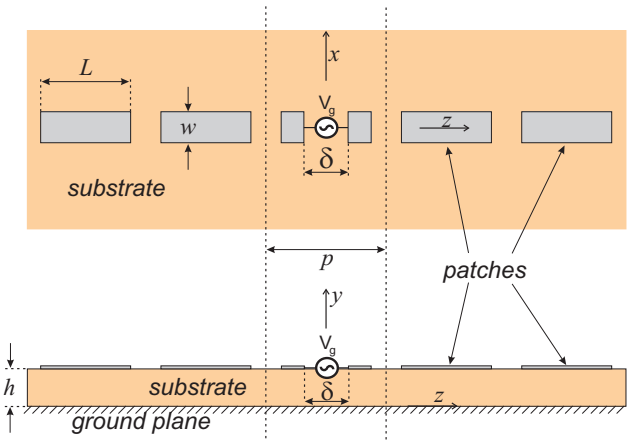


Figure 1: Top and side views of the structure under analysis.

voltage source of finite length δ centered on one of the patches, which in the following will be referred to as the *central* patch. This excitation can be modelled as an impressed electric field in the gap region, namely

$$\mathbf{E}(x,z) = E_g(z)\hat{\mathbf{z}}, \quad E_g(z) = -V_g \mathrm{rect}\left(\frac{z}{\delta}\right). \tag{1}$$

It is important to note that the presence of this impressed field causes the problem to be nonperiodic, thus preventing the direct application of the usual techniques based on Floquet's theorem. Nevertheless, the gap field can be expressed in terms of the following integral:

$$E_g(z) = \frac{p}{2\pi}\int_{-\pi/p}^{\pi/p} E_g^{\infty}(z;k_z)\mathrm{d}k_z, \tag{2}$$

where

$$E_g^{\infty}(z;k_z) = \sum_{n=-\infty}^{\infty} E_g(z-np)e^{-jk_z np}. \tag{3}$$

Note that $E_g^{\infty}(z;k_z)$ is a Floquet-periodic function of z with period p, i.e., a function such that $E_g^{\infty}(z+p;k_z) = e^{-jk_z p}E_g^{\infty}(z;k_z)$. Thus, the problem defined by the excitation of the periodic structure depicted in Fig. 1 by an impressed field of the form $E_g^{\infty}(z;k_z)$ turns into a Floquet-periodic problem that can be readily solved by applying Floquet's theorem. Once this auxiliary Floquet-periodic problem is solved, the solution to the original nonperiodic problem can be found by superposition making use of (2). This technique, in which the original problem is decomposed into a continuous superposition of periodic problems, is known in the literature as *Array Scanning Method* [1, 2].

Let us consider a field of the form (3) impressed on the periodic array of patches depicted in Fig. 1. As the patches are perfect conductors, the tangential electric field on their surface must equal the impressed field. Hence, the following integral equation must be satisfied on the surface of *all* the patches:

$$\hat{\mathbf{y}} \times \mathbf{E}[\mathbf{J}^{\infty}] = E_g^{\infty}\hat{\mathbf{x}}, \tag{4}$$

where $\mathbf{J}^{\infty}$ is the surface current density on the patches. In order to apply the method of moments to equation (4), the current density is next expanded into a linear combination of basis functions. For simplicity, only the longitudinal component of the current density will be considered, as well as only one basis function for its dependence on the x coordinate. Although the integral equation (4) must be enforced on all the patches, it is sufficient to enforce it on one single patch due to the Floquet-periodic nature of the problem.

Next, the integral equation is solved *on the central patch* by applying the method of moments in a Galerkin formulation. Taking now into account that the original nonperiodic excitation was written as a combination of Floquet-periodic impressed fields by means of the linear transformation in (2), after applying the superposition principle, the current density excited by the delta-gap source on the periodic microstrip line can be finally computed as

$$J(x,z) = \frac{p}{2\pi}\int_{-\pi/p}^{\pi/p} J^{\infty}(x,z;k_z)\mathrm{d}k_z \tag{5}$$

Next, some results obtained by using the proposed formulation are shown. Fig. 2 shows the amplitude of the current excited by the delta-gap source on the periodic microstrip line in Fig. 1 at 13 GHz (see the caption for the details of the structure). This value of frequency is located within a passband region since the periodic microstrip line has one real bound mode (BM) above cutoff, whose fundamental spatial harmonic wavenumber is $k_z^{BM} = 1.63k_0$. Far enough from the central patch, the current on the line is accounted for basically by the bound-mode current, since it will be the only non-attenuating component of the current.

Figure 3 shows the current excited on the same periodic microstrip line at 15 GHz. Unlike the passband case above, this value of the frequency is located within the first stopband of the periodic microstrip line. The wavenumbers of the spatial harmonics that constitute the BM are now complex (all of them having the same imaginary part) and lie on the boundary of the Brillouin zones. Specifically, the wavenumber of the fundamental harmonic is $k_z^{BM} = \pi/p - j0.87k_0$. As a consequence, the BM fields and current decay exponentially in z. It is interesting to note that in this case the BM fields do not carry power but they are a purely reactive standing wave.

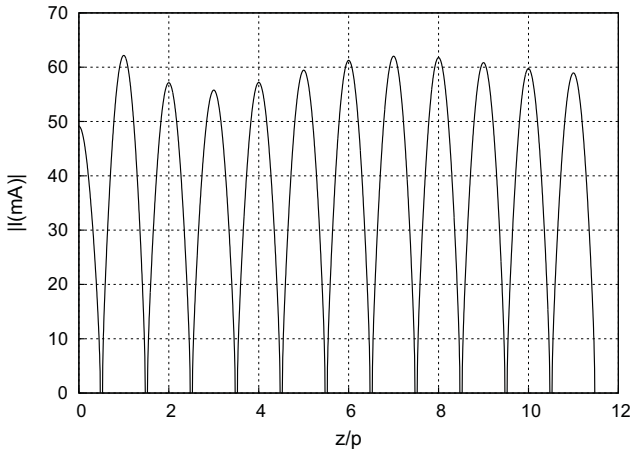


Figure 2: Amplitude of the current excited by a delta-gap voltage source of length $\delta = 0.1\,\text{mm}$ on a periodic microstrip line as in Fig. 1 with $p = 4\,\text{mm}$, $L = 3.8\,\text{mm}$, $w = 0.6\,\text{mm}$, $h = 0.767\,\text{mm}$, and $\varepsilon = 10.2\varepsilon_0$. The frequency is 13 GHz.

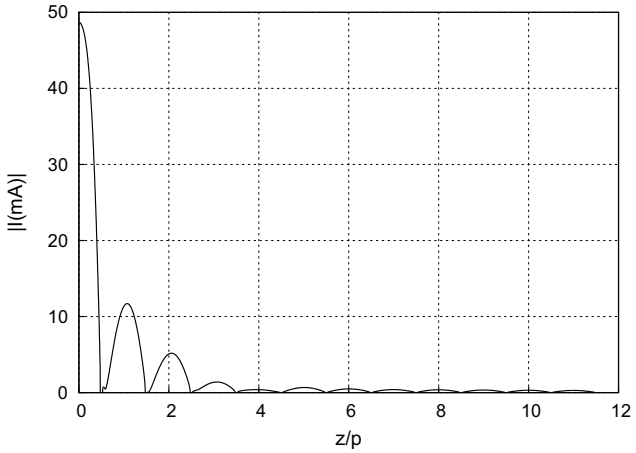


Figure 3: Amplitude of the current excited on the same periodic microstrip line of Fig. 2. The frequency is now 15 GHz.

ACKNOWLEDGMENT

The work of R. Rodríguez Berral y F. Mesa has been supported by the Spanish Ministry of Education and Science and the European Union (FEDER funds), project TEC2004-03214, and by Junta de Andalucía, project TIC.253.

REFERENCES

1. Munk, B. A. and G. A. Burrell, "Plane-wave expansion for arrays of arbitrarily oriented piecewise linear elements and its application in determining the impedance of a single linear antenna in a lossy halfspace," *IEEE Trans. Antennas. Prop.*, Vol. AP-27, 331–343, May 1979.
2. Capolino, F., D. R. Jackson, and D. R. Wilton, "Mode excitation from sources in two-dimensional EBG waveguides using the array scanning method," *IEEE Microw. Wireless Comp. Lett.*, Vol. 15, 49–51, Feb. 2005.

Closed-form Expressions for Layered Media Green's Functions that are Reliable Both in the Near Field and in the Far Field

R. R. Boix[1], **F. L. Mesa**[2], **and F. Medina**[1]

[1]Department of Electronics and Electromagnetism, College of Physics, University of Seville
Av. Reina Mercedes, s/n, 41012, Seville, Spain
[2]Department of Applied Physics 1, School of Computer Engineering, University of Seville
Av. Reina Mercedes, s/n, 41012, Seville, Spain

Abstract— The authors have developed an algorithm for the determination of accurate closed-form expressions of Green's functions for multilayered media. The algorithm is based on the fitting of the spectral domain Green's functions in terms of an asymptotic term plus a ratio of two polynomials. The fitting is carried out via the method of total least squares, which is a non-iterative method demanding very few computational resources. The resulting closed-form expressions for the Green's functions consist of a quasistatic term plus a sum of cylindrical waves. The quasistatic term provides the near-field behavior and the cylindrical waves provide the farfield behavior. As a consequence of this, the closed-form expressions are very accurate irrespective of the distance between the source point and the observation point.

The application of the method of moments (MoM) to the solution of mixed potential integral equations (MPIE) has proven to be an efficient numerical tool for the analysis of planar circuits and antennas [1]. In fact, several commercial software products that are currently used in the design of planar structures (such as "Ansoft Ensemble", "Zeland IE3D", and "Agilent Momentum") are based on the solution of MPIE by means of the MoM. In order to solve the MPIE arising from the analysis of planar structures, it is necessary to calculate the Green's functions (GF) for the scalar and vector potentials in multilayered media. These GF can be determined via numerical computation of infinite integrals that are commonly known as Sommerfeld integrals (SI). However, the highly-oscillatory nature of the integrands involved makes the numerical computation of SI cumbersome and time consuming [1].

Some researches have developed specific techniques that make it possible to accelerate the numerical computation of SI [2, 3]. Unfortunately, these techniques have the drawback that they have to be repeatedly used every time the distance between the source and observation points changes. One technique for the computation of multilayered media GFs that has received much attention in the last fifteen years is the discrete complex image method (DCIM) [4]. This technique makes it possible to obtain closed-form expressions of the GFs without requiring numerical integration. However, the closed-form expressions have been recently found to produce unpredictable errors in the far-field computation of the GFs [5]. Although it is possible to raise the threshold value of the distance between source and observation points that marks the onset of numerical errors, this is achieved at the expense of a considerable CPU time consumption [5]. The sensitivity of the DCIM to far-field numerical errors seems to decrease when a surface waves term is included in the closed-form expression of the GFs [4, 6]. Unfortunately, the surface-waves term contains poles of the spectral domain GFs as well as residues of the spectral domain GFs at these poles, and the accurate determination of both the poles and the residues requires time-consuming algorithms [6].

Very recently, the rational function fitting method (RFFM) has been proposed as an alternative technique to DCIM for the determination of closed-form expressions of multilayered media GFs [7]. In the new technique the GFs are expressed in terms of a quasistatic term plus a linear combination of cylindrical waves. The amplitudes and propagation constants of the cylindrical waves are obtained in the spectral domain by means of a vector fitting algorithm, which is an iterative algorithm. The main drawback of the RFFM is that the closed-form expressions for the GFs introduce unacceptable errors in the near-field [7]. These errors are caused by the Hankel functions representing the cylindrical waves, and in particular, by the non-physical near-field singularities present in these Hankel functions. In this paper, we present a revised version of the RFFM which substantially improves the original approach published in [7]. In one hand, the determination of the amplitudes and propagation constant of the cylindrical waves is carried out via the method of total least squares (TLS), which is a non-iterative algorithm with CPU time and memory requirements that are substantially smaller than those required by the vector fitting algorithm. On the

other hand, in the new version of the RFFM the effect of the cylindrical waves singularities on the closed-form expression of the GFs is mathematically suppressed, and these closed-form expressions do not lead any more to near-field numerical errors.

Let $(x',\, y',\, z')$ be the coordinates of an infinitesimal dipole source embedded in a multilayered medium, and let $(x,\, y,\, z)$ be the coordinates of an observation point placed in that multilayered medium. If we restrict ourselves to formulation C of Michalski and Zheng for MPIE [8], the GF for the scalar potential as well as the elements of the dyadic GF for the vector potential can all be obtained by means of SI of the type

$$S_n\{\widetilde{G}(k_\rho)\} = \frac{1}{2\pi}\int_{0(\mathrm{SIP})}^{\infty} J_n(k_\rho\rho)k_\rho^{n+1}\widetilde{G}(k_\rho)\mathrm{d}k_\rho \quad (n=0,\, 1) \tag{1}$$

where $\rho = \sqrt{(x-x')^2+(y-y')^2}$, $J_n(\bullet)$ is the Bessel function of order n, $\widetilde{G}(k_\rho)$ stands for the spectral counterpart of any of the aforementioned GF, and SIP is an integration path in the first quadrant of the complex k_ρ plane that detours around the poles and branch points of $\widetilde{G}(k_\rho)$ [1]. For an arbitrary multilayered medium, the functions $\widetilde{G}(k_\rho)$ can be obtained by means of the algorithms shown in the Appendix II of [8].

By analogy with the theory presented in [7], in this work the functions $\widetilde{G}(k_\rho)$ are approximated by expressions of the type

$$\widetilde{G}(k_\rho) \approx \widetilde{G}_{\mathrm{as}}(k_\rho) + \sum_{i=1}^{M}\frac{a_i}{k_\rho^2 - p_i^2} \tag{2}$$

In Eq. (2), $\widetilde{G}_{\mathrm{as}}(k_\rho)$ stands for the asymptotic behavior of $\widetilde{G}(k_\rho)$ for large k_ρ. This term is responsible for the near-field behavior of the spatial domain GF $S_n\{\widetilde{G}(k_\rho)$. Concerning the finite pole-residue series of Eq. (2), its spatial domain couterpart is a linear combination of cylindrical waves [7].

When (2) is introduced in (1), the following equations are obtained

$$S_0\{\widetilde{G}(k_\rho)\} \approx S_0\{\widetilde{G}_{\mathrm{as}}(k_\rho)\} - \frac{j}{4}\sum_{i=1}^{M} a_i H_0^{(2)}(p_i\rho) \tag{3}$$

$$S_1\{\widetilde{G}(k_\rho)\} \approx S_1\{\widetilde{G}_{\mathrm{as}}(k_\rho)\} - \frac{j}{4}\sum_{i=1}^{M} a_i p_i H_1^{(2)}(p_i\rho) \tag{4}$$

where $H_n^{(2)}(\bullet)$ are Hankel functions of second kind and order n.

As commented in [7], the Eqs. (3) and (4) lead to closed-form expressions of multilayered media GF provided the integrals $S_0\{\widetilde{G}_{\mathrm{as}}(k_\rho)$ and $S_1\{\widetilde{G}_{\mathrm{as}}(k_\rho)$ can be computed analytically. Therefore, with a view to obtaining closed-form expressions of the GF, the spatial domain counterpart of $\widetilde{G}_{\mathrm{as}}(k_\rho)$ must be analytical. Also, in order to ensure the accuracy of the approximation of Eq. (2), the functions $\widetilde{G}_{\mathrm{as}}(k_\rho)$ must be chosen in such a way that they do not introduce singularities which are not present in $\widetilde{G}(k_\rho)$ (please note that this latter condition is not satisfied by the functions $\widetilde{G}_{\mathrm{as}}(k_\rho)$ defined in [7]). In the current work, the aformentioned constraints on $\widetilde{G}_{\mathrm{as}}(k_\rho)$ have all been fulfilled by expressing $\widetilde{G}_{\mathrm{as}}(k_\rho)$ in terms of the functions defined in Table 1 of [9].

Once the functions $\widetilde{G}_{\mathrm{as}}(k_\rho)$ are chosen, the coefficients a_i and $p_i (i = 1,\, \ldots,\, M)$ must be obtained in such a way that the approximation of Eq. (2) is as accurate as possible. In this paper, the choice of a_i and p_i has been optimized by means of the method of TLS [10]. For that purpose, Eq. (2) has been rearranged so as to express the finite pole-residue series as a ratio of two polynomials. Then, the resulting expression has been sampled at a set of $N > 2M$ values of $k_\rho (k_\rho = k_{\rho,\,j};\, j = 1,\, \ldots,\, N)$, which has led to an overdetermined linear system of equations for the coefficients of the polynomials. The system of equations has been solved by means of the method of TLS, and then, the coefficients a_i and p_i have been subsequently determined in terms of the coefficients of the polynomials. One advantage arising from the application of the method of TLS is that the method gives an estimate for M (the number of terms retained in the pole-residue series) from the number of nonzero singular values obtained in the singular value decomposition [10]. Since the functions $\widetilde{G}(k_\rho)$ may have poles in the real axis of the complex k_ρ plane, the samples $k_{\rho,\,j}\, (j = 1,\, \ldots,\, N)$ have been taken along a path placed in the first quadrant of the k_ρ plane [4].

It should be pointed out that the Hankel functions of Eqs. (3) and (4) have singularities as $\rho \to 0$ (in fact, $H_0^{(2)}(p_i\rho) \propto \ln\rho$ and $H_1^{(2)}(p_i\rho) \propto \rho^{-1}$ as $\rho \to 0$), but these singularities are not necessarily shared by $S_0\{\widetilde{G}(k_\rho)\}$ and $S_1\{\widetilde{G}(k_\rho)\}$. These non-physical singularities introduced by the Hankel functions prevent the approximations of (3) and (4) from giving accurate values of $S_0\{\widetilde{G}(k_\rho)\}$ and $S_1\{\widetilde{G}(k_\rho)\}$ as $\rho \to 0$ (which is recognized in [7]). In order to avoid this problem, we have carried out an expansion of the approximations of (3) and (4) for low ρ, and we have set to zero the coefficients multiplying the singular terms. As a result of this, we have obtained new equations for the coefficients a_i and p_i, and these new equations have been incorporated to the system of linear equations that is solved via the method of TLS. By doing this, the coefficients a_i and p_i are calculated in a way that intentionally avoids the non-physical singularities of Hankel functions as $\rho \to 0$ in (3) and (4).

The results obtained with the approximations of Eqs. (2), (3) and (4) have been compared with accurate values of $\widetilde{G}(k_\rho)$ and $S_n\{\widetilde{G}(k_\rho)\}\,(n = 0, 1)$ (in the case of $S_n\{\widetilde{G}(k_\rho)\}\,(n = 0, 1)$, these accurate values have been obtained by means of brute force numerical integration of SI). Excellent agreement has been found with both the spectral domain approximations and the spatial domain approximations. In the case of the spatial domain approximations, the near-field behavior of $S_n\{\widetilde{G}(k_\rho)\}$ has been found to be dominated by $S_n\{\widetilde{G}_{\rm as}(k_\rho)\}$ since the elimination of Hankel functions singularities prevent the Hankel functions from influencing the values of $S_n\{\widetilde{G}(k_\rho)\}$ as $\rho \to 0$. The far-field behavior of $S_n\{\widetilde{G}(k_\rho)\}$ has been found to be dominated by the cylindrical waves for which $|\mathrm{Im}(p_i)| <<< 1$. This latter result agrees with the discussion following Table 1 of [7].

ACKNOWLEDGMENT

This work has been supported by the Spanish Ministerio de Educación y Ciencia and European Union FEDER funds (project TEC2004-03214), and by Junta de Andalucia (project TIC-53)

REFERENCES

1. Mosig, J. R., "Integral equation technique," *Numerical Techniques for Microwave and Millimeter-Wave Passive Structures*, edited by T. Itoh, 133–213, Wiley, New York, 1989.
2. Michalski, K. A., "Extrapolation methods for Sommerfeld integral tails," *IEEE Trans. Antennas Propagat.*, Vol. 46, 1405–1418, Oct. 1998.
3. Mosig, J. R. and Álvarez-Melcón, "Green's functions in lossy layered media: integration along the imaginary axis and asymptotic behavior," *IEEE Trans. Antennas Propagat.*, Vol. 51, 3200–3208, Dec. 2003.
4. Chow, Y. L., J. J. Yang, D. G. Fang, and G. E. Howard, "A closed-form spatial Green's function for the thick microstrip substrate," *IEEE Trans. Microwave Theory Tech.*, Vol. 39, 588–592, Mar. 1991.
5. Yuan, M., T. K. Sarkar, and M. Salazar-Palma, "A direct discrete complex image method from the closed-form Green's functions in multilayered media," *IEEE Trans. Microwave Theory Tech.*, Vol. 54, 1025–1032, Mar. 2006.
6. Teo, S. A., S. T. Chew, and M. S. Leong, "Error analysis of the discrete complex image method and pole extraction," *IEEE Trans. Microwave Theory Tech.*, Vol. 51, 406–413, Feb. 2003.
7. Kourkoulos, V. N. and A. C. Cangellaris, "Accurate approximation of Green's functions in planar stratified media in terms of a finite sum of spherical and cylindrical waves," *IEEE Trans. Antennas Propagat.*, Vol. 54, 1568–1576, May 2006.
8. Michalski, K. A. and D. Zheng, "Electromagnetic scattering and radiation by surfaces of arbitrary shape in layered media, part I: theory," *IEEE Trans. Antennas Propagat.*, Vol. 38, 335–344, Mar. 1990.
9. Demuynck, F. J., G. A. E. Vandenbosch, and A. R. van de Capelle, "The expansion wave concept-part I: efficient calculation of spatial Green's functions in a stratified dielectric medium," *IEEE Trans. Antennas Propagat.*, Vol. 46, 397–406, Mar. 1998.
10. Rahman, J. and T. K. Sarkar, "Deconvolution and total least squares in finding the impulse response of an electromagnetic system from measured data," *IEEE Trans. Microwave Theory Tech.*, Vol. 54, 416–421, Apr. 1995.

Coupling of Microstrip Lines Exciting the Magnetostatic Surface Waves

Jae-Hyun Lee
Chungnam National University, Korea

Abstract— The coupling of microstrip lines which excite the magnetostatic surface waves is analyzed. The asymmetric excitation of microstrip line and the nonreciprocal propagation characteristic of the magnetostatic surface waves are considered. For the fast numerical calculation, the Jacobi polynomial is used as the basis function to satisfy the edge condition.

1. INTRODUCTION

In order to use a magnetostatic surface wave (MSSW) in the delay line or for signal processing at the microwaves, one needs the characterization of the excitation methods. Microstrip line excitation of the MSSW in the multilayer structure including a ferrite film was calculated approximately by assuming that the current distribution is uniform in the cross-section and has no variation in the transmission direction [1]. Then, the full-wave spectral domain method analysis of the MSSW excited by the microstrip was shown to be numerically possible to account for the complex propagation constant along the transmission direction of the lossless microstrip line [2, 3]. It shows that the current distribution on the microstrip line is asymmetric in the cross-section and the propagation constant is complex which is understood as the power coupled to the radiation of the MSSW. In those analyses, surface current densities on the microstrip line may be expanded by Chebyshev polynomials multiplied by a factor that satisfies the microstrip line edge conditions asymptotically and is equal to the microstrip line with a dielectric substrate.

However, the size of the matrix to calculate the unknown coefficients of Chebyshev polynomials is large, for example, 11×11, in order to obtain the converging solution and so the computational time is great. In the microstrip line with a dielectric substrate, accurate solutions result even if an extremely small size matrix, for example, 2×2, is used. This is because qualitative natures such as the edge condition of the surface current distribution can be incorporated in the choice of basis functions. That is one of the features of the spectral domain method. Much better convergence can be obtained from a small size matrix if we choose a few basis functions which represent physical characteristics of the charge distributions on the strips. So we should choose basis functions representing the asymmetric current distribution on the microstrip line.

In this paper, the edge condition is derived using the Mexiner's method and the Jacobi polynomial may be used as the basis function to satisfy the derived edge conditions. The effects of the nonreciprocal propagation of magnetostatic surface wave.

2. FULL-WAVE FORMULATION

The structure is shown in Fig. 1. A ferrite film is sandwiched between conductors covered with two dielectric layers, and an infinitely long microstrip excites this structure. Each layer is homogeneous and lossless. The magnetization direction of the ferrite coincides with the transmission direction of the microstrip and is taken as the z-axis and the cross-section is taken as the x-y plane.

Since this structure is open and symmetric along the z and x axes, the field solutions may be assumed to have $\exp[-j(k_z z + k_x x)]$ dependences, where k_z and k_x are wave numbers to be determined along the z and the x axes, respectively. One may represent the field and the current as a superposition of these exponential eigenfunctions in the open domain, i.e., the Fourier integral with respect to x, as

$$E(x, y, z) = e^{-jk_z z} \frac{1}{2\pi} \int_{-\infty}^{\infty} \tilde{E}(k_x; y) e^{-jk_x x}\, dk_x \tag{1}$$

where E is the field and $\tilde{E}(k_x; y)$ is the modal amplitude corresponding to k_x.

Taking their integrand from the Fourier integrals, one obtains

$$\begin{bmatrix} \tilde{E}_z(k_x) \\ \tilde{E}_x(k_x) \end{bmatrix} = \begin{bmatrix} \tilde{G}_{zz}(k_x, k_z) & \tilde{G}_{zx}(k_x, k_z) \\ \tilde{G}_{xz}(k_x, k_z) & \tilde{G}_{xx}(k_x, k_z) \end{bmatrix} \begin{bmatrix} \tilde{J}_z(k_x) \\ \tilde{J}_x(k_x) \end{bmatrix} \tag{2}$$

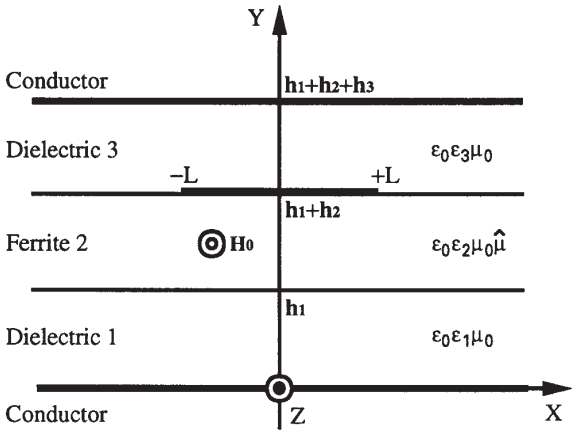


Figure 1: Geometry of a microstrip exciting a ferrite film.

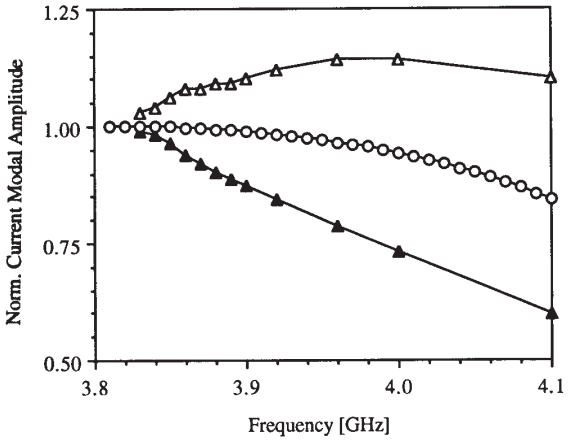


Figure 2: Normalized modal amplitude of the surface current versus frequency for the full-wave calculation in the $+x$ direction (open triangles) and the $-x$ direction (closed triangles), and the magnetostatic approximation with the uniform current distribution (circles).

where $\tilde{G}_{ij}(k_x, k_z)$, $i = x,\ z$ and $j = x,\ z$ are the Fourier transformed components of the dyadic Green's functions, $G_{ij}(x,\ z)$, with respect to x in the plane of $y = h_1 + h_2$ and $\exp(-jk_z z)$ dependences are suppressed.

Surface current densities in the microstrip may be expanded by polynomials satisfying the microstrip edge conditions for the fast numerical calculation. The singularities of the microstrip line edge on the ferrite substrate may be calculated using the Mexiner's method [4, 5]. The calculated field behavior is

$$J_z\,(x) \propto \rho^{-\xi} \tag{3}$$

where

$$\xi = \frac{1}{2} \pm j\frac{1}{2\pi} \ln\left[\frac{(\mu - \kappa + 1)}{(\mu + \kappa + 1)}\frac{(\mu + \kappa)}{(\mu - \kappa)}\right]. \tag{4}$$

From the consideration that the current distribution at the left edge is different from that at the right edge, the surface current distributions, $J_i(x,\ z)$, $i = z,\ x$, can be expanded in terms of the Jocobi polynomials, $P_n^{(\zeta,\eta)}(x)$.

$$J_i\,(x, z) \;=\; e^{-jk_z z} \sum_{n=0}^{N_i} C_{in} J_{in}\,(x) \tag{5}$$

$$J_{zn}(x) \;=\; \begin{cases} \left(1 - \frac{x}{L}\right)^{\zeta} \left(1 + \frac{x}{L}\right)^{\eta} P_n^{(\zeta,\eta)}\left(\frac{x}{L}\right) & |x| < L \\ 0 & |x| > L \end{cases} \tag{6}$$

Substituting the modal (Fourier) amplitudes of the surface current densities, $\tilde{J}_{in}(k_x)$, with unknown coefficients C_{in} into Equation (2), one obtains two equations for $\tilde{E}_i$, $i = x,\ z$. One may satisfy the boundary condition in the microstrip by taking an inner product of $\tilde{E}_i(k_x)$ and $\tilde{J}^*_{im}(k_x)$ as

$$\int_{-\infty}^{\infty} \tilde{J}^*_{im}(k_x)\tilde{E}(k_x)\,dk_x = \sum_j \sum_{m=0}^{N_j} C_{jn} \int_{-\infty}^{\infty} \tilde{J}^*_{im}(k_x)\tilde{G}_{ij}(k_x, k_z)\tilde{J}_{jn}(k_x)\,dk_x \tag{7}$$

where j equals x and z and the asterisk denotes the complex conjugate.

Since the tangential electric fields are zero in the conducting microstrip and the surface current densities are zero outside the microstrip, the left-hand side of Equation (7) becomes zero everywhere

in the plane $y = h_1 + h_2$ due to the Parseval's theorem. For nontrivial solutions of C_{in}, the values of k_z may be obtained numerically by forcing the determinant of these simultaneous integral equations to be zero. With the values of k_z, C_{in} may be obtained by solving the homogeneous simultaneous Equation (7), which gives the surface current densities in Equations (5)–(6).

For full-wave calculation, a microstrip excitation of the three-layer structure covered by conductors, shown in Fig. 1, is chosen with heights of layers, $h_1 = 250\,\text{cm}$, $h_2 = 6.3\,\mu\text{m}$, and $h_3 = 250\,\mu\text{m}$, relative dielectric constants, $\varepsilon_1 = 1$ and $\varepsilon_2 = \varepsilon_3 = 10$, microstrip halfwidth, $L = 25\,\mu\text{m}$, internal magnetic bias field for ferrite, $H_0 = 740\,\text{Oe}$, saturation magnetization, $4\pi M_s = 1700\,\text{Oe}$, and gyromagnetic ratio, $\gamma = 2.8\,\text{MHz/Oe}$.

Numerical calculations of the surface current densities in the microstrip are found for the expansion of J_x and J_z by the Jacobi polynomials, respectively, to ensure numerical accuracy. Calculated current distributions across the microstrip (J_z versus x) are asymmetric [3] and their asymmetry is reduced for the lower-frequency excitation.

From the calculated current distributions, one may obtain its modal amplitudes exciting the magnetic surface wave, as shown in Fig. 2, by taking the Fourier transform of the current distribution and picking the spatial frequency key corresponding to either $+x$ or $-x$ propagation for a given temporal frequency. Fig. 2 shows that the excitation amplitude for the $+x$ direction is larger than the $-x$ direction, while the uniform current distribution yields almost the same excitation amplitude values between the $+x$ and $-x$ directions. The difference in the excitation amplitudes between the $+x$ and $-x$ directions increases as the frequency increases.

Pointing vectors and their integration over the radiating surfaces give the radiating power and its direction carried by the magnetic surface waves. Fig. 3 shows that the magnitude of the power peaks at 3.86 GHz, its direction is -112.5° from the z-axis toward the $-x$ axis, and it is about 10 dB larger than that of the $+x$ direction. The directions of the real power flow are distributed in a sector about 30 degrees closer to the x axis and, as the frequency increases, approach the $\pm x$ direction. The magnetostatic assumptions give the direction of power flow along the x axis, which is different from the full-wave analysis.

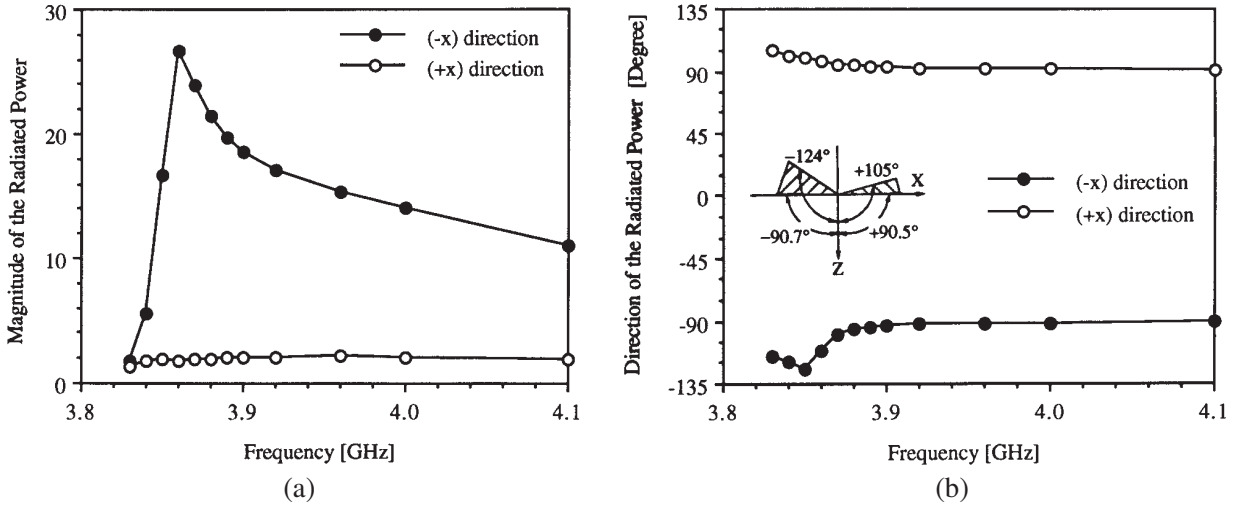


Figure 3: Radiated power carried by the magnetic surface wave (a) Its magnitude and (b) its direction. Plus and minus angles are defined from the $+z$ axis toward the plus and minus x axes, respectively.

3. CONCLUSION

The edge condition of a microstrip line having ferrite film in the frequency range where magnetostatic surface waves are excited is derived using the Mexiner's method. It corresponds with the calculated variation of the surface current distribution. The Jacobi polynomial is chosen as the basis function to satisfy the obtained edge condition. Because the Jacobi polynomial may give the different conditions for each edge, that is an appropriate current basis function. Thus the current basis function having the Jacobi polynomial may be expected to give fast convergence of numerical solution.

ACKNOWLEDGMENT

This work was supported by grant No. R05-2003-000-11957-0 from the Korea Science & Engineering Foundation.

REFERENCES

1. Ganguly, A. K., D. C. Webb, and C. Banks, "Complex radiation impedance of microstrip-excited magnetostatic-surface waves," *IEEE Trans. Microwave Theory Tech.*, Vol. MTT-20, 444–447, June 1978.
2. Lee, J. H. and J. W. Ra, "Full-wave calculation of the radiation impedance of microstrip-excited magnetic surface waves," *Microwave Opt. Technol. Lett.*, Vol. 6, No. 7, 441–444, June 1993.
3. Freire, M. J., R. Marques, and F. Medina, "Full-wave analysis of the excitation of magnetostatic-surface waves by a semi-infinite microstrip transducer-theory and experiment," *IEEE Trans. Microwave Theory Tech.*, Vol. MTT-51, No. 3, 903–907, Mar. 2003.
4. Mittra, R. and S. W. Lee, *Analytical Techniques in the Theory of Guided Waves*, Macmillan Company, New York, 1971.
5. Lee, J. H., "Edge condition of microstrip line used for magnetostatic surface wave transducers," *The Journal of Korea Electromagnetic Engineering Society*, Vol. 15, No. 11, 1062–1068, Nov. 2004 (in Korean).

Analyses of Multimode Forming Process in a Microwave-heating Cavity

Jirun Luo
Institute of Electronics, Chinese Academy of Science, Beijing 100080, China

Abstract— A definition of electromagnetic mode, based on its propagation direction and the distance from the cavity axis with cutting angle θ_B and caustic radius R_c, is given and used to describe multimode forming process in a circular cylindrical cavity for explaining its effect on the quality of field homogeneity. And the effect of polygonal cylindrical cavity as well as the properties, size and shape of the material sample processed in the cavity on multimode forming process is also discussed.

1. INTRODUCTION

Microwave heating of materials has been developing for over 40 years. The researches related to low processing temperature have led to the emergence of a lot of industrial facilities with microwave power sources at frequencies $\leq$ 2.45 GHz [1]. Many systems for microwave heating have been developed to improve the homogeneity of microwave field and temperature distribution, enhance the interaction efficiency of wave with materials, and pursue the industrial applications. The experimental methods vary from the coaxial lines and waveguide techniques used for decimeter wavelength to open quasi-optical cavities and interferometers exploited in millimeter wave range. The experimental activities [2–7] have been extended to microwave-assisted chemical synthesis, the surface processing of some specialized materials, and the sintering of structural and functional ceramics, composites, semiconductor materials, and even some conductive materials. A lot of the research results [2, 3] have shown good industrial perspective of microwave processing of materials from a technical viewpoint.

In the early days, the researches were mainly done in a singe mode resonant cavity made of rectangular waveguide at the frequency of 2.45 GHz or 915 MHz. Usually, Q factor of this cavity is high, which results in that microwave energy concentrates on a small area, so the field around the heated sample is not uniform. Although the rotation and moving of the sample can appropriately improve the heating homogeneity with time average, the method can only used for the experimental study of small size samples. To improve the heating homogeneity in a larger area for satisfying the requirement of industrial applications, scientists enhance the cavity size. For a specific operating frequency, the increase of the cavity size means that more modes can exist in the cavity. So mode stirrer and some special geometrical transition structures have been used to excite these modes in the cavity through changing the wave vector direction of input microwave to promote the generation of different modes. To get sufficient homogeneity of the field and temperature distribution within samples in a multimode cavity, Patterson et al. [8] demand the characteristic size of the oven of about 12 meter if a frequency of 2.45 GHz is used. This is not only a too large size for manufacture, but also needs too high power for keeping enough microwave power density to heat the samples. So too large and complicated microwave system becomes a barrier for most industrial application. With the development of high power gyrotron, millimeter wave radiation has been utilized to improve the homogeneity of the field and temperature distribution inside the heated samples. Because the wavelength of millimeter wave is very short, the same homogeneity of the field and temperature distribution can be realized in a much smaller cavity, compared with that at the frequency of 2.45 GHz. In addition, the rotation and moving of the samples can also be used in the multimode cavity for further improving the homogeneity of the field and temperature distribution inside the heated samples with time average.

Based on above-mentioned description, increasing frequency, enlarging cavity size, rotating and moving the samples, and exciting multimode in the cavity can all improve the homogeneity of the field and temperature distribution. Except for the rotation and moving of the samples, the other three kinds of methods are mutually correlative. The increase of frequency is dependent on the state of production and research of microwave power source. The enlargement of cavity size is determined by mechanical technology. The multimode excitation is affected by the frequency of microwave source, the size and structure of heating cavity, the design of mode stirrer, the wave vector direction of input wave beam, and so on. So the excitation of multimode is much more

important than the other two kinds of methods for improving the homogeneity. One knows that, the higher the frequency, the more modes there are in a size-specified cavity, and the larger the size of a cavity, the more modes there are in the cavity for a specified frequency. Obviously, the effects of the design of heating cavity and mode stirrer as well as the wave vector direction of input wave on the multimode generation are very complicated and important from the point of technical development. In this paper, we will discuss how above-mentioned factors affect multimode generation in detail from structure design of cavity and mode stirrer, methods of mode excitation, and variation of operating frequency in order to better understand multimode forming process.

2. A DEFINITION OF ELECTROMAGNETIC MODE

In order to guide microwave along specific direction or way or excite them in a specific chamber, scientists have designed waveguide and cavity. For the waveguide and cavity with metal wall boundaries, only there are some special electromagnetic field structures. Each field structure can satisfy the boundary condition and all these field structures constitute an orthogonality series, which means that each field structure can exists in the waveguide or cavity either singly or superposition with any other mode at same time. Each field structure possesses specific propagation or resonance characteristics in waveguide or cavity respectively. This field structure is called as 'mode'. When one hopes that microwave do not distorted after it propagates through a waveguide or can be excited as a monochromatic wave in a cavity, he will design the structure and electric parameters of the system to assure single mode operation. If one wishes that microwave field could distribute as evenly as possible in microwave materials processing system, he will design the structure and electric parameters of the system to excite modes as many as possible and change every mode's amplitude and the power density absorbed by the materials for lowering the effect of the operating mode of microwave source on the heating homogeneity. Our discussion will be concentrated on the multimode excitation to improve the homogeneity of field and temperature distribution for microwave heating in this paper.

In a microwave heating system with metallic boundaries, multimode operation means a characteristic size of the heating cavity much bigger than the wavelength of the microwave source, and we can think that the excited modes propagate in the cavity according to geometrical optical ray [9]. The initial TE or TM mode is represented as superpositions of plane waves by decomposition of the Bessel function in a cylindrical cavity, each propagating at the Brillouin angle θ_B relative to the cavity axis:

$$\theta_B = \arctan(X_{mn}/k_z a) \tag{1}$$

where X_{mn} is root of Bessel function (or derivative), $k_z = \sqrt{\left(\frac{\omega}{c}\right)^2 - \left(\frac{X_{mn}}{a}\right)^2}$ is axial wave number, a is radius of the cavity. Although θ_B shows the cutting angle of propagating direction of a ray with cavity axis according to geometrical optics, the minimum distance from the ray to the axis of the cavity can be changed, which means it is difficult to only use θ_B for determining the propagation status of a mode. The requirement of a zero azimuthal electric field at the cavity wall defines the relative phase of these plane waves. In the geometrical optical limit, the transverse location of a plane wavefront can be defined based on that at a particular point of interest the ray direction must coincide with the direction of Poynting vector of the original TE- or TM-mode field distribution. If the point of interest is located at the cavity wall the ray has the distance [10]

$$R_c = \frac{m}{X_{mn}} a \tag{2}$$

from the cavity axis. Hence, if all plane waves are represented by geometrical optical rays, they form a caustic at the radius R_c. So, in the geometrical optical limit, the propagation status of a mode can be determined by Brillouin angle θ_B and caustic with formulae (1) and (2). [11] Fig. 1 shows a geometrical description of the rays propagating in a multimode cylindrical cavity.

Formulae (1) and (2) show the relation of a mode with its propagating direction. In this paper, we use the formulae (1) and (2) as a definition of 'mode' for better understanding the multimode forming process in microwave heating cavity. If one knows Brillouin angle θ_B and caustic radius R_c of a ray in a cylindrical heating cavity, he can determine what mode forms the ray, Vice versa. Based on the formulae (1) and (2), we can not only conveniently study the propagation of a mode in a multimode cavity, but also utilize them to excite the expected modes by designing the propagation direction of a wave beam. Both (1) and (2) are two very important formulae for studying multimode

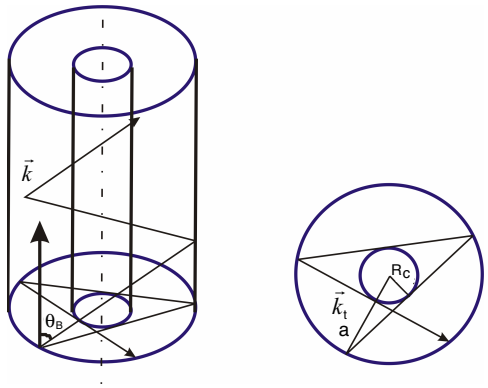


Figure 1: Geometrical description of the ray propagating in a multimode cylindrical cavity.

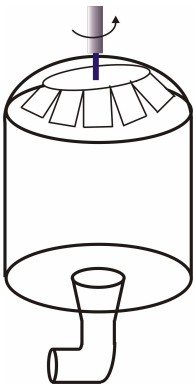


Figure 2: Schematic of multimode excitation by deformation of input waveguide.

excitation to obtain the homogeneity of field and temperature distribution in a microwave heating cavity.

3. MULTIMODE FORMING PROCESS

Multimode Formation is a very important process for improving the homogeneity of electromagnetic field and temperature distribution in a microwave-heating cavity. Only appropriate excitation and controlling for expected modes' amplitude and phase can satisfy the requirements of the homogeneity of the system for microwave heating. There are many methods for multimode generation, which can take place under the process of the coupling and excitation of modes. For example, the tapering transition and bending of input waveguide, the change of the propagation direction for wave beam and the dispersion caused by the nonlinear interaction of microwaves with heated material. In this section, we will discuss the multimode forming process in detail based on the above-mentioned three methods.

3.1. Multimode Excitation through Deformation of Input Waveguide

When microwave propagates in a crossed-section-changed or -bent waveguide, different modes will couple each other. This kind of coupling obeys by the generalized telegraphist's equation:

$$\frac{\mathrm{d}A_{mn}}{\mathrm{d}z} = -i\beta_{mn}A_{mn} + \sum_{m'}\sum_{n'} A_{m'n'}K_{mn}^{m'n'} \tag{3}$$

where the A_{mn} and $A_{m'n'}$ are the amplitudes of the various modes, β_{mn} is the propagation constant of the mode having amplitude A_{mn}, and $K_{mn}^{m'n'}$ is the coupling coefficient between the mn and $m'n'$ mode, and the sum over $m'n'$ excludes mn.

According to (3), if one hopes that the azimuthal mode number is not changed in the process of wave coupling, pure radial variations of waveguide can be used for realizing this mode conversion. However, if one hopes to change the azimuthal mode number, waveguide's bending (or a curvature variation) is required. Usually, conversion component directly connects to the microwave source and load in this kind of conversion process, reflection is a main problem affecting the steady operation of the system. In order to avoid that energy from microwave source faces too strong reflection before transmitting into the heating cavity, the variation of waveguide crossed section or curvature should be tapering. But the tapering transition is not sufficient to obtain a uniform field and temperature distribution, and mode stirring is also a key technology for solving the homogeneity problem in this system. However, appropriate reflection may cause variation of the operating frequency of the source in a small range when a mode stirrer has been used in the heating cavity, which will also promote the more multimode generation and the improvement of field homogeneity. The realization of this method must depend on the base of the enough reflection for multimode generation, the decrement of heating conversion efficiency because of the reflection and the steady operation of the system. And it is difficult to alleviate the effect of field concentrations in the center with this method. Fig. 2 is a schematic of the possible structure for this heating cavity.

3.2. Multimode Formation through Changing the Propagation Direction of Wave Beam

3.2.1. Modes Excited with a Gauss Wave Beam

A more effective method for improving field homogeneity and relieving the effect of the reflection from the interaction system is to use a Gauss wave beam as input microwave energy. If a Gauss wave beam is input to a multimode cylindrical heating cavity, as shown in Fig. 3, the excited modes in the cavity by the beam can be determined according to θ_B and R_c based on formulae (1) and (2). At first, we consider a simple situation, in which the ray OB in the wave beam center is in the same plane with a plane including the cavity axis. So the ray OB, those rays parallel to OB and also included in the plane will excite the same circular electric (or magnetic) mode, supposed as the TE_{0n} (or TM_{0n}). The rays not parallel to OB and also included in the plane will excite the TE_{0m} (or TM_{0m}, $m > n$) modes for BA area and the TE_{0k} (or TM_{0k}, $k < n$) modes for BC area respectively. We can think that the rays not included in this plane all lay in a series of the planes parallel to it, which means that the TE_{pn} (or TM_{pn}), TE_{pm} (or TM_{pm} $m > n$) and TE_{pk} (or TM_{pk} $k < n$) modes (p does not equal to zero and is different for the different planes) will be excited in the cavity. In addition, if the wave beam does not consist of a series of discrete rays, the excited modes will include EH or HE modes even some other modes, which may not be described with exact mathematic form. However, the existence of these modes can well be understood through (1) and (2) from a physical point of view. For a specific cavity structure, the number, amplitude and phase of the modes excited by a Gauss wave beam is only dependent on the cutting angle θ_B and caustic radius R_c of the rays relative to the cavity axis. If input wave beam does not intersect with the cavity axis, the number of the excited modes in the cavity can be same, but some different kinds of modes will be excited, which will affect the field distribution homogeneity in the cavity. In principle, that input wave beam does not intersect with the cavity axis can excite more high order modes, which makes field distribution more uniform by relieving the field concentrations in center. Although the above-mentioned description has shown that a Gauss wave beam can excite a lot of modes in the multimode cavity, it is not sure that the field distribution in the cavity is enough homogeneity for materials processing because much more energy may only be stored in few modes.

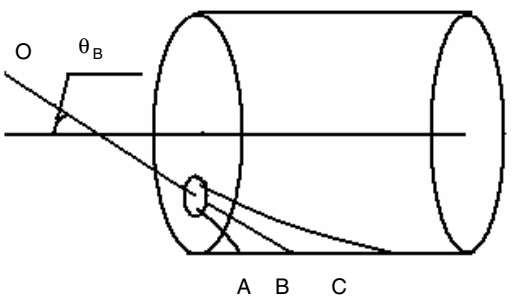


Figure 3: Model for multimode excitation with a Gauss wave beam.

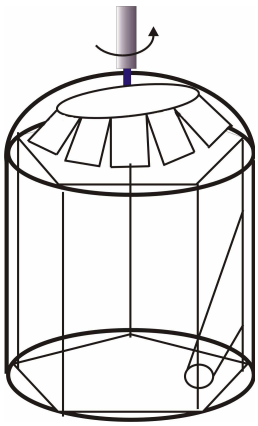


Figure 4: A schematic of polygonal applicator.

3.2.2. Effect of Mode Stirrer

In order to further improve the field distribution homogeneity, mode-stirring technology is utilized to excite more modes through changing the propagation direction of wave beam. Usually, mode stirrer (similar to Fig. 2), installed on the top of the cavity, is similar to a fan and can rotate in 3D space. The effects of mode stirrer on field homogeneity are mainly from two aspects. Firstly, the direction and special shape of its blades as well as its rotation in 3D space will change the reflection direction of the rays of wave beam and may increase the reflection times in the cavity, which excites more modes and may cause time-averaged homogeneity in the cavity. Secondly, periodical motion of the stirrer causes a time-averaged homogeneity of electromagnetic field.

One can design the shape, direction and rotating way of the blades to enlarge or reduce the cutting angle θ_B and caustic radius R_c of a ray relative to the cavity axis through the reflection of the ray on the blades. After the reflection of the ray, the decrease of θ_B means that a mode with

lower radial wave number is excited, and the decrease of R_c means a mode with lower azimuthal wave number is generated. The increment of θ_B and R_c will promote the generation of higher order modes with different radial and azimuthal wave number respectively. If θ_B or R_c of the ray decreases, the time spacing reflected again by the mode stirrer on the ray shortens, which makes it may have more opportunities for exciting other modes so that more uniform field distribution can be formed in the cavity. If θ_B or R_c of the ray increases, the time spacing reflected again by the mode stirrer on the ray will become longer, so the reflection times of the ray on the cavity wall will rises, which may also improve field distribution homogeneity in the cavity by random superposition of phases and amplitudes of the field for the mode. In order to avoid that electromagnetic energy is mainly included few modes, one should excite various different modes in the cavity as many as possible. However, the excitation of high order modes is not as easy as the excitation of lower order modes for a specific ray through a mode stirrer. For example, in a cylindrical cavity (Fig. 3), only when the cutting angle of bottom or top cover with cavity axis is changed to deviate 90°, the ray with θ_B can excite the other mode after it is reflected on the cover. If a mode stirrer is installed on the top of the cavity (Fig. 2), it is difficult to form a cutting angle of larger than 90° between the inner side of the blade of the mode stirrer and the cavity axis even if the stirrer can rotate in 3D space, which may cause field concentrations in the center (near the cavity axis). Although one can do some holes on the blades and design the special shape of the blades and the spacing between two neighboring blades, which makes that diffraction can happens, to try to excite the higher order modes in the cavity, it is difficult to describe the forming process of the modes based on mathematics and physics because of too complex structure of the stirrer, and the improvement for field homogeneity is insufficient for satisfying industrial quality requirements of field homogeneity [11].

3.2.3. Effect of the Cavity Structure Variation

In order to avoid field concentrations in the center, one of effective methods is to make the input angle of the ray of wave beam on the cavity wall become larger, which can make the reflection times of the ray increase and higher order modes excited. However, not every way enlarging the input angle can all improve field homogeneity well because the homogeneity is the result of superposition for the amplitudes and phases of all modes in the cavity. Changing the boundary of the cavity from circular cylindrical surface to polygonal cylindrical one can enlarge the input angle of the ray when the wave beam is input by hitting an edge of the polygonal applicator (Fig. 4), which shows a new opportunity for further improving the quality of the field homogeneity. Because of too complex boundary condition for this new applicator structure, precise knowledge of cavity properties and design studies have to be obtained by 3D numerical field calculation. Although it is not easy to define the cutting angle θ_B and caustic radius R_c of a mode in the polygonal cylindrical cavity according to the formulae (1) and (2), the cutting angle θ_B and the distance between a ray and cavity axis (similar to caustic radius) R_c can still be used to describe concept of 'mode' for input wave beam. A simulation code MiRa [11] had been developed at Forschungszentrum Karlsruhe for studying how to improve field homogeneity in a microwave heating system through designing cavity configuration. As a quantitative measure for the quality of the field within a given volume, the design parameter Δ_d in the code was defined as

$$\Delta_d = \frac{\sigma_{|\vec{E}|^2}}{\langle |\vec{E}|^2 \rangle} \tag{4}$$

where $\sigma_{|\vec{E}|^2}$ is the standard deviation of the electric energy density in the given volume, $\langle |\vec{E}|^2 \rangle$ is the spatial averaged energy density in a considered area. So improving the quality of the field homogeneity at a high field level can be realized by a decreasing value of Δ_d from the design of an appropriate applicator configuration.

Figure 5 shows the numerical simulation results on the field quality of polygonal applicator design study for a) beam hitting the bare surface and b) beam hitting an edge based on MiRa code. It can be seen that, if the beam is hitting the wall of the applicator at the first reflection, the best quality of field homogeneity can be reached for the applicator configuration of "circular cylindrical geometry +stirrer". However, if the beam is hitting an edge of a polygonal applicator at the first reflection, except pentagonal applicator, the qualities of field homogeneity are all better than those of that beam hits the wall of the applicators at the first reflection. A significant minimum for Δ_d can be detected for a hexagonal shape. The experimental results performed at Forschungszentrum Karlsruhe are in good agreement with the simulations based on MiRa code [11].

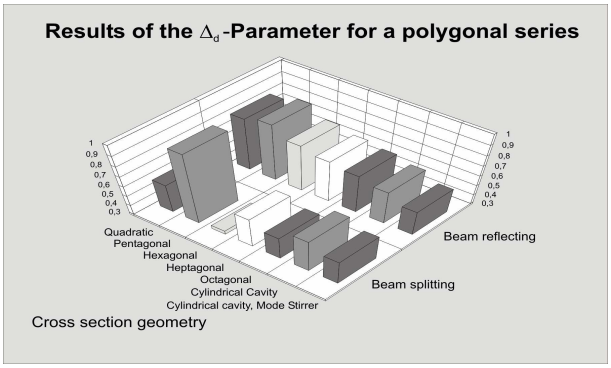


Figure 5: Results on the polagonal design study.

3.3. Effect of Interaction Process of Materials with Wave Beam on Multimode Formation

During the heating process of material samples, the physical, chemical and mechanical properties of the sample may change. So the absorption and reflection of the samples on microwaves will continuously vary because of the nonlinear interaction of microwaves with them, which may change the amplitude, phase and direction of reflected wave, excite new modes and affect the quality of field homogeneity. Usually, the larger the variation of the properties of the sample in the heating process, the larger the effect of the interaction process on the excited modes in the cavity. Because different materials have different nonlinear interaction process with wave beam, the quality of field homogeneity in a specific multimode cavity will depend on what material is processed in it. The reflection from the sample is dependent on the dielectric, conductive and loss properties of the sample as well as its shape and size in the heating process. The phase constant and the attenuation coefficient of a plane wave in the sample can be expressed as [12]:

$$\beta = \omega\sqrt{\mu\varepsilon'}\left\{\frac{1}{2}\left[\sqrt{1+\frac{(\varepsilon''+\sigma/\omega)^2}{\varepsilon'^2}}+1\right]\right\}^{1/2} \tag{5}$$

$$\alpha = \omega\sqrt{\mu\varepsilon'}\left\{\frac{1}{2}\left[\sqrt{1+\frac{(\varepsilon''+\sigma/\omega)^2}{\varepsilon'^2}}-1\right]\right\}^{1/2} \tag{6}$$

Both (4) and (5) show the relation of the phase constant and the attenuation coefficient with angular frequency ω, permeability μ, electric conductivity σ, and real ε' and image ε'' of complex permittivity. It is very clear from (4) and (5) that the propagation and attenuation of the plane wave in the sample are dependent on its electric and magnetic properties as well as microwave source frequency. When the heated sample is nonconductive material, the effect of it on the quality of field homogeneity in the heating process are dependent on the variation of the relative permittivity and loss tangent of the sample. However, for a perfect conductor as heated sample, the skin depth $1/\alpha$ is zero, microwave can only propagate on the surface of the sample and no loss of microwave energy happens inside the sample, which causes the change of the structures and distributions of mode fields in the heating system. In addition, The shrinkage of sample size and the variation of its shape in the heating process will cause the changes of reflection directions of the ray on the sample at different time, which makes that different modes are excited to affect the quality of field homogeneity. It can be seen that the quality of field homogeneity in a microwave-heating cavity is the result of a complicated combination, which includes how the properties, size and shape of the heated materials as well as their variations, the structure design of heating cavity and mode stirrer, and the input way of microwave mutually affect the mode excitations and the reflection process of wave.

4. SUMMARY

In this paper, a definition of electromagnetic mode based on Brillouin angle θ_B and caustic radius R_c has been proposed and applied to analyze the multimode forming process and the effect of it on field homogeneity in microwave-heating cavity. The analyses show that a Gauss wave beam can excite a lot of modes in the multimode cavity and the deviation of the beam related to the

cavity axis can alleviate the field concentrations in center, but microwave energy may only be stored in few modes. Mode-stirring technology can well solve the problem through designing the shape, direction, rotating way and periodical motion of the stirrer to change propagation direction and increase reflection times of the rays for exciting more modes and causing time-averaged homogeneity of field in the cavity. However, there exists difficulty to design the shape and rotating way of the mode stirrer for changing the modes to higher order modes than those formed by the wave beam at the first reflection, which may cause field concentration in center. Changing the configuration of the cavity from circular cylindrical surface to polygonal cylindrical one can enlarge the input angle of the ray when the wave beam is input by hitting an edge of the polygonal applicator, which provides an opportunity to excite higher order modes for alleviating field concentration in center. The simulations with MiRa code developed at FZK have shown that the best quality of field homogeneity can be reached for a hexagonal shape cavity, which is in good agreement with experimental results. In addition, the multimode formation is also affected by the properties, size and shape of the heated materials as well as their variations in the heating process. The homogeneity of field and temperature distributions is determined by combination of the properties and structure of the heated materials as well as their variations, the structure design of heating cavity and mode stirrer, with the input way of microwave.

ACKNOWLEDGMENT

Author would like to acknowledge Prof. M. Thumm, Dr. L. Feher and G. Link for helpful discussion and simulation results.

REFERENCES

1. Schiffmann, R. F., *Microwave: Theory and Application in Material processing III (Ceramic Transactions 59)*, D. Clark et al. (Westerville, OH: The American Ceramic Society), 7–16, 1995.
2. Bykov, Y. V., et al., "High-temperature microwave processing of materials," *J. Phys. D: Appl: Phys.*, Vol. 34, R55-R75, 2001.
3. Clark, D. E., et al., "Microwave processing of materials," *Annu. Rev. Mater. Sci.*, Vol. 26, 299–331, 1996.
4. Zhang, S.-L, et al., "Rapid thermal processing with microwave heating," *Thin Solid Film*, Vol. 246, 151–157, 1994.
5. Roy, R., et al., "Microwave processing: triumph of applications-driven science in WC-composites and ferric tianates," *Cera, Trans.*, Vol. 80, 3–26, 1997.
6. Cherradi, G., et al., "Electric and magnetic field contribution to the microwave sintering of ceramics," *Electrocramics IV*, (eds. Wasner, R. et al), Vol. 2, 1219–1224, RWTN, Aachen, 1994.
7. Roy, R., et al., "Full sintering of powder-metal bodies in a microwave field," *Nature*, Vol. 399, 668, June 17, 1999.
8. Patterson, et al., WO92/1343, PCT
9. Thumm, M. K., et al., "Passive high-power microwave components," *IEEE Trans. on Plasma Science*, Vol. 30, 755, June 2002.
10. Weinstein, L. A., *Open Oscillators and Open Waveguides*, Golem, Press, Boulder, Colorado, 1969.
11. Feher, L., G. Link, and M. Thumm, "Optimized design of an industrial millimeter wave applicator for homogeneous processing of ceramic charges", *Proc. Conf. Microwave and High Frequency Heating 1997*, A. Breccia, R. De Leo, and A. C. Metaxas, Eds., 442–446, Bologna, Italy, 1997.
12. Zhang, K. and D. Li, "Electromagnetic theory for microwaves and optoelectronics," *Springer-Verlag Berlin Heidelberg*, 1998.

Continuous Time Model Predictive Control for a Magnetic Bearing System

Jianming Huang
College of Automation, Chongqing University, Chongqing, China

Liuping Wang and Yang Huang
School of Electrical and Computer Engineering, RMIT University, Melbourne, Australia

Abstract— The nature of active magnetic bearings has many advantages over the conventional bearing, as its operation is energy efficient and potentially leads to cleaner and noise-free environment. However, the successful operation of an active magnetic bearing system requires a complex real-time control system, because of its unstable characteristics, as well as its nature of being a multi-input and multi-output system. This paper presents design and implementation of a continuous time model predictive control algorithm (CMPC) to an active magnetic bearing system (AMB). In this application, the plant continuous time model is identified from experimental data using prediction error method. The control performance of this algorithm is studied using an experimental AMB laboratory system. A host-target development environment of real-time digital control system with hardware in the loop (HIL) is implemented and demonstrated by controlling a nonlinear, open-loop unstable, and multivariable magnetic levitation device.

1. INTRODUCTION

This paper proposes a continuous time model predictive control scheme for an active magnetic bearing system. The AMB, an alternative to the conventional mechanical bearings, is an emerging technology with potential for application to a variety of products such as turbomolecular vacum pumps, flywheel energy storage systems and other high speed rotating machinery. Active magnetic bearing (AMB) systems have been widely applied due to their unique advantages such as: non-contact, the lubricant-free operation, high rotational speed, and the controllability of the bearing characteristics [1–3].

Firstly, the AMB rotor system always requires feedback control for stable levitation of the rotor. Therefore, the control system plays a central role in achieving high performance of the system. Secondly, the magnetic bearing system is strongly nonlinear due to electromagnetic forces, state and control constraints, and electrical characteristics of the power-supply actuator unit [4].

To apply active control to the AMB system, it is necessary to construct a high performance, real-time feedback controller to stabilize and control the suspended rotor. The control design problem is usually quite complicated. This has received much attention in recent years [5].

In this paper, we explore the model predictive control (MPC) algorithms in the application to AMB systems. Model predictive control has many advantages. It can handle constraints explicitly, can handle multivariable problems in a natural way and is model based design in the sense that it uses an explicit internal model to generate predictions of future plant behaviour [6]. However, most of the MPC techniques have focused on discrete time systems. The corresponding continuous time counterpart has received relatively little attention in the development. The continuous-time model predictive control scheme using orthonormal functions (Laguerre functions in our case) developed in [7] is computationally effective and easy to tune to achieve desirable closed-loop performance. These advantages are demonstrated in this application.

Due to fast dynamics of the AMB systems, the application of real-time control is always a challenging task. With the development of today's PC/DSP-based technology, advanced control methods are more often used in applied AMB control algorithms. In our application, the host-target real-time environment with the combination of real time workshop (RTW) and xPC target is implemented to develop and control an experimental AMB laboratory system [8, 9].

2. SYSTEM DESCRIPTION

2.1. MBC 500 Magnetic Bearing System

Figure 1 shows the configuration of the MBC 500 magnetic bearing system. The MBC 500 magnetic bearing system includes a stainless steel shaft or rotor, four magnetic bearing actuators with eight

'horseshoe' electromagnets to levitate the rotor. Hall effect displacement sensors are placed just outside the electromagnets at the end of the rotor. The electromagnets are driven by four current amplifiers which can be actuated by four internal or external controllers. The system has four-degree of freedom with two degrees of freedom at each end of the rotor. These two degrees of freedom are translation (x_1 and x_2) in the horizontal direction perpendicular to the z axis and translation (y_1 and y_2) in the vertical direction.

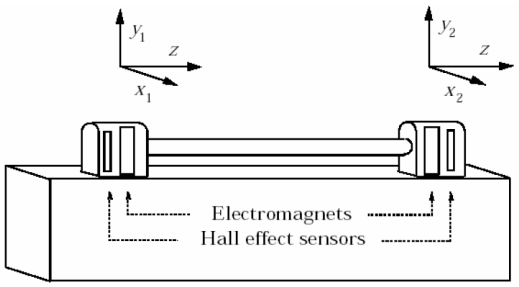


Figure 1: Configuration of MBC 500 magnetic bearing system.

There are four on-board analogue controllers provided for the MBC 500 magnetic bearing which levitate the bearing when connected in the feedback. On the front panel there are four switches for disconnecting each of the controllers so that any one or all of them can be replaced by the external controller [10].

The radial dynamics of the magnetic bearing system constitute a 4 input-4 output MIMO plant with motion in the x-z and y-z planes. The coupling of the motion in the two planes for the open loop plant is due to gyroscopic cross coupling and cross coupling effects from the bearings. In this paper, we assume that these effects be small. Therefore it is reasonable to control the entire system via two subsystems in different plane independently. In this paper, two axes in y-z plane are controlled by on-board controllers and two axes in x-z plane are controlled by a 2 input-2 output continuous time model predictive controller as proposed in this paper.

2.2. Host-target Real-time Control System Architecture

The hardware and software tools used in the development, implementation and testing of the digital control algorithm are discussed in this section. The overall host-target real-time control system architecture is shown in Figure 2, which is similar to the structure discussed in [11].

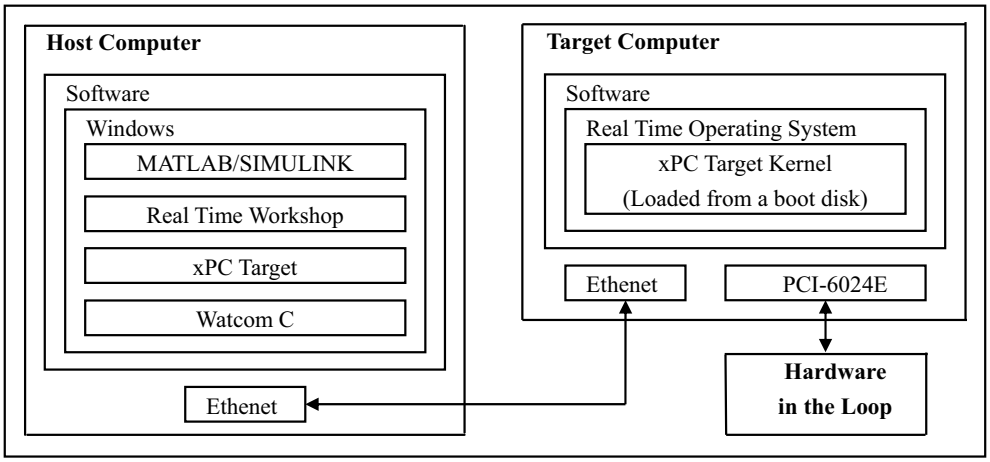


Figure 2: Host-target real-time control system architecture.

The host-target real-time environment is implemented using MathWorks tools. MATLAB is the basic "engine" with add-on components called toolboxes. SIMULINK is a MATLAB add-on that provides a graphical user interface for model development and system simulation. The RTW toolbox is capable of generating real time code for SIMULINK models. The xPC Target toolbox allows access to input/output data directly from a compatible data acquisition card and generates, compiles, and creates real-time executable code for SIMULINK models without the user having to write some low-level code. The real-time code is downloaded to the target computer to control a system in a HIL environment. The Watcom C compiler is used for generating C code.

The target computer is booted using an xPC boot floppy that loads the xPC target real-time kernel. Subsequently, the generated executable real-time code is downloaded to the target PC via the Ethernet. The I/O board and the PCI-6024E card by National Instruments provide the interface between the target computer and the HIL system.

2.3. Implementation

Experimental setup using host-target real time control system architecture to control the MBC 500 magnetic bearing system is shown in Figure 3. SIMULINK block diagram for xPC target to control the plant is shown in Figure 4. PCI-6024E and PCI-6024E1 are I/O blocks for National Instruments PCI-6024E I/O board with A/D channels and D/A channels. The plant outputs are sampled via analog input (A/D) channels and the control signals are generated via analog output (D/A) channels to control the plant. The Embedded MATLAB Function1 block receives the control changes and the plant outputs from which the observed states are generated. The Embedded MATLAB Function block implements the CMPC algorithm. The Embedded MATLAB Function2 block and the Embedded MATLAB Function3 block generate the reference inputs. The signal generator block and the signal generator1 block generate the disturbance inputs.

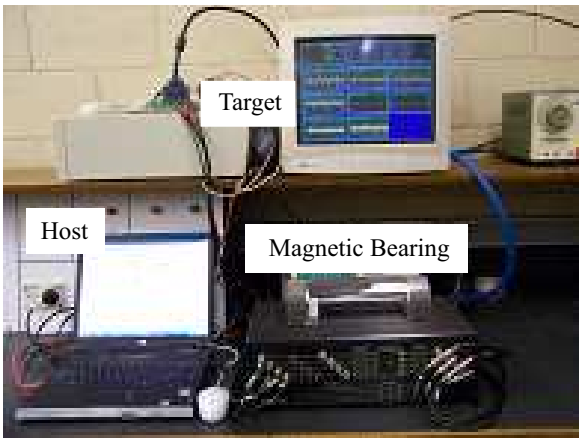


Figure 3: Experimental setup.

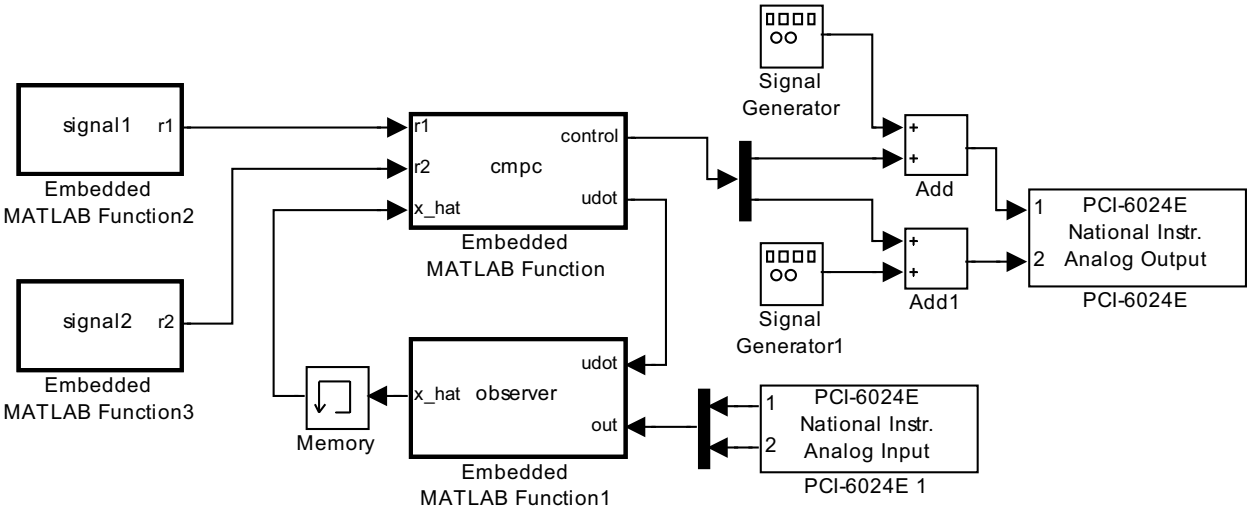


Figure 4: Simulink block diagram for magnetic bearing control system.

3. CONTINUOUS TIME MODEL PREDICTIVE CONTROL

This section provides a brief discussion of the continuous time model predictive control [7] used in this paper. The approach uses orthonormal functions to describe the trajectory of the control variable, and a multivariable state space model is used in the design.

With respect to a set of real functions $l_i(t)$, $i = 1, 2, \ldots$, that is orthogonal and complete over the interval $(0, \infty)$, it is known that an arbitrary function $f(t)$ is expressed in terms of a series expansion as

$$f(t) = \sum_{i=1}^{\infty} \xi_i l_i(t) \tag{1}$$

where ξ_i, $i = 1, 2, \ldots$, are the coefficients.

Specifically, we let $L(t) = [l_1(t) l_2(t) \dots l_N(t)]^{\mathrm{T}}$, where $l_i(t)$'s are Laguerre functions defined as

$$\begin{aligned} l_1(t) &= \sqrt{2p} \times e^{-pt} \\ l_2(t) &= \sqrt{2p}(-2pt+1)e^{-pt} \\ \vdots &= \vdots \\ l_i(t) &= \sqrt{2p}[(-1)^{i-1}\frac{(2p)^{i-1}}{(i-1)!}t^{i-1} + (-1)^i\frac{(i-1)(2p)^{i-2}}{(i-2)!}t^{i-2} \\ &\quad +(-1)^{i-1}\frac{(i-1)(i-2)(2p)^{i-3}}{2!(i-3)!}t^{i-3} + \cdots + 1]e^{-pt} \end{aligned}$$

where the parameter p is positive and often called scaling factor for Laguerre functions and $L(0) = \sqrt{2p}[1 \quad 1 \quad \cdots \quad 1]^{\mathrm{T}}$. Then we can describe the derivative of the control signal using Laguerre functions based series expansion as

$$\dot{u}(t) = \sum_{i=1}^{N} \xi_i l_i(t) = L(t)^{\mathrm{T}}\eta \tag{2}$$

where N is the number of terms used and $\eta = [\xi_1 \quad \xi_2 \quad \cdots \quad \xi_N]^{\mathrm{T}}$ is the vector of coefficients.

Suppose that the plant to be controlled is an r input - q output multivariable system having a state space model

$$\left.\begin{aligned} \dot{X}_m(t) &= A_m X_m(t) + B_m u(t) \\ y(t) &= C_m X_m(t) \end{aligned}\right\} \tag{3}$$

where $X_m(t)$ is the state vector of dimension n.

Let us now define an auxiliary variable

$$Z(t) = \dot{X}_m(t) \tag{4}$$

We then write an augmented state space description of the system (3) as

$$\left.\begin{aligned} \dot{X}(t) &= AX(t) + B\dot{u}(t) \\ y(t) &= CX(t) \end{aligned}\right\} \tag{5}$$

where $X(t) = \begin{bmatrix} Z(t)^{\mathrm{T}} & y(t)^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}}$, $A = \begin{bmatrix} A_m & 0 \\ C_m & 0 \end{bmatrix}$, $B = \begin{bmatrix} B_m^{\mathrm{T}} & 0 \end{bmatrix}^{\mathrm{T}}$, $C = [0 \quad I]$, I is the $q \times q$ unit matrix. Note that the augmented state space description (5) takes the first derivative of the control signal as its input and its output remains the same.

Suppose that in a multi-input and multi-output setting, a set of future setpoint $r(t_i + \tau) = [r_1(t_i + \tau) \quad r_2(t_i + \tau) \quad \cdots \quad r_q(t_i + \tau)]$, $0 \le \tau \le T_p$ are available, where T_p is the prediction horizon. The objective of model predictive control is to find the control law that will drive the predicted plant output $y(t_i + \tau)$ as close as possible, in a least squares sense, to the future trajectory of the setpoint $r(t_i + \tau)$. The cost function is taken as

$$J = \int_0^{T_p} [r(t_i+\tau) - y(t_i+\tau)]^{\mathrm{T}} Q [r(t_i+\tau) - y(t_i+\tau)] d\tau + \int_0^{T_p} \dot{u}(\tau)^{\mathrm{T}} R \dot{u}(\tau) d\tau \tag{6}$$

where Q and R are symmetric matrices with $Q > 0$ and $R \ge 0$.

Let the ith $(0 \le i \le r)$ set of Laguerre functions be $L_i(t)^{\mathrm{T}} = \begin{bmatrix} l_1^i(t) & l_2^i(t) & \cdots & l_{N_i}^i(t) \end{bmatrix}$, and define the convolution integral corresponding to the ith input

$$I_{int}(\tau)^i = \int_0^{\tau} \mathrm{e}^{A(\tau-\gamma)} B_i L_i(\gamma)^{\mathrm{T}} \mathrm{d}\gamma \tag{7}$$

To minimize J, in a least squares sense, the derivative of the optimal control for the unconstrained problem with finite horizon prediction is derived as

$$\dot{u}(t_i) = \begin{bmatrix} L_1(0)^{\mathrm{T}} & 0 & \dots & 0 \\ 0 & L_2(0)^{\mathrm{T}} & \dots & 0 \\ \vdots & & & \\ 0 & 0 & \dots & L_r(0)^{\mathrm{T}} \end{bmatrix} \Pi^{-1} \{\Psi_1 r(t_i) - \Psi_2 X(t_i)\} \tag{8}$$

where $\Pi = \int_0^{T_p} \phi(\tau) Q \phi(\tau)^{\mathrm{T}} d\tau + R$; $\Psi_1 = \int_0^{T_p} \phi(\tau) Q d\tau$; $\Psi_2 = \int_0^{T_p} \phi(\tau) Q C e^{A\tau} d\tau$;
$\phi(\tau)^{\mathrm{T}} = \left(C \left[I_{int}(\tau)^1 \quad I_{int}(\tau)^2 \quad \cdots \quad I_{int}(\tau)^r \right]\right)^{\mathrm{T}}$.

To integrate the $\dot{u}(t)$, we get

$$u(t) = \int_0^t \dot{u}(\tau) d\tau \tag{9}$$

The prediction of the future plant behaviour is built on the availability of the state variable at time t_i. The state variable is estimated by the observer equation that is needed for implementing the CMPC

$$\dot{\hat{X}}(t) = A\hat{X}(t) + B\dot{u}(t) + J_{ob}(y(t) - C\hat{X}(t)) \tag{10}$$

where $\dot{\hat{X}}(t)$ is the estimate of $X(t)$ and J_{ob} is the observer gain.

4. EXPERIMENTAL RESULTS

4.1. System Identification

The CMPC algorithm is a model based approach. The plant model has to be obtained first for the purpose of control design. Here the plant model is estimated using experimental data. Two sets of test signals that have white noise characteristics are applied to two reference inputs. The experimental data of inputs and outputs of the plant is shown in Figure 5. The MATLAB function 'pem' in System Identification Toolbox is used to identify the plant model in discrete-time transfer function form. The identified model is converted into the continuous-time transfer function matrix as follows:

$$\begin{bmatrix} y_1(s) \\ y_2(s) \end{bmatrix} = \begin{bmatrix} \frac{43.98(s-4241)(s+402.1)(s-299.2)}{(s-451.7)(s-252.9)(s+450.4)(s+359.4)} & \frac{-20.38(s-68.23)(s+217.2)(s+3739)}{(s-451.7)(s-252.9)(s+450.4)(s+359.4)} \\ \frac{-19.8(s+4110)(s+183.8)(s-92.44)}{(s-451.7)(s-252.9)(s+450.4)(s+359.4)} & \frac{55.72(s-3762)(s+387.3)(s-317.2)}{(s-451.7)(s-252.9)(s+450.4)(s+359.4)} \end{bmatrix} \times \begin{bmatrix} u_1(s) \\ u_2(s) \end{bmatrix} \tag{11}$$

Figure 6 gives the fit of measured outputs and 1-step ahead predicted model outputs, 98.01% and 94.67% respectively, and the prediction errors for axis x_1 and axis x_2.

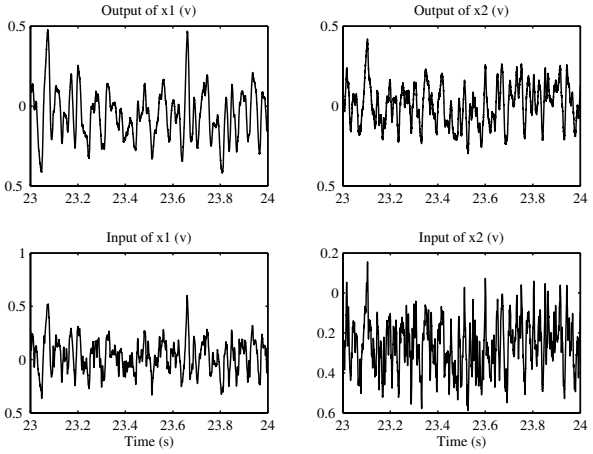


Figure 5: Experimental data of the plant.

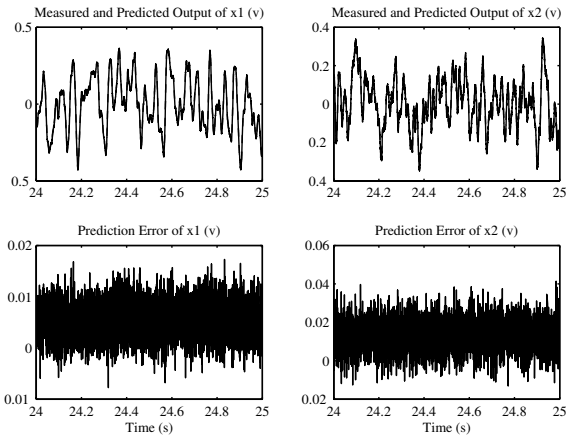


Figure 6: Validation result of the plant.

4.2. Simulation

The CMPC controller with an observer based on the identified plant model is designed. We select number of terms used in the Laguerre model to capture the control signal as $N_1 = N_2 = 2$. The prediction horizon is set to be 40 ms. Two tuning parameters for the closed-loop response speed are selected to be $p_1 = p_2 = 250$. The observer poles are selected close to the region of $-3 \times \max(p_1, p_2)$. The observer is designed using the pole assignment technique by putting the observer poles at [-750 -850 -950 -1050 -1150 -1250]. The simulated results are shown in Figures 7–10.

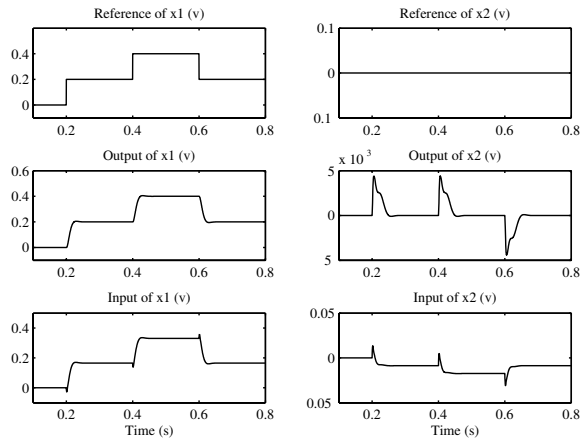


Figure 7: Staircase response of axis x_1.

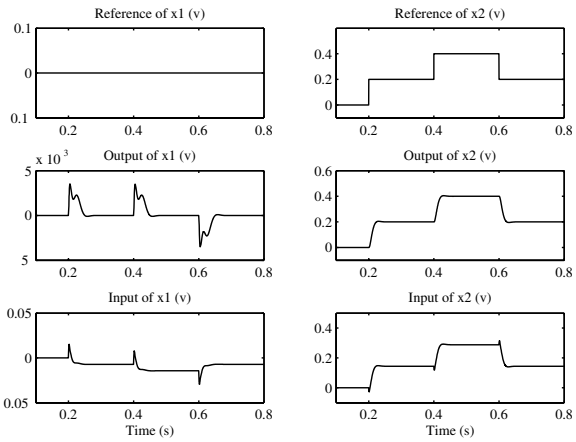


Figure 8: Staircase response of axis x_2.

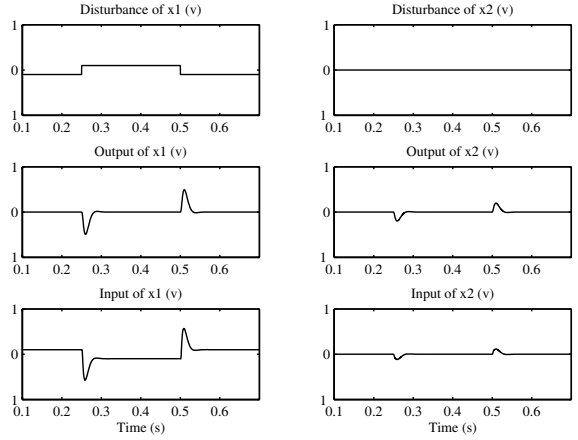


Figure 9: Disturbance response of axis x_1.

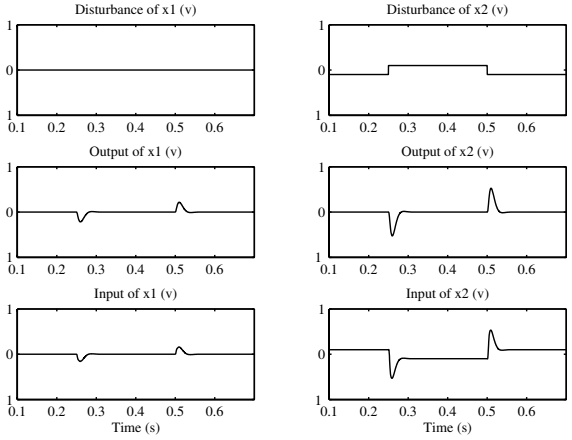


Figure 10: Disturbance response of axis x_2.

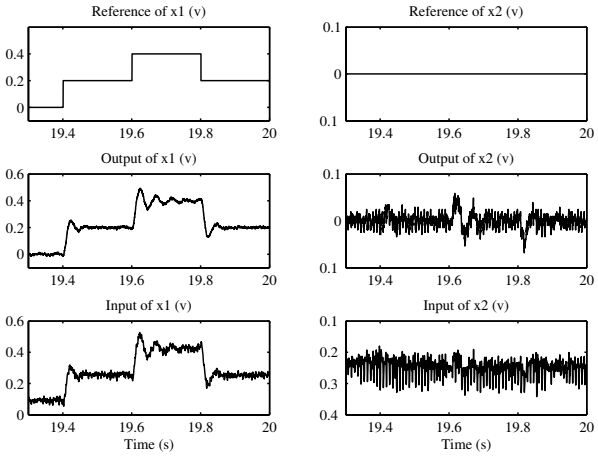


Figure 11: Staircase response of axis x_1.

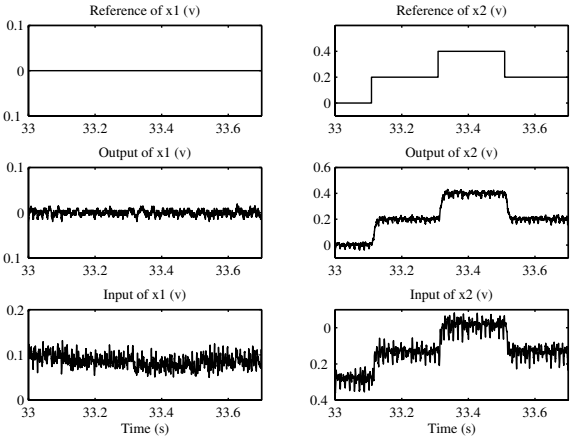


Figure 12: Staircase response of axis x_2.

4.3. Real-time Control

The CMPC algorithm is used to control the MBC 500 magnetic bearing system via the host-target real-time control system architecture shown in Figure 4. The design parameters are the same as the ones in Section 4.2. Sampling time is 0.2 ms. Experimental results are illustrated in Figures 11–14. Figure 11 shows the system output of axis x_1 (voltage equivalent of measured rotor translation in axis x_1 direction). Figure 12 shows the system output of axis x_2 (voltage equivalent of measured rotor translation in axis x_2 direction). Each of the system outputs tracks the corresponding staircase setpoint signal. The interaction is very small for the system output of zero-setpoint axis.

Figures 13 and 14 show the disturbance responses when the square disturbances were applied

to one of the plant inputs respectively. The system outputs converge to the equilibrium points rapidly.

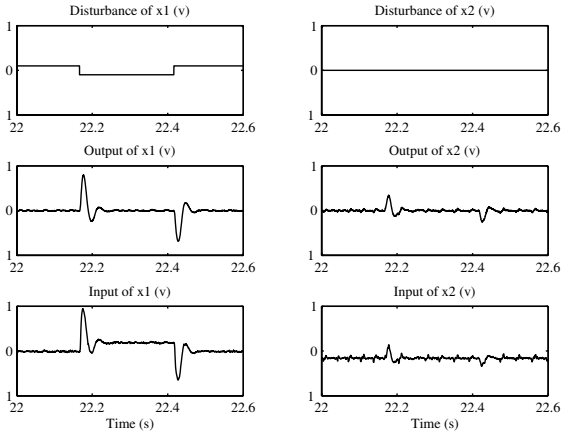


Figure 13: Disturbance response of axis x_1.

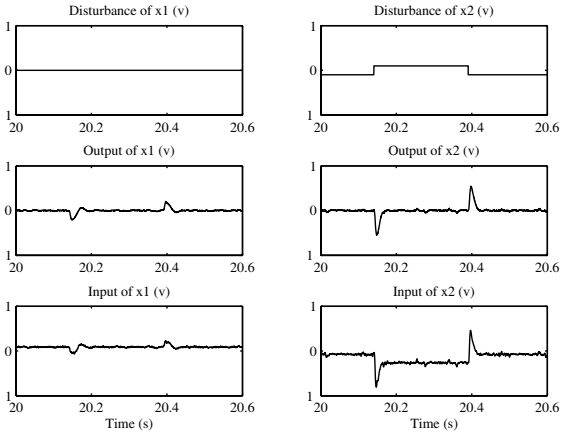


Figure 14: Disturbance response of axis x_2.

5. CONCLUSIONS

This paper has designed and implemented a continuous time model predictive control system on an active magnetic bearing system, which is unstable and in the framework of a multivariable system. The predictive control was designed using Laguerre functions to describe the future control trajectory. As a result, the closed-loop performance of the predictive control system can be tuned through the number of terms of the Laguerre functions in conjunction with the scaling factor. In the implementation of predictive control system, a host-target environment is used where the embedded MATLAB function blocks linking with the SIMULINK models provide a successful development platform for this application.

REFERENCES

1. Schweitzer, G., G. H. Bleuler, and A. Traxler, *Active Magnetic Bearings: Basic, Properties and Applications of Active Magnetic Bearings*, VDF Hochschulverlag, 1994.
2. Mohd-Mokhtar, R., L. P. Wang, L. J. Qin, and T. Barry, "Continuous time system identification of magnetic bearing systems using frequency response data," in *Proceedings of 5th Asian Control Conference*, Melbourne, Australia, 2066–2072, July 2004.
3. Srinivasan, S. and Y. M. Cho, "Modeling and system identification of active magnetic bearing system," in *Proceedings of the 4th IEEE conference on Control Application*, New York, USA, 252–260, September 1995.
4. Ahn, H. J., S. W. Lee, S. B. Lee, and D. C. Han, "Frequency domain control-relevant identification of MIMO AMB rigid rotor," *Automatica*, Vol. 39, No. 2, 299–307.
5. Grega, W. and A. Pilat, "Comparison of linear control methods for an AMB system," *International Journal of Applied Mathematics and Computer Science*, Vol. 15, No. 2, 245–255.
6. Maciejowski, J. M., *Predictive Control with Constraints*, Pearson education limited, 2002.
7. Wang, L. P., "Continuous time model predictive control using orthonormal functions," *International Journal of Control*, Vol. 74, No. 16, 1588–1600.
8. *Real Time Workshop User's Guide*, MathWorks, Inc., MA, 1999.
9. *xPC Target User's Guide*, MathWorks, Inc., MA, 1999.
10. *MBC 500 Magnetic Bearing System Operating Instructions*, Magnetic Moments, LLC, 2004.
11. Shiakolas, P. S. and D. Piyabongkarn, "Development of a real-time digital control system with a hardware-in-the-loop magnetic levitation device for reinforcement of controls education," *IEEE Transactions on Education*, Vol. 46, No. 1, 79–87.

Several Rules about the Magnetic Moment of Rotational Charged Bodies

Guo-Quan Zhou

Department of Physics, Wuhan University, Wuhan 430072, China

Abstract— A strict and delicate analogy relation between the magnetic moment of a rotational charged body and the rotation inertia of a rigid body about a fixed axis has been found in this paper. Based on this analogy relation, many rules and theorems about calculating the rotation inertia of a rigid body can be transplanted to field of electromagnetism and used to calculate the magnetic moment of a rotational charged body. Some principles or theorems are extended, generalized, illustrated, and enumerated. Some related examples are listed in the paper.

1. PREFACE

Reference [1] gives the analogy relation between the magnetic moment [3–5] of a rotational charged body and the rotation inertia of a rigid body about a fixed axis, which is very stimulating and pedagogical. According to this paper, a rotational charged body with an element electric charge of dq, a constant angular speed of ω (the size of $\vec{\omega}$) about a fixed axis, and a rotation radius of $r_\perp$, must have a definite magnetic moment of following magnitude:

$$\begin{aligned}\left|d\vec{P}_m\right| &= dI\pi r_\perp^2 \\ &= (\omega/2\pi)\cdot dq\cdot \pi r_\perp^2 \\ &= (\omega/2)\cdot r_\perp^2\, dq\end{aligned}$$

or

$$d(2P_m/\omega) = r_\perp^2\, dq$$

Its direction is parallel or anti-parallel to $\vec{\omega}$ according to the positive or negative sign of dq respectively. The above expression has a strict and delicate analogy relation to the rotation inertia expression — $dJ = r_\perp^2\, dm$ for an element mass of dm about a definite axis:

$$2P_m/\omega \leftrightarrow J,\ dq \leftrightarrow dm,\ q \leftrightarrow m$$

Let us pay more attention to the analogy relation of $2P_m/\omega \leftrightarrow J$, which is more appropriate than the corresponding one given in the reference [2]. Because only P_m/ω is the physical quantity that independent of ω and only impressed by the distribution of electric charge and position of axis. Thus there is a strict analogy relation between P_m/ω and J. Nevertheless, on the same time we should notice that, charged conductor which is in a state of a electrostatic equilibrium only has its charge distributed on its surface. So we must treat the rotational charged solid conductor as a charged shell, analogical to the hollow rigid body of the same shape. Other analogy relationship between the two circumstances is:

Q(the total charge value) $\leftrightarrow$ m(the total mass),
C_e(the center of elecric charge) $\leftrightarrow$ C_m(the mass center),
σ_e(electrical charge areal density) $\leftrightarrow$ σ_m(mass areal density),
ρ_e(electrical charge body density) $\leftrightarrow$ ρ_m(mass body density),
λ_e(electrical line density) $\leftrightarrow$ λ_m(mass line density).

It is obvious that many computation rules and property theorem of rotation inertia for a rigid body can be transplanted to that of the magnetic moment for a rotational charged body. Just as in reference [6], the rotation inertia for any triangle rigid plane with homogeneous mass distribution about its center-of-mass axis vertical to this rigid plane can be formulated as: $J_{ABC(C_m)} = m(a^2+b^2+c^2)/36$. According to the analogy relation given as above, we can immediately deduce the magnetic moment formula for a rotational charged triangle plate with homogeneous charge distribution about its center-of-charge axis vertical to the plate to be: $P_{mABC(C_e)} = Q(a^2+b^2+c^2)\omega/72$. This paper extends and applies the conclusions drawn from references [6–9] to the case for calculation of magnetic moment of a rotational charged body, tabulates and enumerates related examples of those theorems and makes some further discussion.

2. THE PARALLEL-AXIS THEOREM FOR THE MAGNETIC MOMENT OF A ROTATIONAL CHARGED BODY

The parallel-axis theorem [6–9] of rotation inertia for a rigid body is universal to all cases of any shape and any mass distribution. Then there is some reviews are in order: once we know the rotation inertia of rigid body about its center-of-mass axis, we can calculate its rotation inertia about any other parallel axis. Based on above analogy relation, first of all, we can define the electrical charge center $\vec{r}_c = \dfrac{1}{Q}\int \vec{r}dq$. And suppose the value of magnetic moment of a rotational charged body about one of its axis passing through its charge center is P_{mc} (the corresponding angular speed is ω), the distance between the two parallel axes is d, the total charge value is Q, then following conclusion holds:

$$\frac{2}{\omega}P_m = \frac{2}{\omega_c}P_{mc} + Qd^2 \tag{1}$$

In need of simplicity and conciseness, we can take $\omega = \omega_c$ (but notice that ω and ω_c are independent of each other), and attain a formula:

$$P_m = P_{mc} + \frac{\omega Q}{2}d^2 \tag{1'}$$

This is the parallel-axis theorem for the magnetic moment of a rotational charged body with any shape and any electric charge distribution (a volume, surface, curve or discrete points distribution). No matter what a kind of charged body it is, the above formula always holds — for a rigid body or a kind of liquid matter, a conductor or a kind of electrolysis. Its application examples can be found in Table 1.

There is a very impotent deduction for theorem (1′): if the total charge value of a rotational charged body is zero, then $P_m = P_{mc}$ and it has nothing to do with distance d.

That is to say: for a body with a total zero charge and a nonzero charge density, its magnetic moment about any axis with any direction is equal to that of a charged body about the center-of-charge axis.

3. THE ORTHOGONAL-AXES THEOREM FOR THE MAGNETIC MOMENT OF A ROTATIONAL CHARGED BODY

For a thin slice of rigid plate of zero thickness, its rotation inertia about the **x, y, z** axes respectively satisfies the so-called orthogonal-axes theorem [7, 9]. (provided the **xoy** coordinate plane is located upon the rigid plate):

$$J_z = J_x + J_y$$

Based on the analogy relation, for a thin plate with an arbitrary surface charge distribution, its magnetic moments about the **x, y, z** axes with an angular speed of ω_x, ω_y, ω_z respectively will satisfy the following orthogonal-axes theorem:

$$\frac{P_{mx}}{\omega_x} + \frac{P_{my}}{\omega_y} = \frac{P_{mz}}{\omega_z} \tag{2}$$

For simplicity and conciseness, without loss of generality, we can also assume $\omega_x = \omega_y = \omega_z$, and attain following equation:

$$P_{mx} + P_{my} = P_{mz} \tag{2'}$$

Its application examples can be found to be example (3) in Table 1. Now let us pay more attention to the so-called generalized orthogonal-axes theorem published in one of my papers [7]. For a column rigid body with an arbitrary shape of transversal section and of uniform mass distribution along its generatrix, with a height of L, a mass of m, with its Z axis parallel to its generatrix and the coordinate plane of **xoy** plumb to its generatrix line, amputating it to be two parts, each with a length of L_1 and L_2 respectively ($L_1 + L_2 = L$), as Fig. 1 shows. Then the following generalized orthogonal-axes theorem holds

$$J_x + J_y = J_z + \frac{2}{3}m\left(L_1^2 - L_1L_2 + L_2^2\right)$$

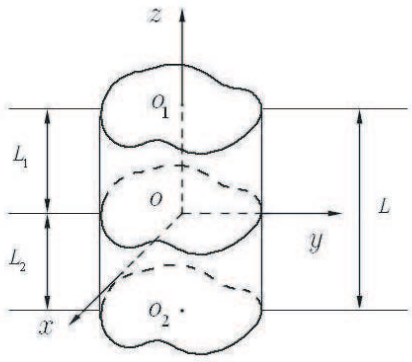


Figure 1: A column with height of L.

According to the analogy relation, for a uniform charged rotational column which has a total charge value of Q and angular speeds of ω_x, ω_y, ω_z about the **x, y, z** axes respectively, its corresponding magnetic moments must satisfy the following orthogonal-axes theorem when following substitution is taken: $\frac{2}{\omega}P_m \leftrightarrow J$, $Q \leftrightarrow m$.

$$P_{mx}/\omega_x + P_{my}/\omega_y = P_{mz}/\omega_z + \frac{1}{3}Q(L_1^2 - L_1L_2 + L_2^2) \tag{3}$$

For simplicity and conciseness, assuming that $\omega_x = \omega_y = \omega_z = \omega$ (notice that ω is not the resultant angular velocity of ω_x, ω_y, ω_z), then

$$P_{mx} + P_{my} = P_{mz} + \frac{\omega}{3}Q(L_1^2 - L_1L_2 + L_2^2) \tag{3'}$$

When $L_1 = 0$, $L_2 = L$, we have

$$P_{mx} + P_{my} = P_{mz} + \frac{\omega}{3}QL^2 \tag{3''}$$

(The end − face orthogonal − axes theorem)

When $L_1 = L_2 = \frac{L}{2}$, we have

$$P_{mx} + P_{my} = P_{mz} + \frac{\omega}{12}QL^2 \tag{3'''}$$

(The orthogonal − axes theorem on the middle transversal section)

We can get its application examples such as No. 4, No. 5, No. 6 in Table. 1 by means of analogy transition from reference [7].

What should be paid more attention to is that when the charged column has a total zero charge value and a nonzero charge density, the generalized orthogonal-axes theorem (3), (3′) reduce to the original orthogonal-axes theorem (2), (2′). That is to say, its magnetic moment is independent of the column length. It is very important and useful to calculate the magnetic moment of a rotational charged body, especially for a charged body with a four-degree symmetry axis (in terms of group theory). The examples Nos. 5 and 6 in Table 1 are its application.

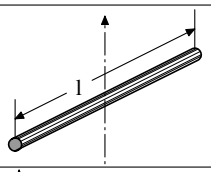
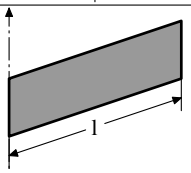
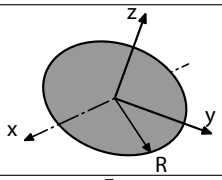
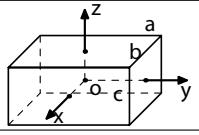
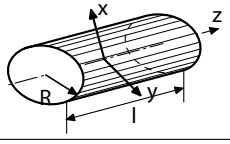
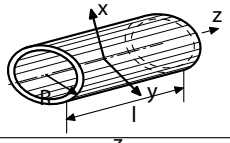
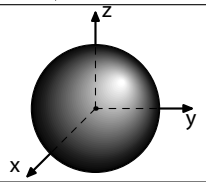
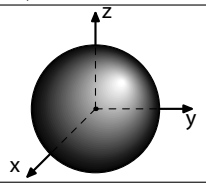
4. THE ORIGIN-MOMENT THEOREM

For those rigid bodies of arbitrary shape and mass distribution, randomly selected coordinate system *O-XYZ*, the rotation inertia J_x, J_y, J_z about the axes X, Y, Z satisfy the following equation

$$J_x + J_y + J_z = 2\int r^2 dm (= 2J_o)$$

Table.1. The table of calculation rules and examples of magnetic moment for rotational charged bodies.

(Provided all matter is evenly charged with total charge value of Q and with a constant volume (surface, linear) charge density)

Examples	**Form of charged bodies**	Rules and theorems	Formula of magnetic moment
No. 1.The thin and straight rod (linear)		The analogy relation and definition	$\vec{P}_m = \frac{1}{24}Ql^2\vec{\omega}$
No.2.The rectangular plate (surface distri.)		The parallel-axis theorem	$\vec{P}_{mx} = \frac{1}{6}Ql^2\vec{\omega}$
No. 3. The circular and thin plate (surface distri.)		The orthogonal-axes theorem	$\begin{cases}\vec{P}_m = \frac{1}{8}QR^2\vec{\omega}_x & (\vec{P}_{my} \text{ the similar}) \\ \vec{P}_{mz} = \frac{1}{4}QR^2\vec{\omega}_z\end{cases}$
No. 4. The cuboid (volume distri.)		The extended orthogonal-axes theorem	$\vec{P}_{mx} = \frac{1}{24}Q(a^2+c^2)\vec{\omega}_x$ $(\vec{P}_{my}, \vec{P}_{mz}$ the similar)
No. 5.The solid Cylinder (volume distri. and axial sym.)		The extended orthogonal axes theorem	$\vec{P}_{mx} = (\frac{1}{8}QR^2 + \frac{1}{24}QL^2)\vec{\omega}_x$ $\vec{P}_{my}$ *the similar*
No. 6.The thin and hollow cylinder (surface distri. and axial sym.)		The extended orthogonal-axes theorem	$\vec{P}_{mx} = (\frac{1}{4}QR^2 + \frac{1}{24}QL^2)\vec{\omega}_x$
No. 7.The solid globe (volume distri. and spherical sym.)		The origin-moment theorem	$\vec{P}_m = \frac{1}{5}QR^2\vec{\omega}$
No. 8.The thin Spherical crust (surface distri. and spherical sym)		The origin-moment theorem	$\vec{P}_m = \frac{1}{3}QR^2\vec{\omega}$

Here J_o is defined as the origin moment [8, 9] (also called the center moment). r is the distance between the mass element and the origin point. The above expression is called origin moment theorem or center moment theorem, which has been mentioned in many references [8, 9]. And its latest manifestation can be found in reference [8].

According to the simple analogy relation given before, we have

$$J_x \to \frac{2}{\omega_x}P_{mx}, \quad J_y \to \frac{2}{\omega_y}P_{mx}, \quad J_z \to \frac{2}{\omega_z}P_{mz} \text{ and } dm \to dq$$

Then we can get the origin-moment theorem related to moments of P_{mx}, P_{my}, P_{mz}

$$\frac{P_{mx}}{\omega_x} + \frac{P_{my}}{\omega_y} + \frac{P_{mz}}{\omega_z} = \int r^2 dq \tag{4}$$

For simplicity and conciseness, assuming that $\omega_x = \omega_y = \omega_z (= \omega)$.
Then

$$P_{mx} + P_{my} + P_{mz} = \omega \int r^2 dq \tag{4'}$$

Especially for those rotational charged bodies of spherical symmetry, their volume charge density and element charge can be expressed as $\rho_e = \rho_e(r)$, $dq = 4\pi r^2 \rho_e(r) dr$.
Then

$$P_{mx} + P_{my} + P_{mz} = 4\pi\omega \int r^4 \rho_e(r) dr$$

But due to the spherical symmetry $P_{mx} = P_{my} = P_{mz}$, we have

$$P_{mx} = P_{my} = P_{mz} = \frac{4}{3}\pi\omega \int r^4 \rho_e(r) dr \tag{4''}$$

Its application examples are enumerated in Table 1, or No. 7, No. 8.

5. CONCLUSIONS

Such a perfect and delicate analogy relation between the magnetic moment of a rotational charged body and the rotation inertia of a rigid body supply us with a powerful tool to calculate precisely the magnetic moments of rotational charged bodies with all kinds of shape. Formulas such as (1′)–(4′) can all be viewed as theorems to deal with all kinds of problem involved in magnetic moment calculation of rotational charged bodies, especially for those rotational bodies of symmetrical charge distribution. Table 1 gives some concrete illustrations of these formulas. Such calculation is very necessary and meaningful for research and teaching of the electromagnetic field and even for investigating of space such as the launching of satellites and spaceships.

ACKNOWLEDGMENT

Supported by National Science Foundation Grant. No. 10375041.

REFERENCES

1. Zhou, G.-Q., "Charge moment tensor and the magnetic moment of rotational charged bodies," *Progress in Electromagnetics Research*, Vol. 68, 156–160, New York, 2007.
2. Xu, D.-F and Z.-S. Yu, "The calculation of magnetic moment for a rotational charged body rotating around a fixed axis," *University Physics*, V16, No. 4, 3–4, Higher Education Press, Beijing, 1997, (Ch).
3. Cheng, D. K., *Field and Wave Electromagnetism*, 210–213, Addison-Wesley Publishing Company, New York, 1983.
4. Lim, Y. K., *Introduction to Classical Electrodynamics*, World Scientific, 1986.
5. Baldassare, D. B., *Classical Theory of Electromagnetism*, Prentic Hall, Engle-wood Cliffs, New Jersey, 1991.
6. Zhou, G.-Q., "Computing rotational inertia of rigid bodies by symmetric operation and dimensional theory," *Journal of Wuhan University*, Vol. 29, No. 5, 90–94, Eng. Ed., 1996, (Ch).
7. Zhou, G.-Q., "The extension of orthogonal axes theorem about moment inertia of rigid body," *Physics Bulletin*, Vol. 8, No. 2, 4–5, 1998, (Ch).
8. Ma, T.-J., *A New Theorem to Calculate the Moment Inertia of Rigid Bodies Physics and Engineering*, Vol. 2, No. 1, 8–9, Qinghua University Press, Beijing, 1995, (Ch).
9. Chen, J.-F., *Classical Mechanics*, 165, Press of China University of Sci. & Tech., He Fei, 1990, (Ch).

The Effects of 884 MHz GSM Wireless Communication Signals on Self-reported Symptoms and Sleep — An Experimental Provocation Study

Bengt Arnetz[1], **Torbjörn Åkerstedt**[2], **Lena Hillert**[2], **Arne Lowden**[2]
Niels Kuster[3], **and Clairy Wiholm**[1]

[1]Wayne State University & Uppsala University, USA
[2]Karolinska Institutet, Sweden
[3]Foundation IT'IS, USA

Abstract— In the current study we assessed possible effects of prolonged (3 hours) exposure to 884 MHz GSM wireless communication signals on self-reported symptoms, cognitive function, and electroencephalographically (EEG) recorded sleep. The study group consisted of 36 women and 35 men. Twenty-two women and sixteen men reported symptoms they specifically related to mobile phone use (SG). The rest of the participants reported no mobile phone-related symptoms (NG).

Potential participants volunteering for the study were evaluated by physicians, including some biochemical assessments, to rule out medical conditions that could interfere with study variables of interest. Once selected, participants spent three different sessions in the laboratory. The habituation session was followed by two subsequent sessions. In these subsequent sessions, subjects were either exposed to sham exposure (sham) or 884 MHz GSM wireless communication signals for 3 hours (an average of 1.4 W/kg including periods of DTX and Non-DTX. Exposure directed to the left hemisphere). Data was collected before, during and following the exposure/sham sessions. Data collected included self-reported symptoms, including headache, cognitive function, mood, and electroencephalographic recordings.

During actual exposure, as compared to sham exposure, sleep initiated one hour after exposure was affected. There was a prolonged latency to reach the first cycle of deep sleep (stage 3). The amount of stage 4 sleep was also decreased in exposed subjects. NG subjects reported more headaches during exposures vs. sham exposure. Neither group (SG and NG) was able to detect the true exposure status more frequently than by chance alone.

The study indicates that during laboratory exposure to 884 MHz wireless signals, components of sleep, believed to be important for recovery from daily wear and tear, are adversely affected. Moreover, participants that otherwise have no self-reported symptoms related to mobile phone use, appear to have more headaches during actual radiofrequency exposure as compared to sham exposure. However, subjects were not able to detect the true exposure status more often than would have been expected by statistical chance alone.

Additional self-reported findings, biochemical, performance and electrophysiological data are currently being analyzed. Possible health implications from the findings will also be further explored.

1. INTRODUCTION

There have been a number of scientific reports concerning the possible relationship between exposure to radio frequency fields (RF) during mobile phone use and self reported symptoms, such as skin sensations, cognitive symptoms, headache, dizziness and sleep disturbances (1–9). However, prior laboratory-based RF exposure studies have been of short-term duration, commonly focusing only on a few outcome variables of interest, and rarely combined self-reported, performance, neurophysiological and neuroendocrine parameters. In order to better understand biological mechanisms behind a possible association between RF exposure and self-reported symptoms, neurophysiological and cognitive responses, we believe a more comprehensive exposure and assessment strategy is needed.

At the previous PIERS meeting in Cambridge, MASS, USA, 2006 we presented the design and methodology of an ongoing double-blind controlled laboratory study with the objective to establish whether RF during mobile phone use had any direct effects on: (1) self-reported symptoms, (2) cognitive symptoms, (3) stress hormones, (4) performance and subsequent sleep and electroencephalogram (EEC). Furthermore, we wanted to establish whether the subjects with self-reported symptoms attributed to mobile phone use (SG) would be able to differentiate exposure conditions from sham, compared to the non symptomatic subjects (NG).

2. METHOD

The exposure set up exposed the left head hemisphere to a GSM signal (884 MHz) at an average of 1.4 W/kg including periods of DTX and Non-DTX (Fig. 1). The exposure was designed to be consistent with worst case exposure occurring in real-life situations, but with extended duration. The exposure laboratory consisted of two separate rooms. In each room, the respective exposure area was shielded with screens, covered with absorbing material. RF and EMF background assessments were conducted prior to initiating the actual study, and quarterly in order to ensure the background field environment in the laboratory was in agreement with specifications in the research protocol. All mobile phone use outside the exposure laboratory was eliminated during exposure sessions. Each exposure session lasted for 3 hours. Three sessions were conducted for each subject, one habituation session and two exposure sessions. The exposure sessions were randomly selected for Sham and RF exposure. During the sessions participants carried out performance and memory tests, scored self-reported symptoms and state of mood. Discrete Likert-type scales were used in the symptom scoring questionnaires. Typically the ratings ranged from 1, indicating "not at all" to 7, "a high degree" or from "disagree strongly" to "agree strongly". After completed exposure the subjects were EEG (electroencephalogram) recorded as they slept in a sleep laboratory. During the subsequent morning additional test were performed before they left the laboratory.

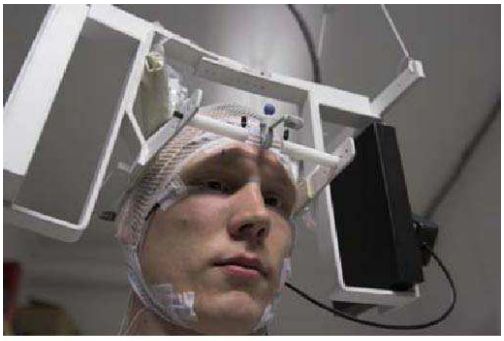


Figure 1: Picture of the actual exposure set up, showing the 884 MHz GSM wireless exposure unit (black box).

3. MATERIAL

The final study group consisted of 71 subjects, age between 18–45 years, where 38 subjects fulfilled the study criteria for SG (22 women and 16 men) and 33 subjects were classified as NG (14 women and 19 men). All subjects reported using their mobile phones daily. The speaking time ranged from five minutes to more than three hours.

4. RESULTS

Sleep initiated one hour after exposure was found to be affected. Under the RF exposure condition, participants exhibited a longer latency to deep sleep (stage 3, meanRF=0.37, (SD=0.33), meanSham=0.27 hours (SD=0.12); F=9.34, p=0.0037). The amount of stage 4 sleep was also decreased (meanRF=37.2 minutes (SD=28), meanSham=45.5 minutes (SD=28); F=10.7, p=0.0019).

For headache, random effects logistic regression was used. Preliminary results reveal a significant interaction effect for exposure and group. The NG reported less headache during sham exposure compared to the SG. The proportion of subjects who reported headache was higher during RF exposure than during sham exposure in the NG but not the SG group. Participants were not able to differentiate RF exposure conditions from sham exposures more often than would have been expected by statistical chance alone. Neither were there any statistically significant differences between the SG and NG in reliable detecting results

5. DISCUSSION

Our results suggest that RF exposure under these conditions is associated with adverse effects on sleep quality within certain sleep stages. The strengths of this study compared to earlier studies are

the longer exposure time during worst conditions and a wider range of outcome variables, including self-reported, neuroendocrine, and neurophysiological variables. There are a number of possible factors that need to be considered when interpreting the results. One being that the exposure laboratory was not fully shielded. However, the shielding walls provided sufficient protection for indirect exposure between the two rooms. Considering this and other possible confounders, we still conclude that there are statistical associations between RF exposure and specific self-reported and neurophysiological variables, pertinent to the current discussions of possible effects from mobile phone-generated RF exposure. Future analysis of the current study will further assess various biological, neurophysiological, and cognitive outcome measures, e.g., spatial memory, performance, and EEG data. We will also attempt to assess the possible clinical relevance of the observed findings.

ACKNOWLEDGMENT

Funding for the study was provided by the Mobile Manufacturers Forum (MMF).

REFERENCES

1. Besset, A, F. Espa, Y. Dauvilliers, M. Billiard, and R. de Seze, "No effect on cognitive function from daily mobile phone use," *Bioelectromagnetics*, Vol. 26, No. 2, 102–8, 2005.
2. Braune, S, A. Riedel, J. Schulte-Monting, and J. Raczek, "Influence of a radiofrequency electromagnetic field on cardiovascular and hormonal parameters of the autonomic nervous system in healthy individuals," *Radiat Res.*, Vol. 158, No. 3, 352–6, 2002.
3. Eliyahu, I., R. Luria, R. Hareuveny, M. Margaliot, N. Meiran, and G. Shani, "Effects of radiofrequency radiation emitted by cellular telephones on the cognitive functions of humans," *Bioelectromagnetics*, Vol. 27, No. 2, 119–26, 2006.
4. Hietanen, M., A. M. Hamalainen, and T. Husman, "Hypersensitivity symptoms associated with exposure to cellular telephones: no causal link," *Bioelectromagnetics*, Vol. 23, No. 4, 264–70, 2002.
5. Hocking, B., "Preliminary report: symptoms associated with mobile phone use," *Occup Med. (Lond)*, Vol. 48, No. 6, 357–60, 1998.
6. Hocking, B., "Symptoms associated with mobile phone use: addendum," *Occup Med (Lond)*, Vol. 48, No. 7, 472, 1998.
7. Koivisto, M., C. Haarala, C. M. Krause, A. Revonsuo, M. Laine, and H. Hamalainen, "GSM phone signal does not produce subjective symptoms," *Bioelectromagnetics*, Vol. 22, No. 3, 212–5, 2001.
8. Oftedal, G, J. Wilen, M. Sandstrom, and K. H. Mild, "Symptoms experienced in connection with mobile phone use," *Occup Med (Lond)*, Vol. 50, No. 4, 237–45, 2000.
9. Rubin, G. J., G. Hahn, B. S. Everitt, A. J. Cleare, and S. Wessely, "Are some people sensitive to mobile phone signals? Within participants double blind randomised provocation study," *BMJ*, Vol. 332, No. 7546, 886–91, 2006.

Behavioral and Cognitive Effects of MW Electromagnetic Field Exposures

Sheila A. Johnston

Independent Neuroscience Consultant, 10 Queen's Mews, London W2 4BZ, UK

Abstract— Published studies of MW exposure thresholds for human adverse perception and cognitive effects with an emphasis on replicated results and the weight of scientific evidence, form the scientific basis for IEEE and ICNIRP exposure limits. Established adverse MW effects are associated with whole body heating at levels that usually increase temperature by approximately 1°C or more. In human cognitive studies with low-level MW exposures, high quality replication studies appear to have eliminated many of the "false positives". Although human cognitive research continues there are no established human cognitive effects at or below standard or guideline limits. In animal spatial memory studies the weight of evidence from six independent research groups is that there is no effect on spatial learning at whole body or local (head) MW exposures within guideline limits and also well above the limits, when whole body temperature did not measurably rise. These replications are of particular note because of the existing large body of scientific evidence of the neurophysiological and neuroanatomical nature of and homology of spatial memory in animals and man. But confirmatory MW exposure, spatial memory experiments in humans are required for health risk assessment. Controversy has followed reports that low level MW exposure caused toxic leakage through the blood brain barrier [BBB] that could lead to cognitive impairment. The BBB area of research remains open while we await the publication of further BBB attempted replication studies. Since changes in temperature are the biologically relevant information, experimental and numerical dosimetry is underway to predict more precisely MW thermophysiological dose-effects on human or animal nervous functions.

1. INTRODUCTION

This paper presents an overview of the interaction of microwaves [MW] electromagnetic fields as external stimuli to the nervous systems and the behavior of humans and laboratory animals. Replicated scientific studies of MW exposure thresholds for human adverse perception and cognitive effects are the scientific basis for IEEE and ICNIRP exposure limits. Reciprocally the standards' thresholds are indicative of and reflect the established scientific evidence for adverse perception and cognitive effects of MWs. These topics are complementary and intertwined. Established adverse MW effects are associated with whole body heating at levels that usually increase temperature by approximately 1°C or more [44]. "Available experimental evidence indicates that the exposure of resting humans for approximately 30 minutes to MWs producing a whole-body SAR of between 1 and 4 W kg^{-1} results in a body temperature increase of less than 1°C" [45]. Animal data indicate a threshold for behavioural responses in the same SAR range [22, 44, 45]. These data form the basis for an occupational exposure restriction of 0.4 W kg^{-1}, which provides a 10X margin of safety [44, 45]. The public standard has an additional 5x margin of safety (0.08 W/kg) for both the IEEE Std C95.1-2005 and ICNIRP guidelines [44–46].

2. MICROWAVE COGNITIVE EFFECTS: HUMAN STUDIES

In recent years, the mobile phone has become extremely popular and is widely used by over 2 billion people. This rapid increase in use of a relatively new technology has resulted in concerns over alleged effects because of exposure of the head to MW energy from the phone antenna. During the period 1999–2006, a variety of studies have been done, both to measure the dose rate to the head and to evaluate the effects of mobile phone exposure on cognitive processes in humans [12, 19, 20, 22, 26, 31–35, 49, 50–53, 58, 59, 64, 70, 71, 80].

The studies that examine human cognitive processes and mobile phone use appear to show no established evidence of memory deficits. Replication studies with standardized protocols, larger samples [30] (multicentered replications; [32]), better experimental controls, double blind conditions, and Bonferroni or other statistical corrections for multiple comparisons [72] appear to have eliminated the "false positives" [46].

Included in cognitive effects are the EEG studies that were previously reviewed in D'Andrea et al., [22]. The reviews by D'Andrea et al. [21, 22] concluded that the evidence for mobile phone

effects on human cognitive performance was very weak. And they concluded that until more studies could be conducted with improved methodology and standardized protocols firm conclusions about cognitive effects, at the low SARs produced by mobile phones, could not be drawn. Since the D'Andrea [21, 22] reviews on EEG there are a few new studies of note: Huber et al. [42, 43] and Loughran et al., [60] showing effects on EEG and Hinrikus et al., [37] reporting the normal EEG is too variable to be able to measure possible small MW effects [46].

Two conclusions can be made. Firstly, the normal EEG appears to be too variable to be able to be a reliable measure of possible small MW effects. For instance, changes in EEG during the night could be attributed to uncontrolled changing brain temperatures that are known to change spike latency of potentials [66]. And the PET scans used to locate the brain site of significant effects of MWs on EEG are known to be unreliable for this purpose [83]. Unless EEG recording and analysis methodology is standardized across experiments and laboratories and improved to avoid experimental artifacts, and the scientific uncertainty in EEG measurements is carefully defined and controlled for, and sham-sham and positive controls used, it appears that further studies should not be considered [46].

3. SPATIAL MEMORY REPLICATIONS AND CONFIRMATIONS

Of particular note is the extensive replicated research (see [48]) that has established scientific evidence of the neurophysiological and neuroanatomical correlates of spatial memory in animals and man. The critical neurotransmission for spatial memory to occur is glutamate binding to glutamate receptors of pyramidal cells in the CA1 region of the hippocampus. Newer research in the literature continues to reconfirm this [13, 11, 61–63]. The effects of MW exposure on spatial memory in animals have also been thoroughly researched with many attempted replications and confirmations [46].

The weight of evidence from six research groups using two species of mammals (Sprague-Dawley rats and C57BL/6J mice) is that there is no effect on spatial learning as tested by six spatial learning paradigms at whole body or local (head) MW exposures within guideline limits and also well above the whole body and partial body limits, when whole body temperature did not measurably rise. Five groups ((i) Sienkiewicz et al., [77]; Jones et al., [47]; (ii) Dubreuil et al., [23–25]; (iii) Yamaguchi et al., [85]; (iv) Cobb et al., [15]; (v) Cassel et al., [14]; Cosquer et al., [16–18]), using spatial memory animal experiments confirmed or replicated each others' results indicating that MW exposures at frequencies from 900 to 2450 MHz had no effect on spatial memory. The positive results of Lai and colleagues [55–57, 84] are not supported by the weight of scientific evidence. Sienkiewicz's group [47] have recently reconfirmed and extended their results on spatial memory in mice exposed to MWs, lending further support to the weight of scientific evidence conclusion.

If the strong evidence for homology of memory is accepted at this time [Nobel Prize 2000], there does not seem to be any further requirement for studies in this area in other strains and species of animals and possibly in man [13, 48, 61–63]. However, it is generally preferred to have confirmatory MW exposure, spatial memory experiments in humans for health risk assessment. Thus, I recommend testing spatial memory in humans during MW exposure.

In conclusion, MW exposure appears to disrupt spatial memory in rats at a dose at or above the whole body threshold of 4 W/kg in a fashion similar to simple tasks of perception [46].

4. THE BLOOD BRAIN BARRIER

Because breaches of the BBB could lead to brain cell death and cognitive impairment [74] BBB studies are of interest in this cognitive review. The BBB prevents high molecular weight substances in the blood from getting into the brain but allows the molecules that are necessary for metabolism to enter. Controversy has followed reports that MW exposure caused toxic leakage through the BBB. In many follow-up studies, most researchers could not replicate the low level BBB permeability changes or could show the effect only at high intensity levels, when the heating of the brain tissue was obvious. D'Andrea et al., [22] reviewed 42 BBB studies. More recently McQuade et al. [65] in a replication of the work of Salford et al. [74, 76] have reported at BEMS 2005, no leakage in the BBB after low SARs [2, 0.2, 0.02, 0.002 W/kg]. Kuribayashi et al. [54], reported the lack of an effect of 1439 MHz local exposure on the BBB in immature rats and young rats. Finally a recently published paper by Cosquer et al., [18] showed no BBB leakage after 2450 MHz MW exposure [2 μs pulse width, 500 pps, whole-body SAR of 2.0 W/kg, $+/-2$ dB and brain averaged SAR of 3.0 W/kg, $+/-3$ dB] as evidenced by no spatial memory deficit after scopolamine methylbromide

administered i.p. This muscarinic antagonist substance does not cross the BBB unless there is leakage in the BBB. If it does it is known to disrupt performance of spatial memory.

The BBB area of research remains open while we are awaiting the publication of results of further BBB attempted replication studies. Because of the recent claims of Salford and Persson in Sweden [1992–2003] of leaks in the BBB and the appearance of dark neurons in the brain, reported to result from MW exposures at incident powers below ICNIRP and IEEE limits, attempted replications of the Salford group studies are underway at the U.S. Air Force Research Laboratory (Mason), in Japan (Ohkubo) and within the French Research programme (Veyret). (Comment: The US Air Force reported in 2006 at the BEMS meeting that they could not replicate Salford results. The Ohkubo and Veyret groups also reported they could not replicate the Salford group's BBB results.) Also replication studies of the Vinogradov group's experiments [82] are being performed in Moscow and at PIOM, France where they have similar exposure systems [funded by MMF] and the ability to test the immune response [81].

5. FUTURE DIRECTIONS IN HUMAN MW STUDIES

There is no known biophysical mechanism based on the modulation of telecommunications signals except the energy absorbed as heat that results in temperature increase [28, 44]. There are two ways to proceed. One is to build on what we know, that is to better define thermal effects. The other is to continue to investigate each modulation for biological effects [10]. We are proceeding with research on both fronts.

To better define thermophysiological effects of MW exposure on humans is an important part of answering remaining questions about possible MW effects since heat is the established mechanism of interaction between tissue and MWs (for low power devices). Thermophysiological knowledge and dosimetric modeling along with cognitive and behavioural tests and traceable exposure dosimetry will help resolve whether or not MWs have any effects thermal or otherwise on human functions below IEEE or ICNIRP limits. Recently a very important workshop established an international collaboration with the primary goal to develop appropriate techniques for predicting the thermophysiological responses of human beings who are exposed to MW fields at specific frequencies, field strengths, and field characteristics and to validate some predictions with existing human thermophysiology data [1].

Masses of data exist that describe the thermoregulatory response changes in the human body as a function of environmental variables, work, exercise, age, fitness, clothing insulation, and other characteristics of each individual. Much of this material is amenable to comparison with data derived from MW-exposed humans [2–9, 36, 67]. If one or more types of models are combined with appropriate physiological and dosimetric data, it will be important to validate the predictions of the models. Data already collected on human volunteers exposed to MW energy at several frequencies and field strengths [2–9] are available to compare with predictions at those same frequencies and field strengths. For example, a 27-node model developed by Stolwijk and Hardy [67]; and Stolwijk, [79] has been used to predict, with reasonable accuracy, the physiological responses of human volunteers exposed to 100, 450, and 2450 MHz CW in controlled thermal environments [27]. Thermophysiological modeling will become a useful tool to determine MW safe exposure levels and hazard thresholds across the RF/MW spectrum for many classifications of people.

6. CONCLUSION

Humans and laboratory animals have shown the capability to detect fairly low levels of MW exposure. Research is continuing to explore the molecular characteristics of heat receptors in the skin and in the CNS [67]. International microwave standards are based on thermal effects, notably adverse behavioral effects, following short-term MW exposures at higher intensities [44–46]. The adverse behavioral effects with a threshold whole-body specific absorption rate (SAR) of 4 W/kg are fairly well understood and serve as the basis (with a 10x lower safety margin for occupational exposure and 50x lower margin for the general public) for numerous safe exposure standards worldwide [44–46].

Dosimetry to establish levels of exposure from mobile telecommunications and to establish guideline limits has been focused on SAR measurement. Much of that work is completed. Microwave dosimetry research is now undertaking microwave thermal modelling to predict the location and amount of heat absorption in humans and the behavioural consequences and establish thermal thresholds for adverse effects on human behaviours. Research to compare SAR and thermal distributions are suggesting similar distributions thus verifying standards based on SAR [29, 38, 39].

Since changes in temperature are the biologically relevant information for setting the standards, in future, the thermophysiological, and behavioural data could be incorporated into the models as the experimental and numerical dosimetry is improved to predict thermal responses to new exposures. This sort of modelling would predict the location and amount of heat absorption in humans and the behavioural consequences and establish accurate human thresholds for adverse effects on human behaviours and cognition. Once this is completed we will be able to set the most biologically relevant and validated science based thermal standards.

This thermophysiological knowledge and precise dosimetric modelling and traceable exposure dosimetry will help further resolve whether or not MWs have any effects thermal or otherwise on human cognitive functions below IEEE or ICNIRP limits.

ACKNOWLEDGMENT

I wish to acknowledge John A D'Andrea, Senior Scientist, Naval Health Research Center Detachment, Brooks Air Force Base, [john.dandreabrooks.af.mil] has made a very significant input to this review. Please note reference 46.

REFERENCES

1. Adair, E. R., "A mechanism by which heat loss response of human subjects are stimulated during whole body RF exposure at resonance," *COST 281 Workshop "Subtle Temperature Effects of RF-EMF"*, London, November 12–13, 2002. http://www.cost281.org/
2. Adair, E. R. and D. R. Black, "Thermoregulatory responses to RF energy absorption," *Bioelectromagnetics Supplement*, Vol. 6, S17–S38, 2003.
3. Adair, E. R., D. W. Blick, S. J. Allen, K. S. Mylacraine, J. M. Ziriax, and D. M. Scholl, "Thermophysiological responses of human volunteers to whole body RF exposure at 220 MHz," *Bioelectromagnetics*, Vol. 26, No. 6. 448–461, Sep. 2005.
4. Adair, E., et al., "Thermophysiological consequences of whole body resonant RF exposure 100 MHz in human volunteers," *Bioelectromagnetics*, Vol. 247, 489–501, 2003.
5. Adair, E., et al., "Thermophysiological responses of human volunteers during controlled whole-body radio frequency exposure at 450 MHz," *Bioelectromagnetics*, Vol. 194, 232–245, 1998.
6. Adair, E. R., K. S. Mylacraine, and B. L. Cobb, "Human exposure to 2450 MHz CW energy at levels outside the IEEE C95.1 standard does not increase core temperature," *Bioelectromagnetics*, Vol. 226, 429–439, 2001.
7. Adair, E. R., K. S. Mylacraine, and B. L. Cobb, "Partial-body exposure of human volunteers to 2450 MHz pulsed or CW fields provokes similar thermoregulatory responses," *Bioelectromagnetics*, Vol. 224, 246–259, 2001.
8. Allen, S. J., E. R. Adair, K. S. Mylacraine, W. Hurt, and J. Ziriax, "Empirical and theoretical dosimetry in support of whole body radio frequency (RF) exposure in seated human volunteers at 220 MHz," *Bioelectromagnetics*, Vol. 26. No. 6, 440–447, Sep. 2005.
9. Allen, S., et al., "Empirical & theoretical dosimetry in support of whole body resonant RF exposure 100 MHz, *Bioelectromagnetics*, Vol. 247, 502–509, 2003.
10. Andersen, J. B., "Signal forms in wireless applications," *Do Sinusoidal versus Non-sinusoidal Waveforms Make a Difference?*, Zurich, Feb. 3, 2005. http://www.cost281.org/documents.php?node=96&dir_session=
11. Bast, T., B. M. da Silva, and R. G. Morris, "Distinct contributions of hippocampal NMDA and AMPA receptors to encoding and retrieval of one-trial place memory," *J. Neurosci.*, Vol. 25, No. 25, 5845–5856, Jun. 22, 2005.
12. Besset, A., et al., "No effect on cognitive function from daily mobile phone use," *Bioelectromagnetics*, Vol. 26, 102–108, 2005.
13. Burgess, N., E. A. Maguire, and J. O'Keefe, "The human hippocampus and spatial and episodic memory," *Neuron*, Vol. 35, 625–641, Review, 2002.
14. Cassel, J. C., B. Cosquer, R. Galani, and N. Kuster, "Whole-body exposure to 2.45 GHz electromagnetic fields does not alter radial-maze performance in rats," *Behav. Brain Res.*, Vol. 1551, 37–43, 2004.
15. Cobb, B., et al., "Radial arm maze performance of rats following repeated low level microwave radiation," *Bioelectromagnetics*, Vol. 251, 49–57, Jan. 2004.
16. Cosquer, B., R. Galani, N. Kuster, and J. C. Cassel, "Whole-body exposure to 2.45 GHz electromagnetic fields does not alter anxiety responses in rats: a plus-maze study including test validation," *Behav. Brain Res.*, Vol. 156, 65–74, 2005.

17. Cosquer, B., N. Kuster, and J. C. Cassel, "Whole-body exposure to 2.45 GHz electromagnetic fields does not alter 12-arm radial-maze with reduced access to spatial cues in rats," *Behav. Brain Res.*, Vol. 161, 331–334, 2005.
18. Cosquer, B., A. P. Vasconcelos, J. Frohlich, and J. C. Cassel, "Blood-brain barrier and electromagnetic fields: effects of scopolamine methylbromide on working memory after whole-body exposure to 2.45 GHz microwaves in rats," *Behav. Brain Res.*, Vol. 161, 229–237, 2005.
19. Croft, R. J., J. S. Chandler, A. P. Burgess, R. J. Barry, J. D. Williams, and A. R. Clarke, "Acute mobile phone operation affects neural function in humans," *Clin. Neurophysiol.*, Vol. 113, 1623–1632, 2002.
20. Curcio, G., M. Ferrara, L. De Gennaro, R. Cristiani, G. D'Inzeo, and M. Bertini, "Time-course of electromagnetic field effects on human performance and tympanic temperature," Neuro Report, Vol. 151, 161–164, 2004.
21. D'Andrea, J., et al., "Behavioral and cognitive effects of microwave exposure," *Bioelectromagnetics Supplement*, Vol. 6, S39–S62, 2003.
22. D'Andrea, J., et al., "Microwave effects on the nervous system," *Bioelectromagnetics Supplement*, Vol. 6, S107–S147, 2003.
23. Dubreuil, D., T. Jay, and J.-M. Edeline, "Does head-only exposure to GSM-900 electromagnetic fields affect the performance of rats in spatial learning tasks?" *Behav. Brain Res.*, Vol. 129, 203–210, 2002.
24. Dubreuil, D., T. Jay, and J. M. Edeline, "Head-only exposure to GSM 900-MHz electromagnetic fields does not alter rat's memory in spatial and non-spatial tasks," *Behav. Brain Res.*, Vol. 145, 51–61, 2003.
25. Dubreuil, D., et al., "Does the radial arm maze necessarily test spatial memory?" *Neurobiol Learn Mem.*, Vol. 79, 109–117, 2003.
26. Edelstyn, N. and A. Oldershaw, "The acute effects of exposure to the electromagnetic field emitted by mobile phones on human attention," Neuro Report, Vol. 13, 119–121, 2002.
27. Foster, K. R. and E. R. Adair, "Modeling thermal responses in human subjects following extended exposure to radiofrequency energy," *BioMedical Engineering Online*, Vol. 3, No. 4, 2004. http://www.biomedical-engineering-online.com/content/3/1/4.
28. Foster, K. R. and M. H. Repacholi, "Biological effects of radiofrequency fields: does modulation matter?" *Radiat. Res.*, Vol. 162, 219–225, 2004.
29. Fujimoto, M., et al., "FDTD-derived correlation of maximum temperature increase and peak SAR in child and adult head models due to dipole antenna," *IEEE Transactions on Electromagnetic Compatibility*, Vol. 48, No. 1, 240–247, Feb. 2006.
30. Hamblin, D. L., R. J. Croft, A. W. Wood, C. Stough, and J. Spong, "The sensitivity of human event-related potentials and reaction time to mobile phone emitted electromagnetic fields," *Bioelectromagnetics*, Vol. 27, No. 4, 265–273, May 2006.
31. Haarala, C., et al., Electromagnetic field emitted by 902 MHz mobile phones shows no effects on children's cognitive function," *Bioelectromagnetics*, Vol. 26, No. S7, S144–S150, Jul. 29, 2005.
32. Haarala, C., L. Bjornberg, M. Ek, M. Laine, A. Revonsuo, M. Koivisto, and H. Hämäläinen, "Effect of a 902 MHz electromagnetic field emitted by mobile phones on human cognitive function: A replication study," *Bioelectromagnetics*, Vol. 24, 283–288, 2003.
33. Haarala, C., M. Ek, L. Bjornberg, M. Laine, A. Revonsuo, M. Koivisto, and H. Hamalainen, "902 MHz mobile phone does not affect short term memory in humans," *Bioelectromagnetics*, Vol. 256, 452–456, Sep. 2004.
34. Hamblin, D. and A. Wood, "Effects of mobile phone emissions on human brain activity and sleep variables," *Int. J. Radiat. Biol.*, Vol. 78, 659–669, 2002.
35. Hamblin, D. L., A. W. Wood, R. J. Croft, and C. Stough, "Examining the effects of electromagnetic fields emitted by GSM mobile phones on human event-related potentials and performance during an auditory task," *Clin. Neurophysiol.*, Vol. 1151, 171–178, 2004.
36. Hensel, H., "Thermoreception and temperature regulation," *Monogr. Physiol. Soc.*, Vol. 38, 1–321, 1981.
37. Hinrikus, H., M. Parts, J. Lass, and V. Tuulik, "Changes in human EEG caused by low level modulated microwave stimulation," *Bioelectromagnetics*, Vol. 25, 431–440, 2004.

38. Hirata, A. and T. Shiozawa, "Correlation of maximum temperature increase and peak SAR in the human head due to handset antennas," *IEEE Trans. Microwave Theory, Tech.*, Vol. 51, No. 7, 1834–1841, 2003.
39. Hirata, A., S. Watanabe, M. Kojima, I. Hata, K. Wake, M. Taki, K. Sasaki, O. Fujiwara, and T. Shiozawa, "Computational verification of anesthesia effect on temperature variations in rabbit eyes exposed to 2.45 GHz microwave energy," *Bioelectromagnetics*, May 24. 2006. http://dx.doi.org/10.1002/bem.20251.
40. Huber, R., T. Graf, K. A. Cote, L. Wittmann, E. Gallmann, D. Matter, J. Schuderer, N. Kuster, A. A. Borbély, and P. Achermann, *Neuro Report*, Vol. 11, 3321–3325, 2000.
41. Huber, R., J. Schuderer, T. Graf, K. Jütz, A. A. Borbely, N. Kuster, and P. Achermann, "Radio frequency electromagnetic field exposure in humans: Estimation of SAR distribution in the brain, effects on sleep and heart rate," *Bioelectromagnetics*, Vol. 24, 262–276, 2003.
42. Huber, R., et al., "Electromagnetic fields, such as those from mobile phones, alter regional cerebral blood flow and sleep and waking EEG," *Journal of Sleep Research*, Vol. 11, 289–295, 2002.
43. Huber, R., et al., "Exposure to pulse-modulated radio frequency electromagnetic fields affects regional cerebral blood flow," *European Journal of Neuroscience*, Vol. 21, 1000–1006, 2005.
44. IEEE Std C95.1-2005, IEEE Standard for Safety Levels with Respect to Human Exposure to Radio Frequency Electromagnetic Fields, 3 kHz to 300 GHz, The Institute of Electrical and Electronic Engineers, New York.
45. ICNIRP International Commission on Non-Ionizing Radiation Protection, "Guidelines for limiting exposure to time-varying electric, magnetic, and electromagnetic fields up to 300 GHz," *Health Physics*, Vol. 74, 494–522, 1998.
46. Johnston, S. A. and J. A. D'Andrea, *Behavioral and Cognitive Effects of Electromagnetic Field Exposures*, B. Greenebaume and F. Barnes, eds., Handbook of Biological Effects of EMF CRC Press, 2006.
47. Jones, N., et al., "Perform B: Behavioral studies with mice," *BEMS Abstract*, 184–85, 2005.
48. Kandel, E., J. Schwartz, and M. Jessell, (editors), *Principles of Neural Science*, 4th Edn., McGraw Hill, London, 2000.
49. Koivisto, M., C. M. Krause, A. Revonsuo, M. Laine, and H. Hamalainen, "The effects of electromagnetic field emitted by GSM phones on working memory," *Neuro Report*, Vol. 11, 1641–1643, 2000.
50. Koivisto, M., et al., "Effects of 902 MHz electromagnetic field emitted by cellular telephones on response times in humans," *Neuro Report*, Vol. 11, 413–415, 2000.
51. Krause, C. M., et al., "Effects of electromagnetic field emitted by cellular phones on the EEG during an auditory memory task: a double blind replication study," *Bioelectromagnetics*, Vol. 251, 33–40, Jan. 2004.
52. Krause, C. M., et al., "Effects of electromagnetic field emitted by cellular phones on the EEG during a memory task," *Neuro Report*, Vol. 11, 761–764, 2000.
53. Krause, C. M., et al., "Effects of electromagnetic fields emitted by cellular phones on the electroencephalogram during a visual working memory task," *Intl. J. Rad. Biol.*, Vol. 76, 1659–1667, 2000.
54. Kuribayashi, M., et al., "Lack of effects of 1439 MHz electromagnetic near field exposure on the blood-brain barrier in immature and young rats," *Bioelectromagnetics*, Vol. 26, No. 7, 578–588, Sep. 2, 2005.
55. Lai, H., "Interaction of microwaves and a temporally incoherent magnetic field on spatial learning in the rat," *Physiol Behav.*, Vol. 82, No. 5, 785–789, 2004.
56. Lai, H., M. A. Carino, A. Horita, and A. W. Guy, "Low-level microwave irradiation and central cholinergic systems," *Pharmacol. Biochem. Behav.*, Vol. 33, 131–138, 1989.
57. Lai, H., A. Horita, and A. W. Guy, "Microwave irradiation affects radial-arm maze performance in the rat," *Bioelectromagnetics*, Vol. 15, 95–104, 1994.
58. Lee, T. M. C., S. M. Y. Ho, L. Y. H. Tsang, S. Y. C. Yang, L. S. W. Li, and C. C. H. Chan, "Effect on human attention of exposure to the electromagnetic field emitted by mobile phones," *Neuro Report*, Vol. 12, 729–731, 2001.
59. Lee, T. M., P. K. Lam, L. T. Yee, and C. C. Chan, "The effect of the duration of exposure to the electromagnetic field emitted by mobile phones on human attention," *Neuro Report*, Vol. 18, 1361–1364, 2003.

60. Loughran S. P., A. W. Wood, J. M. Barton, R. J. Croft, B. Thompson, and C. Stough, "The effect of electromagnetic fields emitted by mobile phones on human sleep," Neuro Report, Vol. 16, No. 17, 1973–1976, Nov. 28, 2005.
61. Maguire, E. A., D. G. Gadian, I. S. Johnsrude, C. D. Good, J. Ashburner, R. S. Frackowiak, and C. D. Frith, "Navigation-related structural change in the hippocampi of taxi drivers," *Proc. Natl. Acad. Sci.*, Vol. 97, 4398–4403, USA, 2000. http://www.pubmedcentral.nih.gov/articlerender.fcgi?tool=pubmed&pubmedid=10716738
62. Maguire, E. A., H. J. Spiers, C. D. Good, T. Hartley, R. S. Frackowiak, and N. Burgess, "Navigation expertise and the human hippocampus: A structural brain imaging analysis," *Hippocampus*, Vol. 132, 250–259, 2003.
63. Maguire, E. A., et al., "Routes to remembering: the brains behind superior memory," *Nat. Neurosci.* Vol. 61, 90–95, 2003.
64. Maier, R., S. E. Greter, and R. Maier, "Effects of pulsed electromagnetic fields on cognitive processes — A pilot study on pulsed field interference with cognitive regeneration," *Acta Neurologica Scandinavica*, Vol. 110, No. 46, 2004.
65. McQuade, J. S., J. H. Merritt, O. Rahimi, S. A. Miller, T. Scholin, A. L. Salazar, M. C. Cook, and P. A. Mason, "Effects of 915 MHz exposure on the integrity of the blood brain barrier," *Bioelectromagnetics*, Vol. 15, No. 3, 171–172, 2005.
66. Moser, E. I., M. B. Moser, and P. Andersen, "Potentiation of dentate synapses initiated by exploratory learning in rats: Dissociation from brain temperature, motor activity, and arousal," *Learn. Mem.*, Vol. 1, No. 1, 55–73 May–June, 1994.
67. Patapoutian, A., et al., "ThermoTRP channels and beyond: mechanisms of temperature sensation," *Nat. Rev. Neurosci.*, Vol. 4, 529–539, 2003.
68. Persson, B. R. R., L. G. Salford, A. Brun, and L. Malmgren, "Increased permeability of the blood-brain barrier induced by magnetic and electromagnetic fields," *Ann. NY Acad. Sci.*, Vol. 649, 356–358, 1992.
69. Persson, B. R. R., L. G. Salford, and A. Brun, "Blood-brain barrier permeability in rats exposed to electromagnetic fields used in wireless communication," *Wireless Network*, Vol. 3, 455–461, 1997.
70. Preece, A. W., et al., "Effect of 902 MHz mobile phone transmission on cognitive function in children," *Bioelectromagnetics Suppl.*, Vol. 7, S138–43, 2005.
71. Preece, A. W., et al., "Effect of a 915-MHz simulated mobile phone signal on cognitive function in man," *Int. J. Radiat. Biol.*, Vol. 75, 447–456, 1999.
72. Regel, S. J., et al., "UMTS base station-like exposure, well-being, & cognitive performance," *EHP*, Vol. 114, No. 8, 1270–1275, Aug. 2006.
73. Reilly, J. P., *Applied Bioelectricity*, Springer, New York, 1998.
74. Salford, L. G., et al., "Nerve cell damage in mammalian brain after exposure to microwaves from GSM mobile phones," *EHP*, Vol. 11, No. 7, 2003.
75. Salford, L. G., A. Brun, J. L. Eberhardt, and R. R. Persson, "Permeability of the blood-brain barrier induced by 915 MHz electromagnetic radiation, continuous wave and modulated at 8, 16, 50, and 200 Hz," *Bioelectrochemistry and Bioenergetics*, Vol. 30, 293–301, 1993.
76. Salford, L. S., A. Brun, K. Sturesson, J. L. Eberhardt, and B. R. R. Persson, "Permeability of the blood-brain barrier induced by 915 MHz electromagnetic radiation, continuous wave and modulated at 8, 50, and 200 Hz," *Microscopy Res. Tech.*, Vol. 27, 535–42, 1994.
77. Sienkiewicz, Z. J., R. P. Blackwell, R. G. Haylock, R. D. Saunders, and B. L. Cobb, "Low-level exposure to pulsed 900 MHz microwave radiation does not cause deficits in the performance of a spatial learning task in mice," *Bioelectromagnetics*, Vol. 21, 151–158, 2000.
78. Stolwijk, J. A. J. and J. D. Hardy, "Temperature regulation in man — a theoretical study," *Pflugers AGPMT*, Vol. 291, No. 2, 129–162, 1966.
79. Stolwijk, J. A. J., "Thermoregulatory response to microwave power deposition, in Microwaves & Thermoregulation, Adair, E. R., Ed. Academic Press, New York, 129, 1983.
80. Van Leeuwen, G., et al., "Calculation of change in brain temperatures due to exposure to a mobile phone," *Phys. Med. Biol.*, Vol. 44, 2367–79, 1999.
81. Veyret, B., *Review of Research*, 2003.
http://www.who.int/peh-emf/research/agenda/en/index.html]

82. Vinogradov, G., "The phenomenon of autoimmunity from the effects of nonionizing microwave radiation," *Electricity and Magnetism in Biology and Medicine*, 649–650, M. Blank, ed., San Francisco Press, Inc., 1993.
83. Wager, T. D. and E. E. Smith, "Neuroimaging studies of working memory: a meta-analysis," *Cogn. Affect Behav. Neurosci*, Vol. 3, No. 4, 255–274, Dec. 2003.
84. Wang, B. and H. Lai, "Acute Exposure to Pulsed 2450-MHz Microwaves Affects Water-Maze Performance of Rats," *Bioelectromagnetics*, 21, 52–56, 2000.
85. Yamaguchi, H., et al., "1439 MHz pulsed TDMA fields affect performance of rats in a T-maze task only when body temperature is elevated," *Bioelectromagnetics*, Vol. 24, 223–230, 2003.

Radiofrequency Exposure and Human Health

J. A. Elder, C-K. Chou, and J. J. Morrissey

Motorola Labs, 8000 W. Sunrise Blvd., Ft. Lauderdale, FL 33322, USA

Abstract— Research on the biological and health effects of radiofrequency (RF) fields has been conducted for more than 50 years and the RF database available in the 1990's proved adequate for the development of the human exposure limits recommended in 1998 by the International Commission on Non-Ionizing Radiation Protection (ICNIRP) [1]. The ICNIRP guidelines are recommended by the World Health Organization (WHO) and have been adopted by more than 35 countries. The database that led to the development of the ICNIRP guidelines has grown, with about 500 studies at mobile phone frequencies including many modulated signals. The WHO database [2] has more than 1500 original, peer-reviewed papers useful for public health risk assessment of RF exposure. The database provides even stronger evidence today that RF exposures within ICNIRP limits associated with mobile telephony pose no known health risks and warrant no special precautions for any segments of the population. WHO has stated that scientific knowledge on electromagnetic fields including RF fields is now more extensive than for most chemicals [3]. Expert scientific organizations, international organizations and government agencies that have reviewed the available database since the publication of the ICNIRP guidelines include the UK Independent Expert Group on Mobile Phones (Stewart Report) (2000); German Commission on Radiological Protection (2001); Australian Communications Authority (2003); French Environmental Health and Safety Agency (2003); Swedish Radiation Protection Authority (2003), Health Council of the Netherlands (2004); UK Advisory Group on Non-Ionising Radiation (2004), UK National Radiological Protection Board (2004); US Food and Drug Administration (2005); and the World Health Organization (2005, 2006). All of these reviews have consistently concluded that there is no credible or convincing evidence that RF exposure within ICNIRP limits causes adverse human health effects. This paper describes a) the extensive database on the biological and health effects of exposure to RF energy, b) the ICNIRP RF safety guidelines, and c) recent conclusions of national and international expert groups that have evaluated the scientific and medical evidence on the potential health effects of RF exposure.

1. RADIOFREQUENCY (RF) DATABASE

The WHO database [2] on biological and health effects of RF energy is extensive. It has more than 2500 scientific items from countries around the world. In addition to reviews, engineering studies and non-peer-reviewed articles, the database has more than 1500 peer-reviewed papers that satisfy criteria for use as a basis to assess the possible public health impacts of RF exposure (see Table 1). The first table shows the number of ongoing, reported-but-not-published research and published entries in the database for each of the following types of scientific studies on RF fields: epidemiological, human, animal, and cellular studies. The WHO database is unique in that it includes summaries of published papers and summaries of ongoing and reported-but-not-published research.

Although all peer-reviewed studies in the RF database (Table 1) are considered relevant to the mobile phone issue, it is of interest to note the large number of studies using radiofrequencies specific to mobile telephony as shown in Table 2. In this table, the number of studies in each of the four types of scientific investigations is shown. There are 749 entries listed in the database using mobile telephony-specific signals and 484 of these have been published. The RF database is available to the public on the WHO website shown in Tables 1 and 2.

The accepted process of evaluating any chemical or physical agent for their potential for causing human health effects can be described as follows. Such assessments place emphasis on human data (epidemiology and other human studies) but also rely on animal data, particularly long-term exposure studies, when human data are weak or nonexistent. In vitro data are only used as supportive evidence of a mechanism if evidence exists in vivo. A recent paper describing a pooled analysis of 12 epidemiology publications investigating tumors in mobile phone users [4] found no association between RF exposure and cancer. This review included papers from several laboratories involved in the INTERPHONE Project, a large multi-centre case-control study of head and neck tumors in mobile phone users directed by the International Agency for Research on Cancer (IARC). At the time of writing this paper, a draft manuscript describing the results from all studies from the

13 countries in the INTERPHONE Project is being finalized and the pooled analysis is expected to be published in 2007.

The weight of scientific evidence[1] from the available epidemiological studies indicates no adverse health effects and this conclusion is strongly supported by results from animal cancer studies; many of these studies have well-defined RF exposure data useful for risk analysis. There are 37 animal cancer studies in the database and the majority were published in the past 10 years. These papers have considerably strengthened the RF database because a number of these studies employed experimental protocols similar to those used by the U.S. National Toxicology Program to determine the carcinogenic potential of chemical and physical agents. The weight of evidence of these animal studies, including studies in which animals were exposed daily throughout their lifetimes, supports the conclusion that RF exposure does not cause or promote tumor formation [5–7]. RF exposure durations have ranged from 2–22 hours per day up to two years, exposure levels in these studies have ranged up to 4 W/kg, and exposure frequencies have ranged from about 400 to 9000 MHz, including a number of studies at radiofrequencies specific to mobile telephony (e. g., 900 and 1800 MHz). In addition, the weight of scientific evidence of the animal cancer studies indicates no effect on survival or body weight at exposure levels less than 4 W/kg, which is regarded as the exposure threshold for adverse effects in animals. These results provide strong evidence that RF exposure does not cause life shortening diseases or general toxicity at exposure levels within ICNIRP limits which are set well below the adverse effect threshold of 4 W/kg [6].

Table 1: Biological and health effect studies of RF exposure in the WHO database.
Studies listed on WHO website under citation listings:
http://www.who.int/peh-emf/research/database/en/

Research Study Type	**Ongoing**	**Reported but not Published**	**Peer-reviewed Publications**
Epidemiology	44	9	233
Human Studies	64	17	170
Animal Studies	49	30	722
Cellular Studies	64	31	401
Totals	**221**	**87**	**1526**

Table 2: Mobile telephony relevant studies in the WHO database.
These studies are listed on WHO website: http://www.who.int/peh-emf/research/database/en/

Research Study Type	**Ongoing**	**Reported but not Published**	**Peer-reviewed Publications**
Epidemiology	38	7	78
Human Studies	60	16	107
Animal Studies	43	20	181
Cellular Studies	55	26	118
Totals	**196**	**69**	**484**

Some reports have suggested that biological effects are frequency specific and are associated with low-frequency modulations of the RF signal. The studies in the extensive RF database use a wide range of frequencies and include many modulations characteristic of mobile telephony and other signals. The weight of scientific evidence shows no frequency-dependent or modulation-dependent response suggestive of adverse health effects. This statement is supportive of the scientific consensus that all established adverse health effects are thermal effects.

[1]Weight of scientific evidence is the outcome of assessing the published information about the biological and health effects from exposure to RF energy. This process includes evaluation of the quality of test methods, the size and power of the study designs, the consistency of results across studies, and the biological plausibility of dose-response relationships and statistical associations (see page 12 in [9])

Much of the peer-reviewed literature has been summarized in eleven review papers published in a Special Issue of Bioelectromagnetics in December 2003 [8]. These reviews cover epidemiology, cancer, survival, central nervous system, behavioral and cognitive effects, thermoregulation, ocular effects, auditory responses, teratogenesis, effects on blood cells and functions of the cardiac, endocrine and immune systems and in vitro studies on mammalian cell toxicity, genotoxicity and cellular transformation, i.e., cancer-related changes. All of these reviews are available to the public and can be downloaded from the Wiley Interscience website (http://www3.interscience.wiley.com/cgi-bin/jhome/34135) [Open website, click on "Issues" and then click on Bioelectromagnetics, Volume 24, Supplement 6(S6), December 2003.] The publication of the recently revised IEEE C95.1-2005 RF safety standard [9] includes a comprehensive review of more than 1300 primary peer reviewed papers on RF biological effects. Overall, the reviews published in 2003 in Bioelectromagnetics [8] and the IEEE review [9] support the conclusion that the only adverse effects of RF exposure are thermal effects and the adverse effect threshold level is 4 W/kg, the basis for both the IEEE C95.1-2005 standard [9] and the ICNIRP guidelines [1].

2. THE ICNIRP GUIDELINES

The process for developing an RF exposure guideline requires a review of all available scientific evidence on the subject, including papers describing thermal and non-thermal effects, short-term and long-term exposures, cancer and other biological and health endpoints, and epidemiological and other human studies. The purpose of the review includes identification of all established adverse health effects and identification of the threshold exposure level causing the effect. The RF database available in the 1990's proved adequate for the development of the human exposure limits recommended in 1998 by ICNIRP [1]. There is broad scientific consensus that RF exposure at high levels can cause adverse health effects due to a significant increase in temperature of body tissues or in the whole body. This consensus provides a reliable baseline for the development of exposure limits that protect all segments of the population from established adverse effects.

The ICNIRP guidelines are based on established adverse effects of RF exposure and the only replicated adverse effects are caused by RF heating (thermal effects). These effects are well-understood and exhibit a clear RF exposure threshold level. Adherence to the ICNIRP exposure guidelines ensures RF exposure levels remain below this tissue heating threshold. The ICNIRP guidelines [1] are recommended by the World Health Organization and have been adopted by more than 35 countries.

Recently, WHO summarized important points in the ICNIRP guidelines concerning the magnitude of safety factors and the protection provided to people of all ages as follows: "The ICNIRP guidelines were developed to limit human exposure to electromagnetic fields (EMF) under conditions of maximum absorption of the fields, which rarely occurs, and the limits incorporate large safety factors to protect workers and even larger safety factors to protect the general public, including children. Thus, the limits in the ICNIRP guidelines are highly protective and are based on all the available scientific evidence" [10].

3. EXPERT REVIEWS

Many national and international expert groups have evaluated the evidence on the potential health and biological effects of RF fields and have been consistent in their conclusions that RF exposures within the ICNIRP limits pose no known health risks. This section is a compilation of the conclusions of expert groups over the past 6 years regarding the ICNIRP guidelines, mobile phones and base stations.

The report by the Independent Expert Group on Mobile Phones [11] in the United Kingdom (UK), commonly know as the Stewart Commission, concluded that "the balance of evidence to date suggests that exposures to RF radiation below NRPB and ICNIRP [International Commission on Non-Ionising Radiation Protection] guidelines do not cause adverse health effects to the general population."

In Germany, the Commission on Radiological Protection (SSK) [12] issued the following statement: "the SSK concludes that even after assessing the recent scientific literature, there are no new scientific findings with respect to proven adverse impact on health that give rise to doubts regarding the scientific assessment underlying the protective concepts of ICNIRP or the European Council Recommendation."

The French Environmental Health and Safety Agency (AFSSE) [13] concluded that "the general analysis of current scientific data on exposure to base station waves shows no health risk linked to

mobile phone base stations."

With specific reference to health aspects of third generation (3G) phones, the Australian Communications Authority (ACA) [14] concluded: "The weight of national and international scientific opinion is that there is no substantiated evidence that exposure to low level radiofrequency EME causes adverse health effects. This view has been backed by every major review panel ..."

The independent UK Advisory Group on Non-Ionising Radiation (AGNIR) [15] examined recent experimental and epidemiological evidence for health effects due to exposure to RF transmissions, including those associated with mobile telephone handsets and base stations. This review focused on the scientific data made available since the publication of the IEGMP report in 2000 mentioned above. In their 2004 report, the AGNIR concluded: "In aggregate the research published since the IEGMP report does not give cause for concern. The weight of evidence now available does not suggest that there are adverse health effects from exposures to RF fields below guideline levels, but the published research on RF exposures and health has limitations, and mobile phones have only been in widespread use for a relatively short time."

Also in 2004, the Electromagnetic Fields Committee of the Health Council of the Netherlands (HCN) [16] expressed its support for the ICNIRP limits and concluded that "no health problems can be expected to occur as a direct result of exposure to those fields. Furthermore, the Committee feels that there are no health-based reasons for limiting the use of mobile phones by children."

The 2004 review by the UK National Radiological Protection Board (NRPB) [17] was an update of both the above mentioned AGNIR review and the May 2000 report from the IEGMP (as known as the Stewart Report). The NRPB advice stated: "Since then, the widespread development in the use of mobile phones world-wide has not been accompanied by associated clearly established increases in adverse health effects. Within the UK, there is a lack of hard information showing that the mobile phone systems in use are damaging to health. It is important to emphasis this crucial point."

In response to the NRPB report, the US Food and Drug Administration (FDA) [18] released the following statement: "FDA agrees with the NRPB on its conclusions that there is 'no hard evidence of adverse health effects on the general public' from exposure to radiofrequency energy while using wireless communication devices ... With regards to the safety and use of cell phones by children, the scientific evidence does not show a danger to users of wireless communication devices including children."

More recently, WHO issued a clarification statement on children and mobile phones reading in part as follows: "In 2000 WHO issued a fact sheet (#193) on Mobile Phones and their Base Stations. In the section under 'Precautionary measures' it states 'Present scientific evidence does not indicate the need for any special precautions for the use of mobile phones. If individuals are concerned, they might choose to limit their own or their children's RF exposure by limiting the length of calls, or by using 'hands-free' devices to keep mobile phones away from the head and body. Not only is the information provided in this WHO fact sheet still valid, but the precautionary measures suggested are still those recommended by the International EMF Project. For further information readers are referred to: http://www.who.int/emf'" [10].

At its annual symposium on Cell Phones and Cancer, the European Cancer Prevention Organization [19] developed a consensus statement including the conclusion that "The European Cancer Prevention Organization states that, in 2005 there is insufficient contemporary proof with regard to increased cancer risk to change mobile phoning habits."

In summary, the current website of the World Health Organization [3,10] has the following statements:

"Despite extensive research, to date there is no evidence to conclude that exposure to low level electromagnetic fields is harmful to human health."

"To date, all expert reviews on the health effects of exposure to RF fields have reached the same conclusion: There have been no adverse health consequences established from exposure to RF fields at levels below the international guidelines on exposure limits published by the International Commission on Non-Ionizing Radiation Protection (ICNIRP, 1998)."

REFERENCES

1. ICNIRP. International Commission on Non-Ionizing Radiation Protection, "Guidelines for limiting exposure to time-varying electric, magnetic, and electromagnetic fields (up to 300 GHz)," *Health Physics*, Vol. 74, No. 4, 494–522, 1998.
2. WHO database: http://www.who.int/peh-emf/en/.

3. World Health Organization, Conclusions from Scientific Research, http://www.who.int/peh-emf/about/WhatisEMF/en/index1.html.
4. Lahkola, A, K. Tokola, and A. Auvinen "Meta-analysis of mobile phone use and intracranial tumors," *Scand J Work Environ Health*, Vol. 32, 171–177, 2006.
5. Heynick, L. N., S. A. Johnston, and P. A. Mason, "Radio frequency electromagnetic fields: Cancer, mutagenesis, and genotoxicity," *Bioelectromagnetics*, Supplement 6, S74–S100, 2003.
6. Elder, J. A., "Survival and cancer in laboratory mammals exposed to radiofrequency energy," *Bioelectromagnetics*, Supplement 6, S101–S106, 2003.
7. Dasenbrock, C., "Animal carcinogenicity studies on radiofrequency fields related to mobile phones and base stations," *Toxicology and Applied Pharmacology*, Vol. 207, 342–346, 2005.
8. "Reviews of the effects of RF fields on various aspects of human health," *Bioelectromagnetics*, Supplement 6, S1–S213, 2003.
9. IEEE C95.1-2005, "IEEE standard for safety levels with respect to human exposure to radio frequency electromagnetic fields, 3 kHz to 300 GHz," *IEEE*, 237, 3 Park Avenue, New York, NY 10016-5997, USA.
10. World Health Organization, Children and Mobile Phones: Clarification Statement, 2005. http://www.who.int/peh-emf/meetings/ottawa_june05/en/index4.html.
11. Independent Expert Group on Mobile Phones (IEGMP), Mobile Phones and Health, Chairman Sir William Stewart, National Radiological Protection Board, Chilton, Didcot, Oxon OX11 0RQ, United Kingdom, April 2000.
12. German Commission on Radiological Protection (SSK), Limit values and Precautionary Measures to Protect the Public against Electromagnetic Fields, September 2001.
13. French Environmental Health and Safety Agency (AFSSE), Statement on Mobile Phones and Health, Ref. 1/2002, 16 April 2003.
14. Australian Communications Authority (ACA). Consumer Fact Sheet. Electromagnetic Radiation and 3G Mobile Phones, December 22, 2003.
15. Advisory Group on Non-Ionising Radiation (AGNIR), "Health effects from radiofrequency electromagnetic fields," Report of an Independent Advisory Group on Non-ionising Radiation, Vol. 14, No. 2, National Radiological Protection Board, 2003.
16. Health Council of the Netherlands (HCN), *Electromagnetic Fields: Annual Update 2003*, Electromagnetic Fields Committee, Health Council of the Netherlands, The Hague, Publication 2004/01, January 15, 2004.
17. National Radiological Protection Board (NRPB), Mobile Phones and Health 2004, Vol. 15, No. 5, Chilton, Didcot, Oxon OX11 0RQ, United Kingdom, 2004.
18. Food and Drug Administration (U.S. FDA), FDA Cell Phone Facts: Consumer Information on Wireless Phones http://www.fda.gov/cellphones/, January 13, 2005.
19. European Cancer Prevention Organization (ECPO), Cell Phones and Cancer, ECP Consensus EMF 2005. http://www.ecpo.org/.

Unusual Transimission Properties of Waves in One-dimensional Random System Containing Left-handed Material

Yunxia Dong and Xiangdong Zhang
Department of Physics, Beijing Normal University, 100875, China

Abstract— Propagation properties of electromagnetic waves in a one-dimension random system containing left-handed-material are studied by the transfer matrix method. The statistics of the Lyapunov exponent and its variance of the transmitted waves are also analyzed. The nonlocalized modes are not only found in such a disordered system, the Anderson localization states with short localization length can also be easily realized due to the existence of low frequency resonant gap. Furthermore, our results also show that a single-parameter scaling is generally inadequate even for the complete random system with left-handed-material when the frequency we consider is located in a gap.

1. INTRODUCTION

During the past decade wave propagation in random media has been the subject of intensive studies due to its fundamental importance and potential application in photonics, medical, and biological research [1]. Random multiple scattering of light waves shows many similarities to the propagation of electrons in solid. Various electron transport phenomena such as Anderson localization [2, 3], coherent backscattering or weak localization [4] and universal conductance fluctuations [5] have now also been found to exist for light waves in disordered dielectrics. The investigations of transport behavior in disordered 1D photonic crystal (PC) have also shown that there can exist two types of localized states inside a band gap [6]. The first type do not satisfy the single-parameter scaling law in the case of small disorder and when the disorder is large (complete random), all the states inside the gap belong to the normal-type localized states, satisfying the single-parameter scaling theory [6].

The above investigations all focus on random dielectric or metal system. Recently, negative refraction and left-handed materials (LHMs) have attracted a great deal of attention from both the theoretical and experimental sides due to their unusual wave propagation phenomena [7–9]. Multilayered periodic structures that include LHMs have also been analyzed [10–14]. A natural question to ask is what kind of phenomenon will occur when the wave transmits through a random system produced by the above periodic PC including the LHMs? In order to answer this interesting question, in this work, we investigate the transmission and statistical properties of the wave through such a random system.

2. RANDOM SYSTEMS AND THEORY

We consider a randomly layered sample with random thickness for each layer. Such a disordered system can be obtained by completely randomizing the ordered PC, which consists of the alternate layer of air $\varepsilon = \mu = 1$ and the negative-n material, with $\varepsilon(\omega)$ and $\mu(\omega)$ given by [14]

$$\begin{aligned} \varepsilon(f) &= 1 + \frac{5^2}{0.9^2 - f^2} + \frac{10^2}{11.5^2 - f^2}\,, \\ \mu(f) &= 1 + \frac{3^2}{0.902^2 - f^2} \end{aligned} \tag{1}$$

here f is the frequency measured in GHz. Numerical values of $\varepsilon(\omega)$ and $\mu(\omega)$ are given in Fig. 1(a). The wave transmission in such a system can be tackled in exact manner by the transfer-matrix method [6]. Therefore, the localization length (ξ) of wave transmitting through such a finite disorder system can be obtained as

$$\xi = \frac{2\,(L_2 - L_1)}{<\ln T_1> - <\ln T_2>}\,. \tag{2}$$

Where L_1 and L_2 are the different thicknesses of the samples, and T_1 and T_2 are the corresponding transmission coefficients. $<\ln T>$ represents configurational average of the transmissivity (T) for the disordered sample with thickness L [15]. Based on them, the statistical properties of the wave transmitting through the above disorder system such the Lyapunov exponent (LE) and its variance var (γ) can also be obtained.

3. NUMERICAL RESULTS AND DISCUSSIONS

We first calculate the band structure of the above disordered system, which the results are shown in Fig. 1(b). The solid line in Fig. 1(c) represents the transmittance through a stack of 25 unit cells. The widths of the air and the negative-n layers are taken as 10 mm. The band structure and the transmittance clearly show two band gaps. The first is the resonant gap of the low frequency and the second is the general Bragg gap, which are agreement with those in Ref. [14]. In order to make a comparison, we also consider 1D PC consisting of the air and the pure dielectric material. That is to say, the negative-n layers in the above system are replaced by the dielectric layers with $\varepsilon = 3.4$. It is well known that only Bragg gaps exist in such a case.

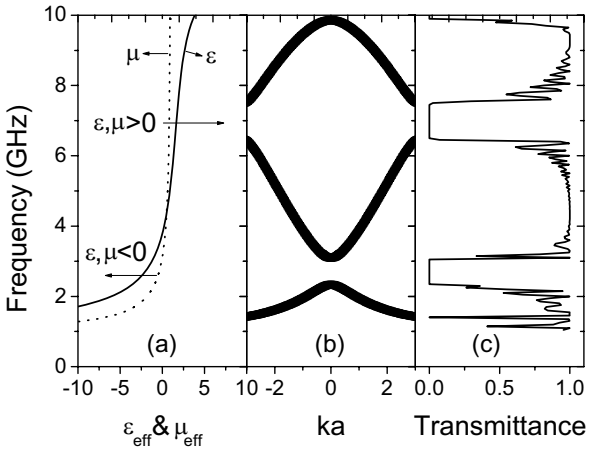


Figure 1: (a) Effective ε and μ of the negative-n material, as given by Eq. (5); (b) Dispersion relationship of a photonic crystal with alternate layers of air (10 mm thick) and the negative-n material (10 mm thick) with material parameter as shown in (a); (c) Transmittance through 25 unit cells, corresponding to the band structure in (b).

In the following, we introduce disorder in the above two periodic structures. Thus, the width of the nth layer (negative-n or dielectric) is taken to be random variable $a_n = a_0(1 + \delta)$, where δ is a random number between $(-1.0, 1.0)$. We first calculate the transmission coefficients of the wave through two disordered systems with different configurations. The average transmittances $<\ln T>$ with 1000 different configurations for each frequency are plotted in Fig. 2(a) and (b). Fig. 2(a)

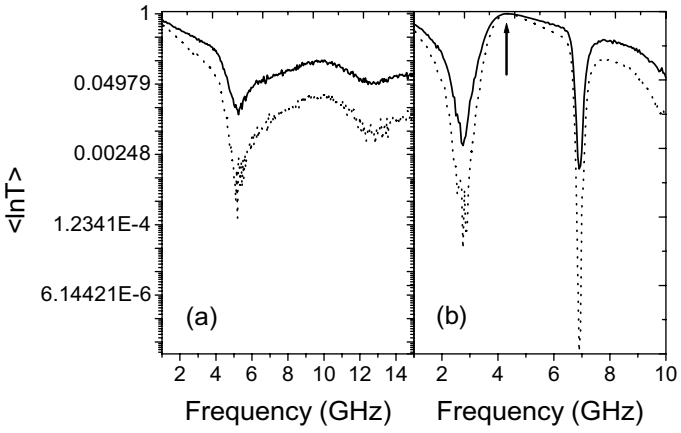


Figure 2: The average transmission $<\ln T>$ through a complete random photonic crystal with the dielectric materials (a) and the negative-n materials (b). Solid lines and dotted lines correspond to the systems with 50 and 100 layers, respectively. The other parameters are identical to those above.

represents the case with pure dielectric layers and (b) for that with negative-n layers. Solid lines correspond to the case of 50 layers (0.5 m) and dotted lines to the case of 100 layers (1.0 m).

For the case of pure dielectric layer, the average transmittance decreases with the increase of the sample thickness for all frequencies and the localizations always occur for all frequencies with

the increase of the size of the sample. The situation becomes different for that with the negative-n layer. We find that a high-transmission peak, which does not depend on the thickness of the disordered sample, always exists in this case. The position of the peak corresponds to the frequency with 4.35 GHz. Both of the effective permittivity and permeability are 0.5, which has been marked as an arrow in Fig. 2(b). The value of the transmission coefficient corresponding to the peak is close to 1. Such a nonlocalized mode is originated from the zero reflection of wave on the interface of the dielectric and negative-n layers [15].

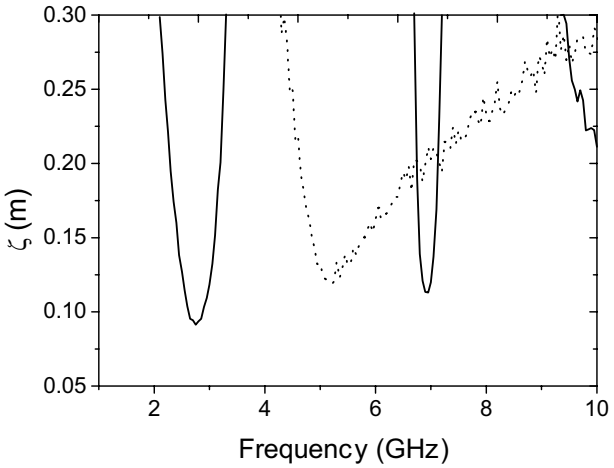


Figure 3: Localization length (ξ) as a function of frequency. Solid line and dotted line correspond to the complete random systems with the negative-n materials and the pure dielectric materials, respectively. The other parameters are identical to those in Fig. 2.

In fact, the localization of the wave in such a system can also exhibit some new features. First, it is easier to obtain strong Anderson localization with the short localization length due the existence of the low frequency gap. The localization length as a function of the frequency for such a case is plotted as solid line in Fig. 3. The minimum localization length corresponding to the first gap is about 0.09 m. It is significantly smaller than that of the pure dielectric case (dotted line in Fig. 3). The reason is that the low frequency resonant gap is insensitive to the disorder. That is to say, the strong Anderson localization is easer to be observed in such a disordered system.

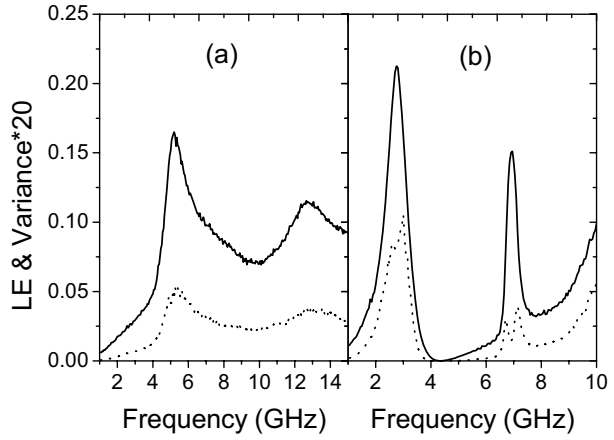


Figure 4: Lyapunov exponent (solid lines) and its variance (dashed lines) as a function of frequency for the case with the dielectric materials (a) and the negative-n materials (b). The other parameters are identical to those in Fig. 2.

Furthermore, the statistical properties of the wave through the system with the negative-n materials (case I) are also different from those of the dielectric case (case II). The average LE and its variance as a function of the frequency are plotted in Fig. 4 as solid lines and dotted lines, respectively. The results are obtained by the average of 1000 different configurations for each

frequency. Figs. 4(a) and 4(b) represent the results of case II and case I, respectively. In the case II, the changes of LE and it variance as a function of the frequency are identical. In the case I, as a function of frequency, the variance exhibits a dip inside the gap between two maxima. The presence of such a dip shows that the single-parameter scaling is inadequate in such a case even for complete random. The physical origination can still be understood by the existence of the resonance, which the random systems including the negative-n materials can not be simply homogenized as those with only dielectric components.

In the above calculations, we do not consider the effect of absorption. To study the effects due to the absorption, we introduce an imaginary dielectric constant ($i\varepsilon''$) for each negative-n layer. And we find that the statistical properties do not change with the introduction of the absorption, although the values of the variances decrease. In fact, other propagation properties of wave through the disordered negative-n system in the presence of the absorption have also been investigated, similar conclusions have been obtained.

4. CONCLUSIONS

We have investigated the propagation and statistical properties of the electromagnetic waves in a one-dimension random system containing left-handed-material by the transfer matrix method. The average transmission coefficient, localization length, Lyapunov exponent and its variance of the transmitted waves have been calculated for two disordered systems (pure dielectric and negative-n). We find that the nonlocalized modes exist in the disordered system with negative-n materials. Furthermore, the Anderson localization states with short localization length can also be easily obtained in such a system due to the existence of the low frequency resonant gap. Our results also show that the single-parameter scaling is generally inadequate even for the complete disordered system with negative-n materials when the frequency we consider is located in the gap. In addition, the effect of the absorption on these unusual properties has also been discussed.

REFERENCES

1. See, for example, *Scattering and Localization of Classical Waves in Random Media*, edited by P. Sheng, World Scientific, Singapore, 1990.
2. John, S., *Phys. Rev. Lett.*, Vol. 53, 2169, 1984; *Phys. Rev. Lett.*, Vol. 58, 2486, 1987.
3. Genack, A. Z. and N. Garcia, *Phys. Rev. Lett.*, Vol. 66, 2064, 1991; Vol. 66, 1850, 1991.
4. Kuga, Y. and A. Ishimaru, *J. Opt. Soc. Am. A*, Vol. 8, 831, 1984; Albada, M. v. and A. Lagendijk, *Phys. Rev. Lett.*, Vol. 55, 2692, 1985; Wolf, P. and G. Maret, *Phys. Rev. Lett.*, Vol. 55, 2696, 1985.
5. Scheffold, F. and G. Maret, *Phys. Rev. Lett.*, Vol. 81, 5800, 1998.
6. Deych, L. I., D. Zaslavsky, and A. A. Lisyansky, *Phys. Rev. Lett.*, Vol. 81, 5390, 1998; Deych, L. I., A. A. Lisyansky, and B. L. Altshuler, *Phys. Rev. Lett.*, Vol. 84, 2678, 2000; *Phys. Rev. B*, Vol. 64, 224202, 2001; Luan, P.-G. and Z. Ye, *Phys. Rev. E*, Vol. 63, 066611, 2001.
7. Veselago, V. G., *Sov. Phys. Usp.*, Vol. 10, 509, 1968.
8. Smith, D. R., W. J. Padilla, D. C. Vier, S. C. Nemat-Nasser, and S. Schultz, *Phys. Rev. Lett.*, Vol. 84, 4184, 2000; Shelby, R. A., D. R. Smith, and S. Schultz, *Science*, Vol. 292, 77, 2001.
9. Pendry, J. B., *Phys. Rev. Lett.*, Vol. 85, 3966, 2000.
10. Zhang, Z. M. and C. J. Fu, *Appl. Phys. Lett.*, Vol. 80, 1097, 2002.
11. Ziolkowski, R. W. and E. Heyman, *Phys. Rev. E*, Vol. 64, 056625, 2001.
12. Nefedov, I. S. and S. A. Tretyakov, *Phys. Rev. E*, Vol. 66, 036611, 2002.
13. Wu, L., S. L. He, and L. F. Shen, *Phys. Rev. B*, Vol. 67, 235103, 2003.
14. Li, J., L. Zhou, C. T. Chan, and P. Sheng, *Phys. Rev. Lett.*, Vol. 90, 083901, 2003.
15. Dong, Y. and X. Zhang, *Phys. Lett. A*, Vol. 359, 542, 2006.

Numerical Simulation and Analysis on Mode Property of Photonic Crystal Fiber with High Birefringence by Fast Multipole Method

Wei Song, **Yuanyuan Zhao**, **Yu Bao**, **Shuguang Li**, **Zhiyuan Zhang**, **and Tianfu Xu**

Physics Department, College of Science, Yanshan University, Qinhuangdao 066004, China

Abstract— In this paper the mode property of a photonic crystal fiber with triangle-lattice array in a silica matrix has been simulated by fast multipole method. The PCFs studied in this paper have a silica core, obtained by introducing a defect, that is by removing three holes in the center of the fiber transverse section. The model fields, effective index and confinement loss about the fundamental mode in the fibre are analysed and compared. It is demonstrated that lower confinement loss and higher birefringence can be realized in the condition of fewer rings of air holes. At the wavelength of 1.55 μm the confinement loss is 7.3×10^{-6} dB/m and the birefringence is 1.65×10^{-3} in this fiber. The simulation results show that birefringence of this triangle-lattice PCF is dominated by inner-ring air holes in the fibre effectively. The simulation results in this paper have important meaning for instructing the fabrication of birefringent photonic crystal fiber.

1. INTRODUCTION

In recent year, great interest has been focused on the study of photonic crystal fiber (PCF). By using the structure adjustability of the PCF, zero dispersive wavelength moving towards short-wave spectroscopy, chromatic dispersion compensating, dispersion truncating, high nonlinearity and birefringence can be realized [1–15]. The mode birefringence of PCF mainly comes from the geometer structure of fiber and the usual method to cause birefringence is employing double-core or multi-core structure, changing the fiber core or the shape of the air holes and altering the distribution of the air holes. High birefringence PCF can be used in optical fiber sensor, interferometer and polarizer, etc. In addition, by designing the high birefringence and intensive nonlineality PCF required, we can make a fiber not only high birefringence, but also intensive nonlineality. We can also make the large mode property polarization maintaining PCF [15]. This can make the property of PCF integrate, so it can be used in the Raman magnification, ex-continous spectum of the polarization, four-wave frequency mixing and the crossing phase modulation. The development of the birefringence PCF will promote the study and appliance of the new photoelectric device. In the progress of making the PCF fiber, we find that maintaining the ideal structure of the fiber and the number of air holes layers of the PCF cladding is a contradiction [2, 4, 17]. Generally speaking, the more layers the PCF cladding have, the smaller the confinement loss will be, but at the same time increasing the layers will make the distortion of the air holes a big problem when the fibers are produced. On the fundament of the traditional triangular arrayed air holes, this article proposes a birefringence structure PCF which has C_{2v} symmetry. And by using the fast multipole method, its modal property is numerically simulated. It is found that in this PCF, there will be several conductive modes, and its fundamental mode has very low confinement loss and high birefringence.

2. BASIC THEORY

The ordinary solution about multipole theory at the complex boundary electrostatic field has been deduced and developed by Zheng Qinhong. Nie Zaiping [22] analyzed the three dimensional vector scattering of complicated object by using the fast multipole method. These examples show that the multipole method is an effective way to analyze electromagnetic field theory. Using the multipole method to simulate the dispersion and the loss property of microstructured fiber is originally proposed by T. P. White and B. T. Kuhlmey [23–25] in Sydney University. Zhao Mingzhu [26] also simulated the PCF by using multipole method.

This method is adopted when the air holes of microstructured fiber is cylindrical. By using this method, we can get real part and imaginary part of the effective refractive index of the fiber and the mode propagation constant. And then by using the real part, we can analyze the dispersion, or using the imaginary part to analyze the confinement loss of limited cladding air holes. We can also use this method to get its propagation constant corresponding to the input wavelength (or frequency). Accordingly, we can consider the material dispersion in the progress of simulation by using the Sellmeier formula. In this paper, the theory of T. P. White [23, 24] is used for reference

and fast multipole method is much adopted. Using different multipole cutoff order to meet the varied diameter of air holes and wavelength ratio, we not only guarantee the calculation precision, but also speed the computation velocity. So fast multipole method is a very efficient way to simulate the PCF property.

The electromagnetic field in the fiber can be decomposed to two types. One is axis Z along the fiber: E_Z and H_Z. The other is sectional part: E_t and H_t. In fact, when magnetic field multiple a parameter, it will have the same Maxwell equation form to electic field, that is $K = ZH$, where Z is impedance of free space. Electromagnetic field can be demonstrated as follows:

$$\vec{E} = \vec{E}(r,\phi)\exp[i(\beta z - \omega t)] \tag{1}$$

$$\vec{K} = \vec{K}(r,\phi)\exp[i(\beta z - \omega t)] \tag{2}$$

where ω is angular frequency, propagation constant b is a plural. The imaginary part of b denotes the attenuation along the Z-axis.

The longitudinal part of electromagnetic field ($V = E_z$ or $V = K$) meet the Helmholtz equation

$$(\nabla^2 + (k_\perp)^2)V = 0 : \tag{3}$$

$k_\perp$ is shown by $k_\perp^l$ and $k_\perp^e$ respectively in the air holes and base material. They could be expanded to cylindrical function. In the lth air hole, its vertical electric field E_Z can be expanded in the cylindrical coordinate:

$$E_z = \sum_{m=-\infty}^{\infty} a_m^{(l)} J_m\left(k_\perp^i r_l\right)\exp(im\phi_l)\exp(i\beta z) \tag{4}$$

In the medium next to the lth air hole, the vertical electric field can be shown as follow:

$$E_z = \sum_{m=-\infty}^{\infty}\left[b_m^{(l)} J_m\left(k_\perp^e r_l\right) + c_m^{(l)} H_m^1\left(k_\perp^e r_l\right)\right]\times\exp(im\phi_l)\exp(i\beta z) \tag{5}$$

$k_T = k_\perp^i = (k_0^2 n_i^2 - \beta^2)^{1/2}$, $k_\perp^e = (k_0^2 n_e^2 - \beta^2)^{1/2}$, the refractive index of air n_i =1, n_e is the refractive index of quartz material, k_0 is the wavevector in free space, r_L and j_L are the coordinate of regional coordinate system $\vec{r}_l(r_l,\phi_l) = \vec{r} - \vec{c}_l$, $\vec{c}_l$ is the center of the air hole. The expression of magnetic field part K_z is similar to the electric field. Be careful when compute the squadratic root of the plural number, as to the mode which is attenuated traveling along the +Z axis, the real part and imaginary part of the propagation constant must meet this rules: $\Re(\beta) > 0$, $\Im(\beta) > 0$.

By using the boundary condition of electromagnetic field on the interface of the air hole, we can get the expression of $a_m^{(l)}$, $b_m^{(l)}$ and $c_m^{(l)}$. In the real computation by choosing the proper cutoff value corresponding to the air hole diameter of the optical fiber cladding and the wavelength ratio M, we can optimize speed and precision of the calculation. Further more, taking in the concept of fast multipole method to programme can increase the computation speed. The effective refractive index of the mode n_{eff} can be obtained by propagation constant b.

Then, the imaginary part of the n_{eff} can be used to get the fiber confinement loss (the unit is dB/m).

$$L = \frac{20}{\ln(10)}\frac{2\pi}{\lambda}\Im(n_{eff})\times 10^6 \tag{6}$$

The unit of l is micrometer (mm), the dispersion coefficient can be got from the real part

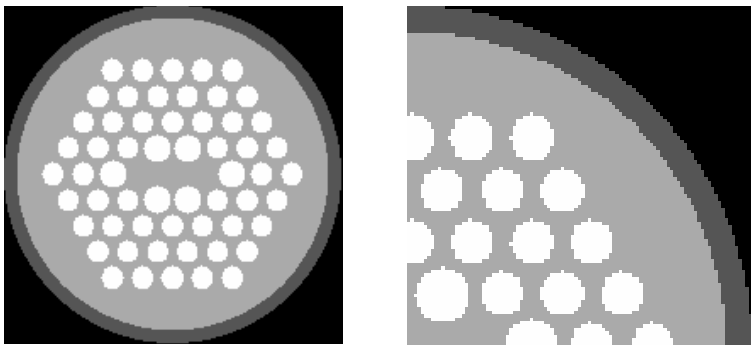
$$D = -\frac{\lambda}{c}\frac{d^2\Re(n_{eff})}{d\lambda^2} \tag{7}$$

The birefringence index B [9–11] is:

$$B = n_{eff}^s - n_{eff}^f \tag{8}$$

n_{eff}^s and n_{eff}^f are the two cross polarization of the fundamental mode which is corresponding to the slow axis and fast axis. The polarization beat length between the two cross polarization is:

$$L_B = \lambda/B \tag{9}$$



(a) (b)

Figure 1: (a) The section of birefringence PCF and (b) minimal sector.

3. SIMULATION AND ANALYSIS

The section of birefringence PCF simulated in this paper is shown in Figure 1. PCF cladding is composed of regular triangular arrayed air holes. Three holes along the center of X axis are removed to form the fiber core. The pitch $\Lambda = 2.3\,\mu\text{m}$, air-holes diameter $d = 1.6\,\mu\text{m}$. In the most inner layer, there are 6 more bigger air holes, their diameter is $d_0 = 2.0\,\mu\text{m}$. As far as the structure of the fiber is concerned, on one hand, because the fiber core can generate birefringence for its asymmetrical in the X axis and Y axis direction; on the other hand according to the theory of group, this structure has the $C_{\text{v}2}$ symmetry. Taking advantage of these symmetries can simplify the calculation and enhance the precision and speed of computation. The minimal sector is $\phi = 0 \sim \pi/2$

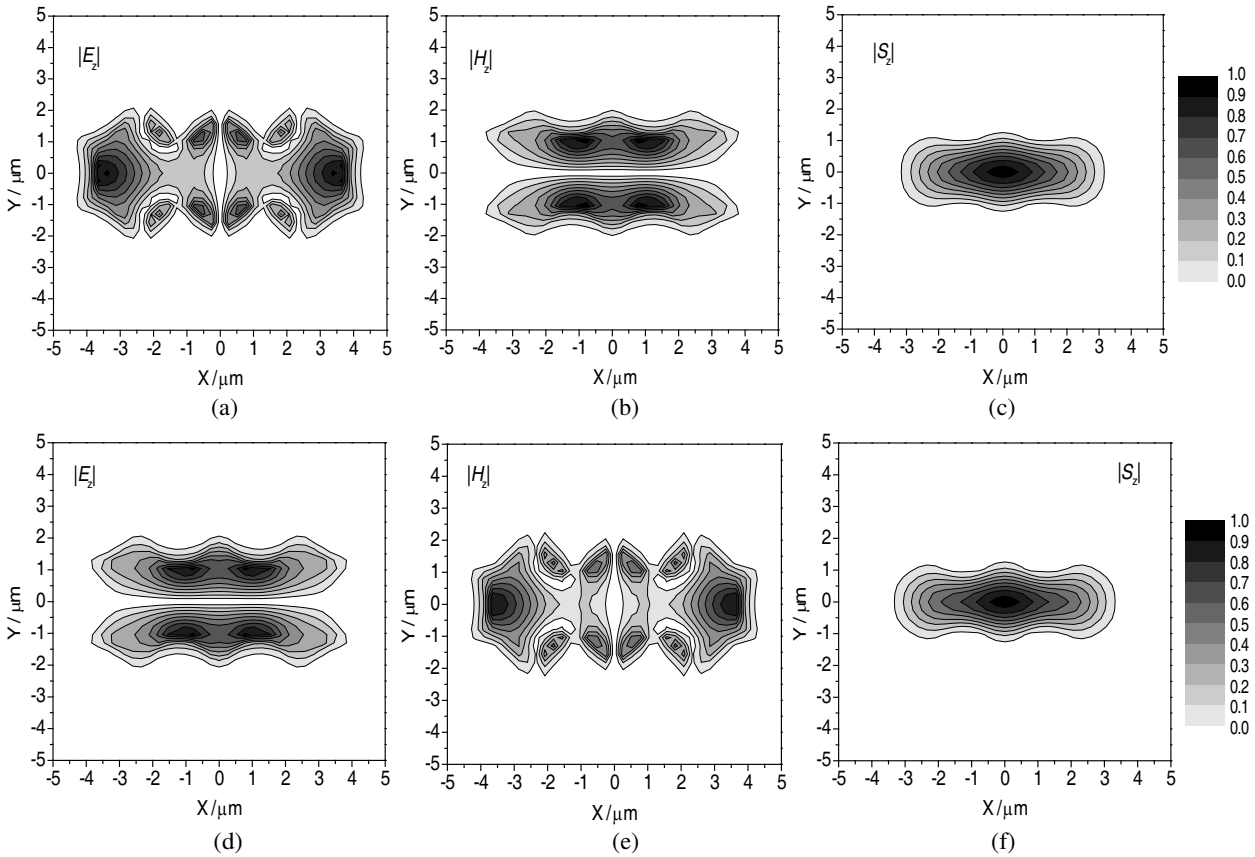


Figure 2: The fundamental mode distribution of the birefringence PCF at the wavelength ($\lambda = 1.55\,\mu\text{m}$), where figures (a) and (b) and (c) respectively stand for their slow axis mode field $|E_{\text{z}}|$ and $|H_{\text{z}}|$ and $|S_{\text{z}}|$, (d) and (e) and (f) stand for their fast axis mode field $|E_{\text{z}}|$ and $|H_{\text{z}}|$ and $|S_{\text{z}}|$ respectively.

when simulated.

Figure 2 shows the fundamental mode distribution of normalized mode field of the PCF shown in Figure 1, where figures (a), (b) and (c) respectively stand for their slow axis mode field $|E_z|$ and $|H_z|$ and $|S_z|$, (d), (e) and (f) stand for their fast axis mode field distribution $|E_z|$ and $|H_z|$ and $|S_z|$ respectively. From these pictures, it can be seen that every mode's $|E_z|$, $|H_z|$, $|S_z|$ distribution has the C_{2v} symmetry. The $|S_z|$ of the fundamental mode has a peek center. The modal field of the fundamental mode distribution shows that $|E_z|$ of slow axis and $|H_z|$ of fast axis have similar mode field distribution and the $|H_z|$ of the slow axis and $|E_z|$ of fast axis have the similar mode field distribution too. The Z part of Poynting vector of the slow axis and fast axis has similar distribution of $|S_z|$.

Figure 3 gives the effective refractive index and Fundamental Space Mode [5, 6] (FSM) of the fundamental mode of PCF. In the figure, curve1 and 2 represent effective refractive index of slow axis and fast axis of fundamental mode. FSM represents the effective refractive index of cladding fundamental space mode. It can be seen from this figure that the effective refractive index of the fundamental mode of slow axis is bigger than the fast axis.

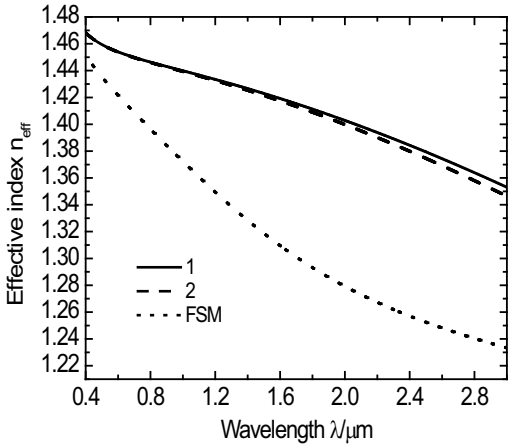


Figure 3: The effective index of PCF of fundamental and cladding mode changing corresponding to the wavelength.

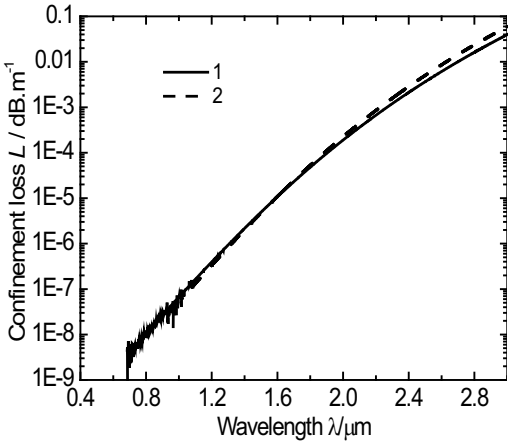


Figure 4: The confinement loss of PCF fundamental mode corresponding to wavelength.

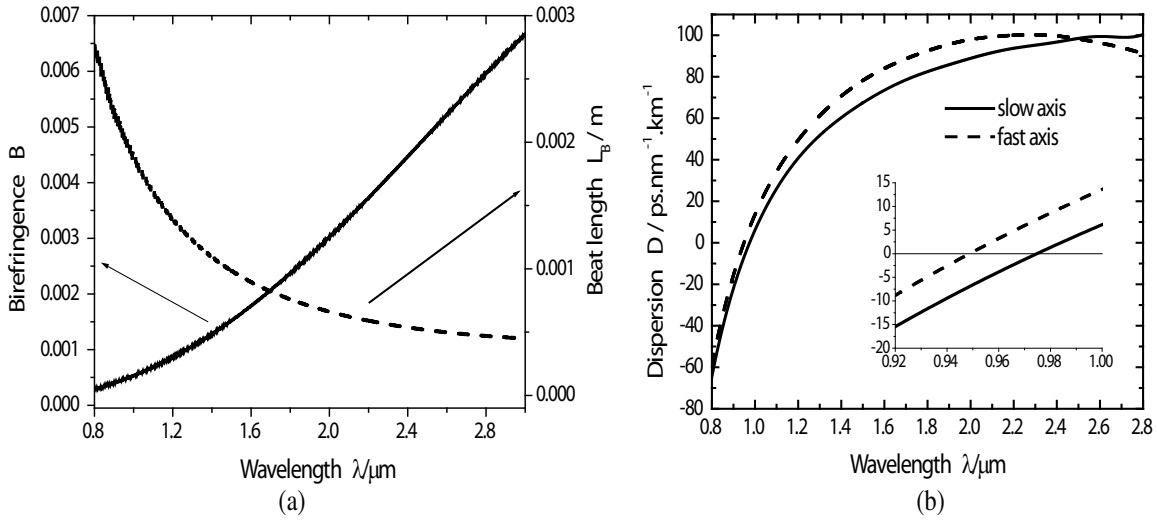


Figure 5: (a) the fundamental mode birefringence B and beat length L_B of PCF as functions of wavelength. (b) the dispersion coefficient D of slow axis and fast axis as function of wavelength.

Figure 4 shows the changing rule of the confinement loss L corresponding to the wavelength of the PCF fundamental mode. In the figure, curve1 and 2 represent the confinement loss of the slow axis and fast axis. It can be seen in this figure that in the low loss window where=1.55 μm of optical

communication, the slow axis mode and the fast axis mode are 7.3×10^{-6} and 7.2×10^{-6} dB/m. This loss is very small. The confinement loss of fundamental mode at 1.55 μm is far bellow the traditional loss 0.2 dB/km in optical communication. We can also see that the fiber's fundamental mode of short infrared and middle infrared band (0.8 μm∼3.0 μm) we simulate have confinement loss bellow 0.06 dB/m. If there is another layer of air hole, its loss can be lower. So this design of structure is certainly a type of PCF with low confinement loss.

Figure 5(a) shows the rule of fundamental mode birefringence B and beat length L_B of PCF as functions of wavelength. Figure 5(b) gives the dispersion coefficient D of slow axis and fast axis as functions of wavelength. At 1.55 μm, birefringence of fundamental mode is $B = 1.65 \times 10^{-3}$. This belongs to high birefringence fiber. Between infrared band, the birefringence is increased with the increasing of wavelength. Correspondingly, the beat length decreases with the increase of wavelength. It can be seen from the Figure 5(a) that the dispersion coefficient of fast axis module and slow axis corresponding to the fundamental mode has difference. The zero dispersion coefficient of fast axis and slow axis are 0.95 μm and 0.975 μm. So there is modal dispersion between fast axis and slow axis.

4. CONCLUSION

Based on traditional structure of triangular arrayed air holes, this paper promotes a type of birefringence PCF which has C_{2v} symmetry property. By using the fast multipole method, the fundamental mode distribution, effective refractive index and the confinement loss of this birefringence PCF are simulated. It is found that the structure of PCF only with several layers of the cladding air-holes can realize very low confinement loss. The result has instructive significance for the manufacture of the birefringence PCF. It can promote the progress of the manufacture of infrared band photoelectric device appliance by using PCF.

REFERENCES

1. Knight, J. C., T. A. Birks, et al., "All-silica single-mode optical fiber with photonic crystal cladding [J]," *Optics. Letters*, Vol. 21, 1547–1549, 1996.
2. Li, S., L. T. Hou, Y. L. Ji, and G. Y. Zhou, "Supercontinuum generation in holey microstructure fibres with random cladding distribution by femtosecond laser pulses [J]," *Chin. Phys. Lett.*, Vol. 20, No. 8, 1300–1302, 2003.
3. Li, S., X.-D. Liu, and L. T. Hou, "The study of waveguide mode and dispersion property in photonic crystal fibers [J]," *Acta Physica Sinica*, Vol. 52, No. 11, 2811–2817, 2003.
4. Li, S., Y. L. Ji, G. Y. Zhou, and L. T. Hou, "Supercontinuum generation in holey microstructure fibers by femtosecond laser pulses [J]," *Acta Physica Sinica*, Vol. 53, No. 2, 478–483, 2004.
5. Li, S., X.-D. Liu, and L. T. Hou, "Vector analysis of dispersion for fundamental cladding mode in photonic crystal fibers [J]," *Acta Physica Sinica*, Vol. 53, No 6, 1873–1879, 2004.
6. Li, S., X.-D. Liu, and L. T. Hou, "The study of dispersion compensating property in photonic crystal fibers [J]," *Acta Physica Sinica*, Vol. 53, No. 6, 1880–1886, 2004.
7. Chi, H., Q.-J. Zeng, et al., "Analysis on dispersion characteristics of photonic crystal fiber [J]," *J. Infrared Millim. Waves*, Vol. 22, No. 2, 149–153.
8. Li, S., X.-D. Liu, and L. T. Hou, "Numerical simulation and analysis on photonic crystal fibers with closing to zero and flattened chromatic dispersion [J]," *Chinese Journal of Lasers*, 2004, Vol. 31, No. 6, 713–717.
9. Peyrilloux, A., T. Chartier, A. Hideur, et al., "Theoretical and experimental study of the birefringence of a photonic crystal fiber [J]," *J. Lightwave Technology*, Vol. 21, No. 2, 536–539, 2003.
10. Kunimasa, S. and K. Masanori, "Photonic bandgap fibers with high birefringence [J]," *IEEE Photonics Technology Letters*, Vol. 14, 1291–1293, 2002.
11. Kerbage, C., P. Steinvurzel, P. Reyes, et al., "Highly tunable birefringent microstructured optical fiber [J]," *Optics letters*, Vol. 27, No. 5, 842–844, 2002.
12. Ortigosa-Blanch, A., A. Díez, M. Delgado-Pinar, et al., "Ultrahigh birefringent nonlinear microstructured fiber [J]," *IEEE Photonics Technology Letters*, Vol. 16, 1667–1669, 2004.
13. Ju, J., W. Jin, and M. S. Demokan, "Properties of a highly birefringent photonic crystal fiber [J]," *IEEE Photonics Technology Letters*, Vol. 15, 1375–1377, 2003.
14. Kakarantzas, G., A. Ortigosa-Blanch, T. A. Birks, et al., "Structural rocking filters in highly birefringent photonic crystal fiber [J]," *Opt. Lett.*, Vol. 28, 158–160, 2003.

15. Folkenberg, J. R., M. D. Nielsen, N. A. Mortensen, et al., "Polarization maintaining large mode area photonic crystal fiber [J]," *Opt. Express*, Vol. 12, 956–960, 2004.
16. Hoo, Y. L., W. Jin, C. Shi, H. L. Ho, D. N. Wang, and S. C. Ruan, "Design and modeling of a photonic crystal fiber gas sensor," *Applied Optics*, Vol. 42, No. 18, 3509–3515.
17. Zhou, G.-Y., Z.-Y. Hou, S.-G. Li, and L. T. Hou, "Mathematical model for fabrication of micro-structure fibres [J]," *Chin. Phys. Lett.*, Vol. 22, No. 5, 1162–1165, 2005.
18. Zheng, Q., "The multipole theory of the calculation of complex f electrostatic field," *Journal of University of Electronic Science and Technology of China*, Vol. 26, No. 6, 599–604.
19. Zheng, Q., H. Cao, and F. Xie, "Analysis of resonant cavit using multipole theory," *High Power Laser and Particle Beams*, Vol. 13, No. 1, 89–92.
20. Zheng, Q., W. Dai, F. Xie, and Y. Luo, "The calculation of two dimetional Poisson equation using multipole theory," *Journal of Yunnan Normal University*, Vol. 21, No. 2, 42–46.
21. Zheng, Q., F. Xie, and W. Lin, "Waveguide eigenvalue analysis using multipole theory," *Chinese Journal of Electrics*, Vol. 28, No. 3, 112–114.
22. Nie, Z., J. Hu, H. Yao, and H. Wang, "Analysis of chromatic dispersion of complex object three dimetional using fast multipole theory," *Chinese Journal of Electrics*, Vol. 27, No. 6, 105–109.
23. White, T. P., B. T. Kuhlmey, R. C. McPhedran, et al., "Multipole method for microstructured optical fibers. I. Formulation [J]," *J. Opt. Soc. Am. B*, Vol. 19, 2322–2330, 2002.
24. Kuhlmey, B. T., T. P. White, G. Renversez, et al., "Multipole method for microstructured optical fibers. II. Implementation and results [J]," *J. Opt. Soc. Am. B*, Vol. 19, 2331–2340, 2002.
25. White, T. P., R. C. McPhedran, L. C. Botten, et al., "Calculations of air-guided modes in photonic crystal fibers using the multipole method [J]," *Optics Express*, Vol. 9, 721–732, 2001.
26. Zhu, Z. and T. G. Brown, "Multipole analysis of hole-assisted optical fibers [J]," *Optics Communications*, Vol. 206, 333–339, 2002.

A Comparison of Performance of Four Methods in Solving Time Domain Integral Equations for Arbitrarily Shaped Conducting Bodies

Xinpu Guan, Shaogang Wang, Yi Su, and Junjie Mao
The Third Department of School of Electronic Science and Technology
National University of Defense Technology, Changsha 410073, China

Abstract— The objective of this paper is to generate a temporal response of three dimensional conducting structures by solving time domain electric-field integral equation. Four methods including Initial MOT/ averaging MOT/ IMPT/ MOO are used to solve the integral equation. Comparison of performance of four methods is presented to illustrate validity for solving late-time instability by simulation of four targets including two simple and two complex conducting bodies.

1. INTRODUCTION

Broadband analysis has been receiving much interest in studying many practical transient electromagnetic problems, with increasing interest in short pulse radar design for high resolution and target identification problems. The time-domain approach is preferred for broadband electromagnetic analysis because the program needs to be executed for each frequency in the frequency-domain. Electric field integral equation (EFIE) is widely used for the numerical analysis of electromagnetic radiation and scattering for its capability of handling either open or closed arbitrarily curved structures of finite extent. The most popular method to solve a TD-EFIE is the marching on in time (MOT) method [1] using triangular patches and vector basis functions proposed by rao-wilton-glisson (rwg). Like method of MOM [2], the MOT method discretizes the scatters or targets into segments for two-dimensional models or patches for three-dimensional models. And the time axis is divided into equal increments or time steps. A gaussian pulse is generally used as an incident source.

However, as pointed out by many researchers, some of the TD-EFIE formulations associated with the MOT method suffers from late-time instability, which usually takes the form of an exponentially increasing oscillation that alternates in sign at each time step. In order to eliminate the instability of the Mot methods, several formulations have been presented for the solution of the time-domain integral equation to calculate the electromagnetic scattering from arbitrarily shaped three dimensional structures using triangular patch modeling techniques. Rynne and Vechinski [3] proposed a averaging methods, Rao and Sarkar [4–6] introduced an implicit MOT technique, and Young-seek Chung and Sarkar [7–9] proposed a marching on in order method through expanding temporal coefficients of Hertz vector by weighted Laguerre polynomials. In this paper, four methods including Initial MOT/ Improved averaging MOT/ IMPT/ MOO are used again to solve the electric field integral equation, and the comparison of performance of four methods are implemented.

2. COMPARISON OF FOUR METHODS

Let S denote the surface of a closed or open perfectly conducting arbitrary shaped scatterer for which we wish to formulate the time-domain scattering problem. The incident wave induces a surface current $J(r,t)$, since the total tangential electric field is zero on the conducting surface at all times, we have

$$\left[\frac{\partial A}{\partial t}+\nabla\phi\right]_{\tan}=\left[E^i\right]_{\tan};\quad r\in S \tag{1}$$

where A and ϕ are the magnetic vector potential and the electric scalar potential given by

$$A(r,t)=\frac{\mu}{4\pi}\int_s\frac{J(r',\tau)}{R}ds';\quad \phi(r,t)=\frac{1}{4\pi\varepsilon}\int_s\frac{\delta(r',\tau)}{R}ds';\quad \tau=t-\frac{R}{c};\quad R=\left|r-r'\right|;$$

and where μ and ε are the permeability and permittivity of space, c is the velocity of wave propagation in that space, and r and r' are the arbitrarily located observation point and source point. The surface-charge density δ is related to the surface divergence of J by the equation of continuity.

$$\nabla\cdot J(r,t)=-\frac{\partial\delta(r,t)}{\partial t}; \tag{2}$$

The purpose of four methods which we mentioned is to solve the formulation (1), and to obtain surface current density $J(r,t)$ and far scattered electric field $E^s(r,t)$. All these time domain methods need to separate the spatial and the time variables and treat each one of them independent of the other. We do not introduce the detailed procedure of four methods here, but discuss their difference and characteristic.

1) In basic MOT method (including initial MOT and the averaging method), the second order time derivatives of the vector potential and the first order time derivatives of the incident impulse plane wave are approximated by finite differences. In implicit method, only the first order time derivatives is approximated by first order finite differences. In MOO method, the first and second order time derivatives of hertz vector are expressed analytically.

2) The expanded form of the electric current J on the scattering structure is different. In the basic MOT method, the current J on S is approximated as

$$\bar{J}(\bar{r},t) \cong \sum_{j=-\infty}^{\infty} \sum_{n=1}^{N} I_{n,j} \vec{f}_n(\bar{r}) T_j(t), \tag{3}$$

where $\vec{f}_n(r)$ and $T_j(t)$ are spatial basis functions and temporal basis functions, each unknown coefficient $I_{n,j}$ represents the value of the component of the surface current normal to the nth edge at the time instant t_j. In the implicit method, the current is approximated as

$$\bar{J}(\bar{r},t) \cong \sum_{n=1}^{N} I_n(t) \vec{f}_n(\bar{r}), \tag{4}$$

$I_n(t)$ represents temporal function. In the MOO method, the current is approximated as

$$J(r,t) = s \sum_{n=1}^{N} \sum_{u=0}^{\infty} (0.5 P_{n,u} + \sum_{i=0}^{u-1} P_{n,i}) \varphi_u(\bar{t}) \vec{f}_n(r), \tag{5}$$

where s is the time scale factor, $\varphi_u(\bar{t})$ is the Weighted laguerre polynomials, and $P_{n,u}$ are the coefficients of weighted laguerre polynomials which are used to expand the temporal function of the hertz vector.

3) The recursion formula and iterative process of four methods are different. In basic MOT method, the explicit iterative scheme is used by marching on in time to compute unknown coefficient $I_{n,j+1}$ through the recursion formula: $\alpha_m \cdot I_{n,j+1} = V_{m,i} - \beta_{m,j}$. The averaging method is based on the initial MOT method and using formulation $\tilde{I}_{m,j} = (\tilde{I}_{m,j-1} + 2I_{m,j} + I_{m,j+1})/4$ to obtain an approximate averaged value $\tilde{I}_{m,j}$. In implicit method, assume the current as an approximately linear function, when $t_{j-1} < \tau_{mn}^{pq} < t_j$, where $t_j \le t_{I-1}$, $I_n(\tau_{mn}^{pq})$ can be expressed as

$$I_n(\tau_{mn}^{pq}) = (1-\delta) I_n(t_{J-1}) + \delta I_n(t_J), \quad \delta = (\tau_{mn}^{pq} - t_{j-1})/\Delta t; \tag{6}$$

When $t_{i-1} < \tau_{mn}^{pq} < t_i$, we have

$$I_n(\tau_{mn}^{pq}) = S_{mn}^{pq} I_n(t_{i-1}) + (1 - S_{mn}^{pq}) I_n(t_I), \quad S_{mn}^{pq} = R_{mn}^{pq}/c\Delta t; \tag{7}$$

In the last, we can obtain a matrix equation as $[\alpha_{mn}][I_n(t_i)] = [\beta_m]$. In the MOO method, an explicit iterative scheme is used by marching on in order to calculate unknown coefficient $P_{n,u}$ through the recursion formula $[\bar{\alpha}_{mn}]\{P_{n,v}\} = \{V_m^v\} + \{\bar{\beta}_m^{v-1}\}$, $v = 0, 1, 2, \ldots$. Furthermore, in the last matrix equations of all methods, the inverse of the matrix needs to be calculated once at the beginning of the computation step in the implicit method and the MOO method.

4) In basis MOT methods, it requires that the time sampling interval satisfy $c \cdot dt \le R_{\min}$ to avoid the instability phenomena. But it is not restricted for the implicit method and MOO method.

3. NUMERICAL SIMULATIONS

In this section, four perfectly conducting structures are tested using forenamed methods in order to validate their stability performance. In all of the following examples, the incident electric field is assumed to be a linearly polarized plane wave with a gaussian profile in time of the form

$$E^i(r,t) = E_0 \frac{1}{T\sqrt{\pi}} e^{-[1/T(ct - ct_0 - r\cdot\hat{k})]^2}; \tag{8}$$

where T is the pulse width, $\hat{k}$ is the unit vector in the propagation direction of the incident wave, $\vec{r}$ is a position vector relative to the origin, c is the velocity of light in the air. For comparison, we present the results obtained by the IDFT of the frequency domain solution by MOM. In the implicit method, a central difference scheme is employed.

For the first example, we consider a PEC sphere with radius 1.0 m centered at the origin, as shown in Fig. 1. The number of triangular patches is 160 and the number of unknowns is 240. In this case, we choose the time step as $c \cdot \Delta t = 0.5R_{\min}$ for the basis MOT and $c \cdot \Delta t = 2.2R_{\min}$ for the implicit method and the MOO method, where $R_{\min} = 0.1551317$ m. We use $T = 1$ lm and $c \cdot t_0 = 6.0$ lm for all methods.

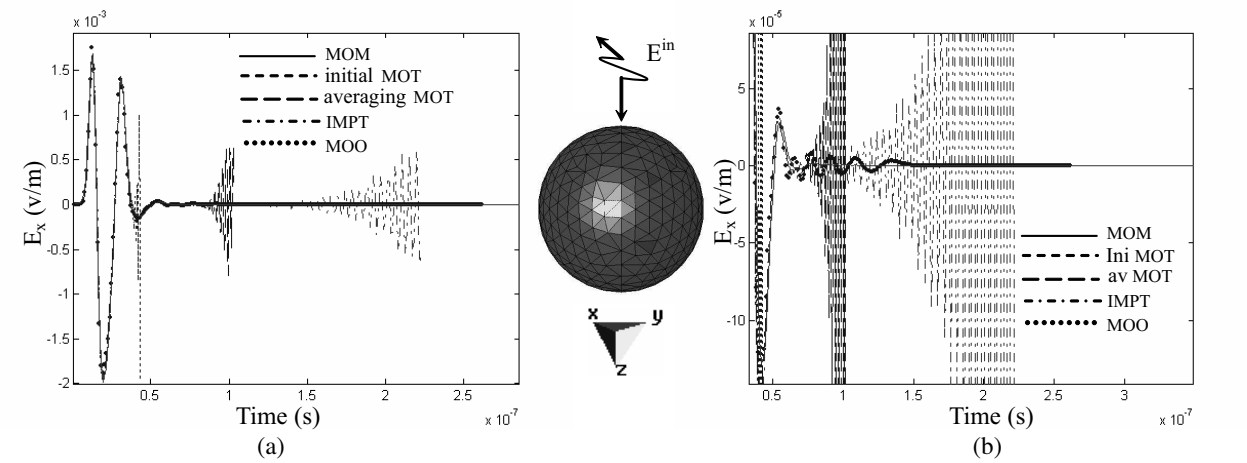


Figure 1: Backscattered far-field response of a conducting sphere 1.0 m in radius illuminated by a Gaussian plane wave. (a) x-component of the electric field at $(0, 0, -100)$ m. (b) enlarge late-time domain in Fig. (a).

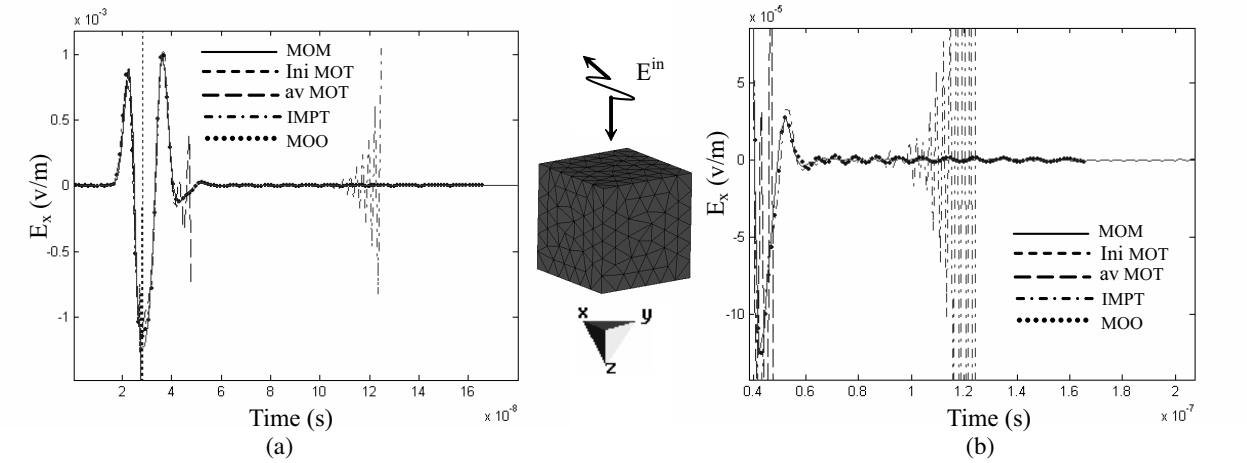


Figure 2: Backscattered far-field response of a conducting cube 1.0 m on a side illuminated by a Gaussian plane wave. (a) x-component of the electric field at $(0, 0, -100)$ m. (b) enlarge late-time domain in Fig. (a).

The second model is a PEC cube, 1.0 m on a side, centered at the origin, as shown in Fig. 2. The surface is made of 76 triangular patches and 114 unknowns. In this case, we choose the time step as $c \cdot \Delta t = 0.5R_{\min}$ for the basis MOT and $c \cdot \Delta t = 2.2R_{\min}$ for the implicit method and the MOO method, where $R_{\min} = 0.1666665$ m. We use $T = 1$ lm and $c \cdot t_0 = 6.0$ lm for all methods.

Figure 1 and Fig. 2 illustrate the simulated results of the sphere and cube by four methods. We observe that the agreement between the results from four time domain methods and IDFT of the frequency domain solution is very good in early time. However, as can be seen, the results from these time domain methods display late-time oscillations. The results of averaged method extend the span of stable region than the initial MOT method, and the results of implicit method are better than the averaging method. However, these methods cannot resolve the late time exponentially increasing oscillation. There are no unconvergent phenomena in the results of MOO method. From

Fig. 1(a) and Fig. 2(a), and also the enlarged Figs. 1(b) and 2(b), it is shown that a slack decreasing oscillation exists.

We also calculated the backscattered electric field of two complex shaped scatters in order to test the applicability of these methods.

For the third example, we consider a metallic plane model. The length is 0.2 m, and we take the top position as origin, as shown in Fig. 3. The surface is made of 356 triangular patches and 534 unknowns. In this case, we choose the time step as $c \cdot \Delta t = 1.5R_{\min}$, $T = 0.06$ lm and $c \cdot t_0 = 0.4$ lm for the implicit method and the MOO method, where $R_{\min} = 0.00121094$ m.

The fourth example is a metallic missile model. The length is 0.15 m, and we take the top position as origin, as shown in Fig. 4. The surface is made of 344 triangular patches and 516 unknowns. In this case, we choose the time step as $c \cdot \Delta t = 25R_{\min}$, $T = 0.04$ lm and $c \cdot t_0 = 0.3$ lm for the implicit method and the MOO method, where $R_{\min} = 0.000499$ m.

We cannot obtain a convergent results for two complex structure models using the basic MOT method (including initial and averaging MOT). Fig. 3 and Fig. 4 represents the simulated result of the plane and missile models by using the IMPT and MOO methods. We observe that the agreement between the results from two time domain and IDFT of the frequency domain solution is very good in early time. However, as can be seen, the results from implicit method display late-time instability with exponentially increasing oscillations. The MOO method can solve the unconverged problem in late time, but there still exists a slack decreasing oscillation problem as

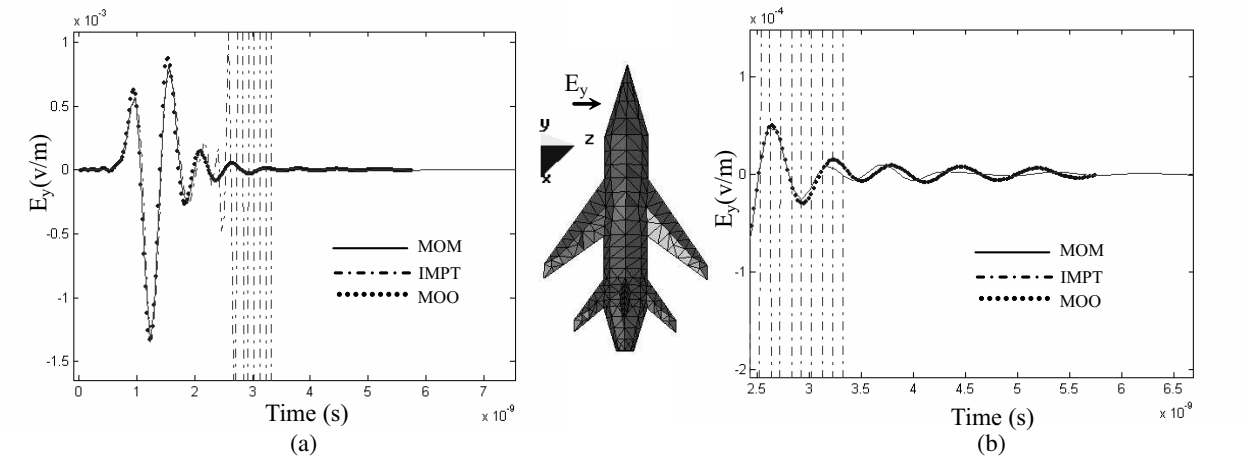


Figure 3: Backscattered far-field response of a conducting plane model 0.2 m in length illuminated by a Gaussian plane wave. (a) x-component of the electric field at $(0, 0, -100)$ m. (b) enlarge late-time domain in Fig. (a).

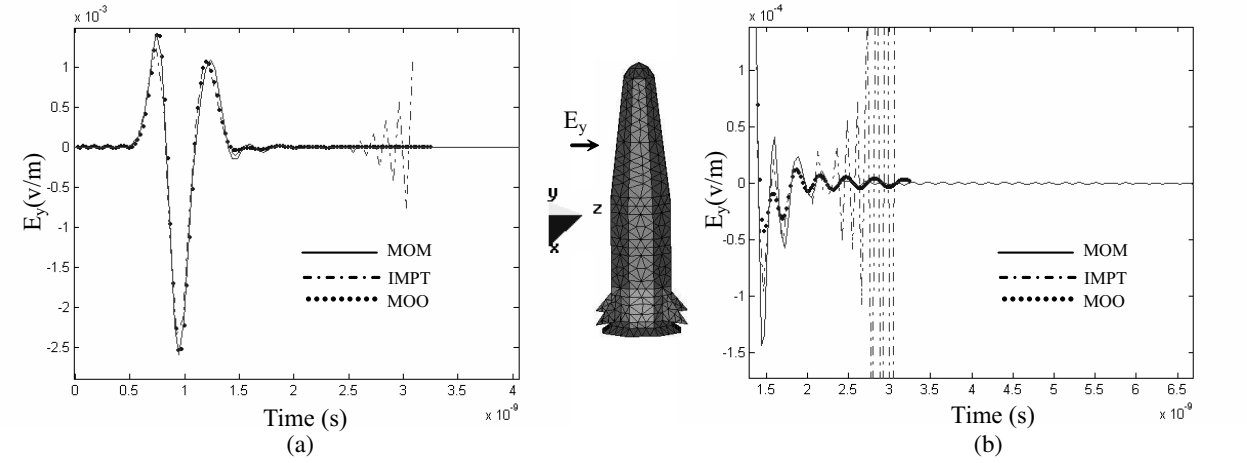


Figure 4: Backscattered far-field response of a conducting missile model 0.15 m in length illuminated by a Gaussian plane wave. (a) y-component of the electric field at $(0, 0, -100)$ m. (b) enlarge late-time domain in Fig. (a)

can be seen from Fig. 3(b) and Fig. 4(b).

4. CONCLUSION

In this paper, we calculated the backscattered electric field for four targets using four time domain methods respectively. By comparing their performances we can know that: 1) the solution from the initial MOT method easily becomes unstable with exponentially increasing oscillations and the improved averaging method extend the span of stable region than initial MOT method but still becomes unstable easily. Both methods cannot be used in complex structures. 2) The solution from the IMPT method extend the stable region more than the averaging method, but still can not solve the unconvergence phenomena emerged in late time. However, the IMPT method can be used to calculate the complex targets. 3) The MOO method solved the unconverged problem in late time, but there exists a slack decreasing oscillation problem. The slack decreasing oscillation still affects the veracity of the late-time response. We have found that combining a few frequency data will amend this phenomenon, and we are going to work on this problem.

REFERENCES

1. Rao, S. M. and D. R. Wilton, "Transient scattering by conducting surfaces of arbitrary shape," *IEEE Trans. on Antennas and Propagation*, Vol. 39, No. 1, 56–61, Jan. 1991.
2. Rao, S. M. and D. R. Wilton, "Electromagnetic scattering by surfaces of arbitrary shape," *IEEE Trans. on Antennas and Propagation*, Vol. AP-30, No. 3, 409–418, May 1982.
3. Vechinski, D. A. and S. M. Rao, "A stable procedure to calculate the transient scattering by conducting surfaces of arbitrary shape," *IEEE Trans. Antennas and Propagation*, Vol. 40, No. 6, 661–665, June 1992.
4. Rao, S. M., D. A. Vechinski, and T. K. Sarkar, "Transient scattering by conducting cylinders-implicit solution for the transverse electric case," *Microwave Opt. Technol. Letters*, Vol. 21, 129–134, 1999.
5. Jung, B. H., T. K. Sarkar, and Y. S. Chung, "An accurate and stable implicit solution for transient scattering and radiation from wire structures," *Microwave and Optical Technology Letters*, Vol. 34, No. 5, Sep. 5, 2002.
6. Jung, B. H. and T. K. Sarkar, "Time domain electric field integral equation with central finite difference," *Microwave and Optical Technology Letters*, Vol. 31, No. 6, 429–435, Dec. 2001.
7. Zhong, J. and T. K. Sarkar, "A stable solution of time domain electric field integral equation for thin-wire antennas using the laguerre polynomials," *IEEE Trans. Antennas and Propagation*, Vol. 52, No. 10, 2641–2649, Oct. 2004.
8. Chung, Y. S. and T. K. Sarkar, "Solution of time domain electric field integral equation using the laguerre polynomials," *IEEE Trans. Antennas and Propagation*, Vol. 52, No. 9, 2319–2328, Sep. 2004.
9. Jung, B. H., Y. S. Chung, and T. K. Sarkar, "Time domain EFIE, MFIE, CFIE formulations using laguerre polynomials as temporal basis functions for the analysis of transient scattering from arbitrary shaped conducting structures," *Progress in Electromagnetics Research*, Vol. PIER 39, 1–45, 2003.

A Method to Reduce the Oscillations of the Solution of Time Domain Integral Equation Using Laguerre Polynomials

Xinpu Guan, **Shaogang Wang**, **Yi Su**, **and Junjie Mao**

School of Electronic Science and Technology, National University of Defense Technology
Changsha 410073, China

Abstract— In this paper, the reason of oscillations appeared in late time low energy region of time domain electric filed integral equation using laguerre polynomials is analyzed. And then, we propose an approach to reduce the oscillations by combining frequency and time domain. According to the Fourier transform relation between frequency responses and time responses, the solution of TDIE method with Laguerre expansion is absolutely convergent. Replace the spectrum of oscillation parts of time domain data with accurate frequency data by MOM. The stable time domain data can be obtained by performing IFFT to the revised frequency data. Numerical results are presented to illustrate the efficiency of this approach.

1. INTRODUCTION

For solving the time domain integral equation, the marching on in time (MOT) method [1] is usually employed. However, as pointed by many researchers, some of the TD-EFIE formulations associated with the MOT method suffers from late-time instability, which usually takes the form of an exponentially increasing high frequency oscillation. In order to eliminate the instability of the MOT methods, the approximate averaging method [2] and the implicit scheme based on backward finite difference [3] approximation or central finite difference [4] methodology are proposed. Even though employing these methods, the solution obtained by using MOT has still exponentially increased late-time oscillations but only extend the span of the stable region. Young-seek Chung and Sarkar [5–7] proposed a marching on in order method of using the expansion by weighted Laguerre polynomials for temporal coefficients of Hertz vector. The advantages of using the weighted Laguerre polynomials are the solution is independent of the time discretization, and all the computations are carried out using spatial variables. So the exponentially increasing solution will not emerge in late time. But there still exist some oscillations in late-time low energy region due to oscillations characteristic of weighted Laguerre polynomials. Oscillation phenomenon effects on obtaining other electromagnetic characteristic from the late time scattering data. In this paper, the reason of oscillations appeared in late time low energy region of time domain electric filed integral equation using laguerre polynomials is analyzed, and then, we consider absolute convergence of the time solution and the Fourier transform relation between temporal response and frequency response. An approach based on discrete Fourier transforms to improve the stabilities in late time is proposed.

2. THE REASON OF OSCILLATIONS IN LATE-TIME

It is assumed that we have obtained the coefficients for the Hertz vector $P_{n,i}\ \ i = 1, 2 \cdots u$. Details of the method can be found in [6]. We can derive the scattered electric field as:

$$E^s(r,t) = \sum_{v=0}^{u} a_v \varphi_v(\tilde{t}) \tag{1}$$

where $\tilde{t} = st - s\dfrac{r - \hat{r}\cdot\bar{r}'}{c}$, $\varphi_v(t,s) = e^{-s\cdot t/2} L_v(s\cdot t)$, $a_v = \dfrac{\eta s^2}{4\pi r c}\sum\limits_{n=1}^{N}\left(0.25P_{n,v} + \sum\limits_{i=0}^{v-1}(v-i)P_{n,i}\right)\iint_s f_n(r)\times\hat{r}ds\times\hat{r}$, s is time scale factor, $\tilde{t}$ is the scaled delay time. $L_v(t)$ is Laguerre polynomial for order v, $\varphi_v(t)$ is weighted Laguerre polynomial. $f_n(r)$ is the RWG basis function [8], η is the wave impedance in the medium surrounding the scatter.

So the last result of the electric field can be presented by the sum of Laguerre polynomials with different order. We know that when the signals with instantaneous jump pass through band-limited system, or the signals are expanded with Fourier series of finite orders, Gibbs phenomena will emerge. In the same way, oscillations will occur in late time low energy domain while the scattered electric fields are expanded by Laguerre polynomials with finite orders. The only difference

is the frequency in truncation place. It is a frequency span in the truncation place with Laguerre expansion, unlike Fourier expansion where the frequency is a point. The weighted Laguerre polynomials are plotted in Fig. 1.

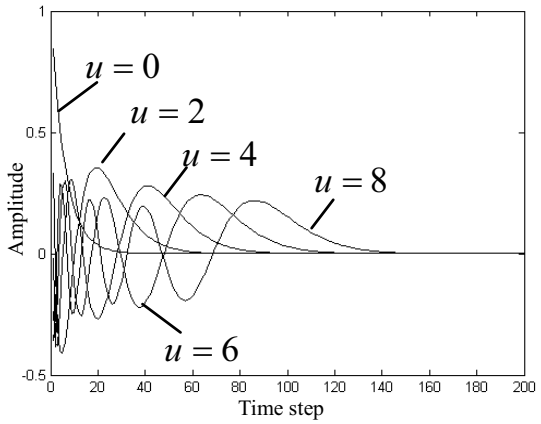


Figure 1: Weighted laguerre polynomials of different orders.

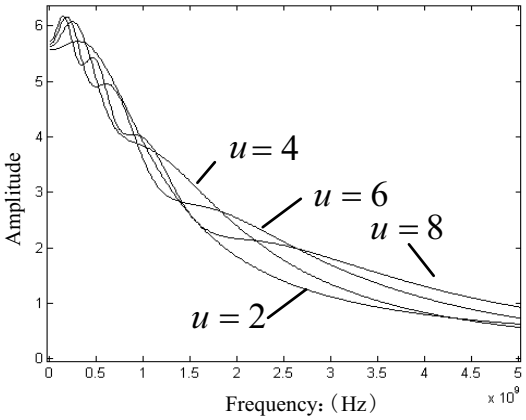


Figure 2: Frequency spectrum of the weighted laguerre polynomials of different orders.

3. IMPROVED THE INSTABILITY OF LATE TIME COMBINE FREQUENCY RESPONSE

The weighted Laguerre polynomials decay to zero as time goes to infinity and therefore the solution do not blow up for late times. So we can apply Fourier transform to the solution.

The Fourier transform of the temporal electric field is given by

$$E^s(f) = \int_{-\infty}^{\infty} E^s(r,t) \cdot e^{-j\omega t} dt = \int_{-\infty}^{\infty} \sum_{v=0}^{u} a_v \varphi_v(\tilde{t}) \cdot e^{-j\omega t} dt = \sum_{v=0}^{u} a_v \int_{-\infty}^{\infty} \varphi_v(\tilde{t}) \cdot e^{-j\omega t} dt \quad (2)$$

From formula (5), we can see that the spectrum of temporal electric field also can be spanned by spectrum of the weighted Laguerre polynomials with different orders. The spectrums of the weighted Laguerre function of different order, are plotted in Fig. 2, are oscillatory too. So there should exist an oscillation region in the frequency spectrum transformed from time data.

It is accordant for time response and frequency response through the Fourier transform, and we can obtain frequency response of arbitrary frequency by using MOM [8, 9] method. If we replace the oscillation region with right frequency data by MOM method, and then make Inverse-Fourier transform for entire spectrum, a stable time data can be obtained.

Because the incident fields forms are different in frequency domain and time domain, we need to compensate frequency data by MOM for amplitude and phase to keep consistency in the data of Fourier transform of time domain data.

The computational cost is proportional to the number of the frequency point which need replacing. If the oscillation region of spectrum is wide, the computational time will increase. In order to reduce computational cost, we only need to calculate partial frequency data from MOM, and then smooth it through interpolation/extrapolation of frequency domain data.

3.1. Amplitude and Phase Supplement

It is assumed that the scattering system of the target is a linear system. And its pulse response is $h(t)$, and frequency response is $H(f)$. In time domain, we use incident Gaussian plane wave

$$E^{in1}(r,t) = \vec{E}_0 \frac{1}{T\sqrt{\pi}} e^{-[1/T(ct-ct_0-\vec{r}\cdot\hat{k})]^2}; \quad (3)$$

where $\hat{k}$ is the unit vector in the propagation direction of the incident wave, T is pulse width, c is the velocity of light in the air, $\vec{r}$ is a position vector relative to the origin, and t_0 represents a time delay of the peak from the origin. In frequency domain, we use incident plane wave

$$E^{in2}(r,f) = \vec{E}_0 e^{-j\vec{k}\cdot\vec{r}} \quad (4)$$

where $\vec{k}$ is the propagation vector, $\vec{r}$ is a position vector relative to the origin.

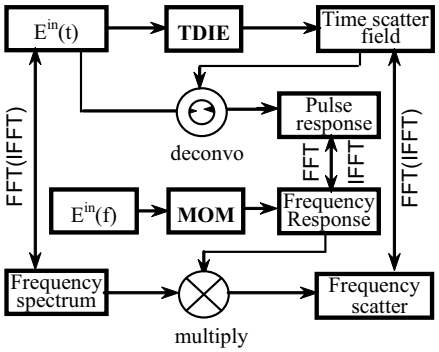


Figure 3: Relation between time domain and frequency domain.

After excited by incident field, the scattering electric fields are denoted as $E^{s1}(t)$ in time domain and $E^{s2}(f)$ in frequency domain. According to the relation between time response and frequency response, plotted in Fig. 3, we have

$$E^{s1}(t) = E^{in1}(t) \otimes h(t) \tag{5}$$

$$E^{s2}(f) = E^{in2}(f) \cdot H(f) \tag{6}$$

The Fourier transform of temporal scattering field is

$$E^{s1}(f) = E^{in1}(f) \cdot H(f) \tag{7}$$

Do Fourier transform for the incident Gaussian plane wave, by Fourier transform formula $e^{-ax^2} \rightarrow \frac{1}{\sqrt{2a}} e^{-\frac{\xi_2}{4a}}$ and vector formula $\vec{k} = \frac{\omega}{c}\hat{k}$, and assume that $x = t - t_0 - \frac{\vec{r}\cdot\hat{k}}{c}$.

The Fourier transform of the incident Gaussian plane wave can be written as

$$\begin{aligned} E^{in1}(r,f) &= \int_0^\infty E^i(r,t) \cdot e^{-j\omega t} dt \\ &= \vec{E}_0 \left(1/T\sqrt{\pi}\right) e^{-j\omega\left(t_0 + \vec{r}\cdot\hat{k}/c\right)} \int_{t_0+\vec{r}\hat{k}/c}^\infty e^{-c^2/T^2\cdot x^2} \cdot e^{-j\omega x} dx \\ &= \vec{E}_0 \left(1/\sqrt{2\pi}c\right) \cdot e^{-T^2\omega^2/4c^2} e^{-j\omega t_0} e^{-j\vec{k}\vec{r}} \end{aligned} \tag{8}$$

Inserting (4) into (6), and inserting (8) into (7), and compare them, we have

$$E^{s1}(f) = \left(1/\sqrt{2\pi}c\right) \cdot e^{-T^2\omega^2/4c^2} e^{-j\omega t_0} E^{s2}(f) \tag{9}$$

Therefore, if we make a compensation with amplitude $\left(1/\sqrt{2\pi}c\right) \cdot e^{-T^2\omega^2/4c^2}$ and phase $e^{-j\omega t_0}$ for frequency scattering electric field by MOM, then we can replace the spectrum of temporal scattering electric field with the frequency data from MOM.

3.2. Interpolation/Extrapolation of Frequency Domain Data

To maximize the use of the given information and to smooth the data in the frequency domain, we use a method based on the Hilbert transform as described in [10, 11]. The method is an iterative technique to extrapolate/interpolate frequency domain data relying on fact that the underlying time-domain data is causal, the real and imaginary parts of the frequency domain response have to be related through the Hilbert transform. The details of the method can be found in [10, 11].

Figure 4 compares the real and imaginary parts of the interpolated data with original data. Clearly, padded with zeros to the sample point 221 to 250, the reconstruction is quite accurate using this technique.

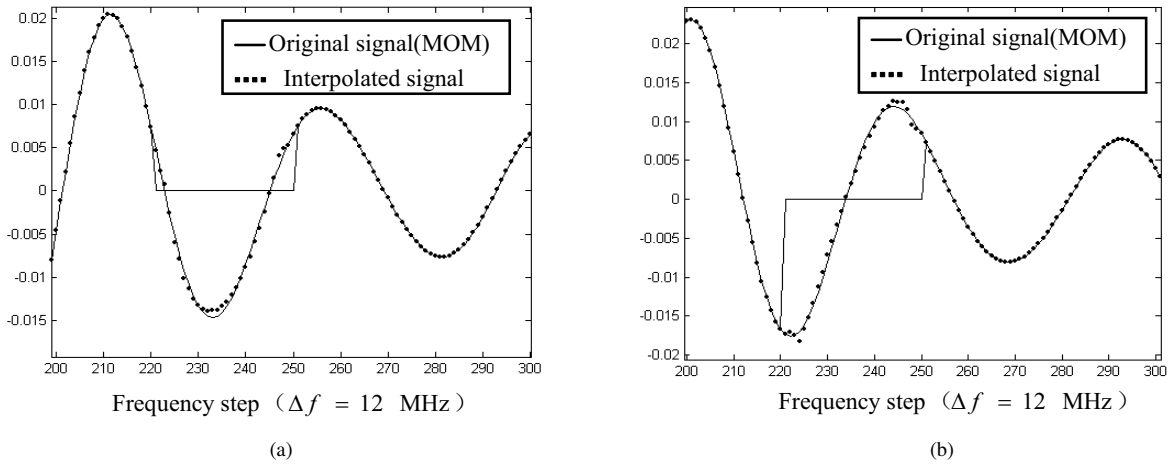


Figure 4: Comparison of the reconstruction using the Hilbert methods with original signal. (a) real part, (b) imaginary part.

3.3. The Operation Procedure

Assume that we have a sequence of time domain data with length N_t, defined to be $x(i)$, $i = 1 \cdots N_t$.

1) Perform a N_t point FFT on this time domain data and define the resulting sequence to be $E^{s1}(k)$, $k = 1 \cdots N_f$,
2) Find the oscillation samples $E^{s1}(N_1 : N_2)$, select samples point $[N_3 : N_4]$ from $[N_1 : N_2]$ ($N_1 < N_3 < N_4 < N_2$, and $N_2 - N_4 + N_3 - N_1 < N_t/5$). Calculate scattering field sequence $E^{s2}(N_3 : N_4)$ by using MOM, replace the $E^{s1}(N_3 : N_4)$ with $(1/\sqrt{2\pi} \cdot e^{-T^2\omega^2/4c^2} e^{-j\omega t_0} \cdot E^{s2}(N_3 : N_4)$. ω is the frequency in samples point $[N_3 : N_4]$.
3) Smooth the new N_f point frequency data with Hilbert transform. Define the resulting sequence to be $E^{\text{snew}}(k)$, $k = 1 \cdots N_f$.
4) Perform a N_f point IFFT on $E^{\text{snew}}(k)$.

4. NUMERICAL EXAMPLES

In this section, two perfectly conducting structures are tested to validate the proposed method, all the structures are illuminated by the incident field with $\vec{E}_0 = \hat{x}$, $\hat{k} = -\hat{z}$. All time domain data are obtained from time domain electric filed integral equation using Laguerre polynomials, and frequency domain data from MOM.

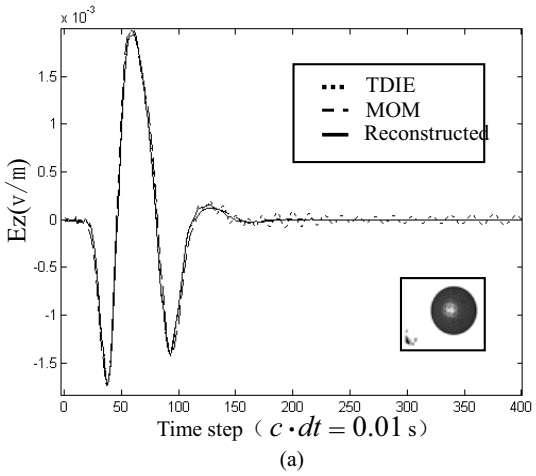
For the fist example, we consider a PEC sphere with radius 0.06 m that is centered at the origin as shown in Fig. 5. The number of triangular patches is 160 and the number of unknowns is 240. In this case we use $T = 0.1/c$ s, $t0 = 0.6/c$ s, $dt = 0.01/c$ s, and time scale factor $s = 1.0 \times 10^9$. The frequency step $\Delta f = 1.5 \times 10^7$ Hz. (The time domain data to be Fourier transformed can be formed by concatenating an appropriate number of zeros, to control the frequency step).

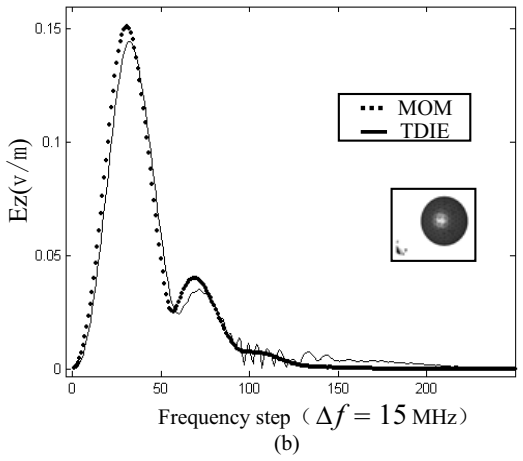
Figure 5(b) shows the frequency spectrum from MOM and the Fourier transform of the time domain data from Laguerre method. We can see that there exists oscillating region from sample 90 to 130. We calculate frequency data from sample 101 to 110, and reconstruct time data using the procedure described in Section 3.3. Figs. 5(a) and (c) shows the reconstructed result compared with the original time data and the accurate IFFT data of frequency domain data from MOM. We can see that the reconstructed data reduced the oscillation emerged in late time. It is almost equal to the IFFT data of frequency domain data from MOM.

The second target is a PEC plane model with length 0.2 m, front-wingspan is 0.1125 m, back-wingspan is 0.07 m, and up-wingspan is 0.03 m, as shown in Fig. 6. The number of triangular patches is 686 and the number of unknowns is 1029. In this case we use $T = 0.06/c$ s and $t0 = 0.6/c$ s, $dt = 0.0075/c$ s, and time scale factor $s = 8.0 \times 10^9$. The frequency step $\Delta f = 1.2 \times 10^7$ Hz.

Figure 6(b) shows the frequency spectrum from MOM and the Fourier transform of the time domain data from Laguerre method. We can see that there exists oscillation region from sample 80 to 140. Replaced frequency samples 101 to 120 by MOM, and reconstructed time data using the procedure are described in Section 3.3. From Figs. 6(a) and (c), we can see that the reconstructed

data reduce the oscillation emerged in late time. It is almost equal to the IFFT data of frequency domain data from MOM too.





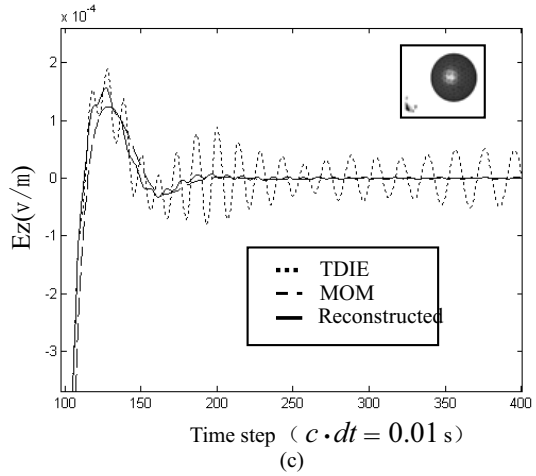
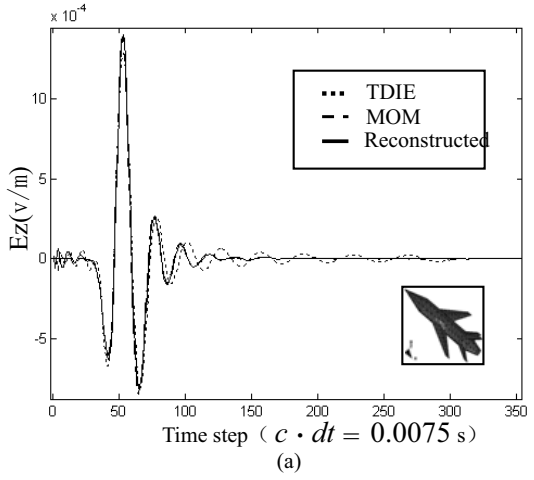
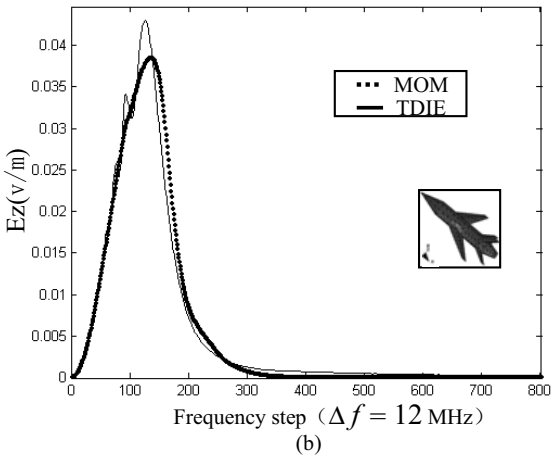


Figure 5: Simulation of PEC sphere. (a) scattering field in time domain at the point(0,0,1000 m), (b) spectrum of scattering field at the point (0,0,1000 m), (c) scattering filed in late time domain.





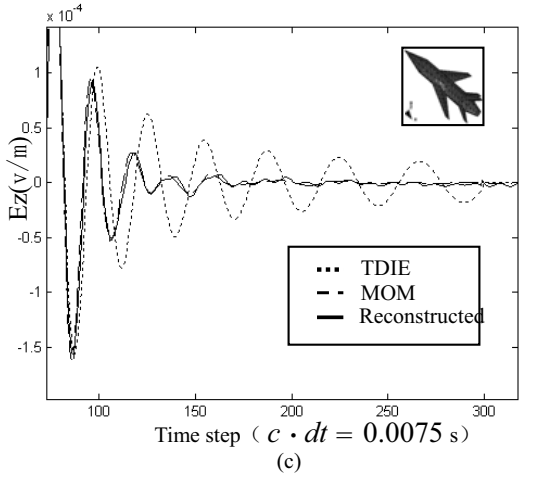


Figure 6: Simulation of PEC plane model. (a) scattering field in time domain at the point(0,0,1000 m), (b) spectrum of scattering field at the point (0,0,1000 m), (c) scattering filed in late time domain.

5. CONCLUSIONS

In this paper, we have reduced the oscillation in late time domain by using TDIE with Laguerre polynomials through combining a few partial frequency domain data with entire time domain data. Two examples have demonstrated the good reconstructed result by using a few frequency data to replace the dilapidate spectrum of time domain data. It is important to obtain other electromagnetic characteristic from the late time scattering data.

REFERENCES

1. Rao, S. M. and D. R. Wilton, "Transient scattering by conducting surfaces of arbitrary shape," *IEEE Trans on Antennas and Propagation*, Vol. 39, No. 1, 56–61, Jan 1991.
2. Vechinski, D. A. and S. M. Rao "A stable procedure to calculate the transient scattering by conducting surfaces of arbitrary shape," *IEEE Trans. Antennas and Propagation*, Vol. 40, No. 6, 661–665, June 1992.
3. Rao, S. M. and T. K. Sarkar, "An efficient method to evaluate the time domain scattering from arbitrarily shaped conducting bodies," *Microwave And Optical Technology Letters*, Vol. 17, No. 5, 321–325, April 1998.
4. Jung, B. H. and T. K. Sarkar, "Time domain electric field integral equation with central finite difference," *Microwave and Optical Technology Letters*, Vol. 31, No. 6, 429–435, Dec. 2001.
5. Ji, Z. and T. K. Sarkar, "A stable solution of time domain electric field integral equation for thin-wire antennas using the laguerre polynomials," *IEEE Trans. Antennas and Propagation*, Vol. 52, No. 10, 2641–2649, Oct. 2004.
6. Chung, Y. S. and T. K. Sarkar, "Solution of time domain electric field integral equation using the laguerre polynomials," *IEEE Trans. Antennas and Propagation*, Vol. 52, No. 9, 2319–2328, Sep. 2004.
7. Jung, B. H. and T. K. Sarkar, "Transient electromagnetic scattering from dielectric objects using the electric field integral equation with laguerre polynomials as temporal basis functions," *IEEE Trans. Antennas and Propagation*, Vol. 52, No. 9, 2329–2340, Sep. 2004.
8. Rao, S. M. and D. R. Wilton, "Electromagnetic scattering by surfaces of arbitrary shape," *IEEE Trans. Antennas and Propagation*, Vol. AP-30, No. 3, 409–418, May 1982.
9. Cui, T. J. and W. C. Chew, "Efficient MLFMA, RPFMA and FAFFA algorithms for EM scattering by very large structures," *IEEE Trans. Antennas and Propagation*, Vol. 52, No. 3, 759–770, March 2004.
10. Narayana, S., G. Rao, and R. Adve, "Interpolation /extrapolation of frequency domain responses using the Hilbert transform," *IEEE Trans. Microwaves Theory and Techniques*, Vol. 44, No. 10, 1621–1627, Oct. 1996.
11. Narayana, S., T. K. Sarkar, and R. Adve, "A comparison of two techniques for the interpolation/extrapolation of frequency domain responses," *Digital Signal Processing*, Vol. 6, 51–67, 1996.

An Integer-N Frequency Synthesizer Applied to MB-OFDM UWB 5th Band Group

Cheng-Chan Tien[1], **Tsung-Mo Tien**[2], **and Christina F. Jou**[1]
[1]Chiao-Tung University, Taiwan, China
[2]Cheng-Kung University, Taiwan, China

Abstract— In this paper, we will discuss an integer-N frequency synthesizer operated at 10 GHz around the MB-OFDM UWB system 5th band group with only 1.5 V supply voltage which owns a varying frequency ranging from 9.29 GHz to 10.9 GHz and delivers −5.9 dBm output power. A single NMOS cross-coupled-pairs forms the core circuit of the voltage-controlled oscillator and delivers a −97 dBc/Hz phase noise at 1 MHz from the carrier signal. The frequency settling time is about 25 µs and the whole circuit power consumption is 23.55 mW.

1. INTRODUCTION

In a MB-OFDM UWB system, the extremely wide frequency band is divided into five band groups while the 5th group divided into two frequency bands [1]. Since frequency synthesizers designed at such a high frequency 10 GHz are not popular yet, and a synthesizer that can operates from 3.1 GHz ∼ 10.6 GHz needs not only more control signals than usual but also raising the circuit implementation complexity, we will try to prove that the highest band is possibly realizable by using the conventional PLL architecture.

In this paper, a synthesizer shown in Fig. 1 with a varying frequency ranging from 9.5 GHz ∼ 10.6 GHz is designed, and a fully programmable multi-modulus frequency divider (FPMMFD) [2] scales down the oscillating frequency to meet the target we set.

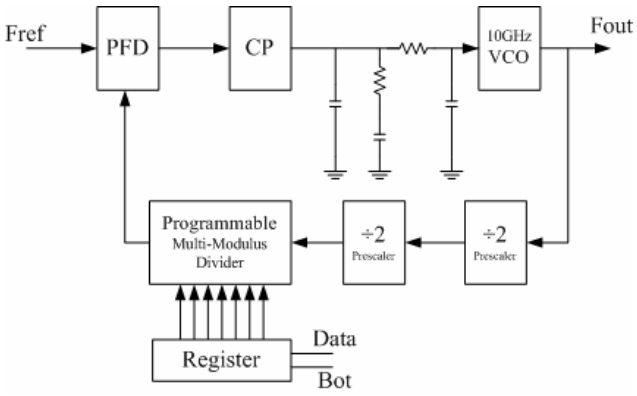


Figure 1: 10 GHz frequency synthesizer with multi-modulus divider.

2. 10 GHZ VCO

Instead of using complementary cross-coupled pairs as usual, we adopt a single NMOS couple pair for the voltage controlled oscillator as shown in Fig. 2. In this work and consequently outputs, two differential signals which can effectively alleviate the common-mode noise coupled from the substrate. A NMOS coupled circuit, which is free from the upper operating region limits, that bothers a lot under the cascading condition. As long as well adjusting the size of NMOS can we obtain a greater output waveform.

We get rid of the traditional single-in single-out inductor often used in NMOS-only voltage controlled oscillator, and choose a center-tapped structure as shown in Fig. 3 [3] in behalf of high Q value and improvement in economizing the circuit area.

3. VCO FULLY PROGRAMMABLE MULTI-MODULUS FREQUENCY DIVIDER

The most common choices of frequency dividers in frequency synthesizer are phase-switching circuit and programmable pulse-swallow counter; however these two architectures have lower flexibility.

So in this design we adopt the fully programmable multi-modulus frequency divider (FPMMFD) which is not only easy to be implemented, but can also effectively reduces the possibility of dividing error since that the delay time of every stage only related to next stage compared with other divider architecture.

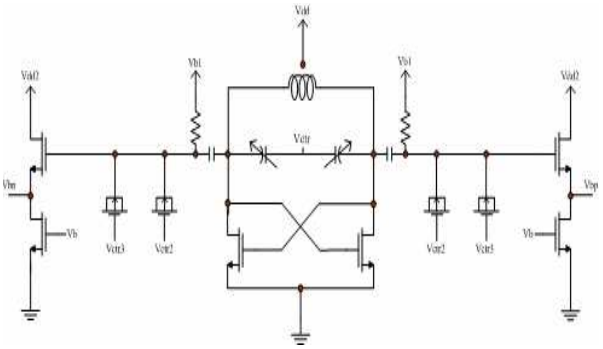


Figure 2: NMOS cross-coupled VCO.

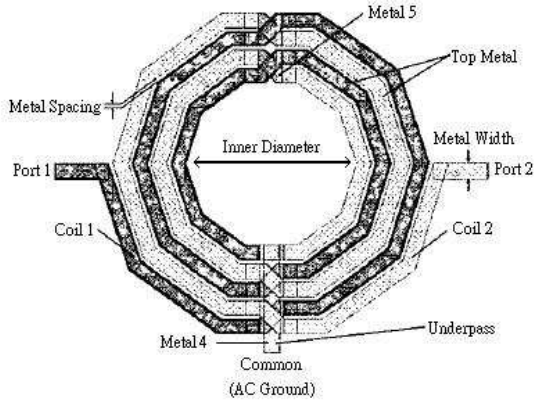


Figure 3: A differential-in center-tapped inductor.

Figure 4 shows a fully programmable multi-modulus divider that consists of 7 asynchronously cascaded dual modulus divide-by-2/3 dividers [4, 5] is put to use in order to achieve both high-speed frequency division and moderate power consumption. The first two stages assumed to operate at high frequency. Therefore, both two circuits are realized in a differential source coupled logic (SCL) and the others are accomplished as digital devices.

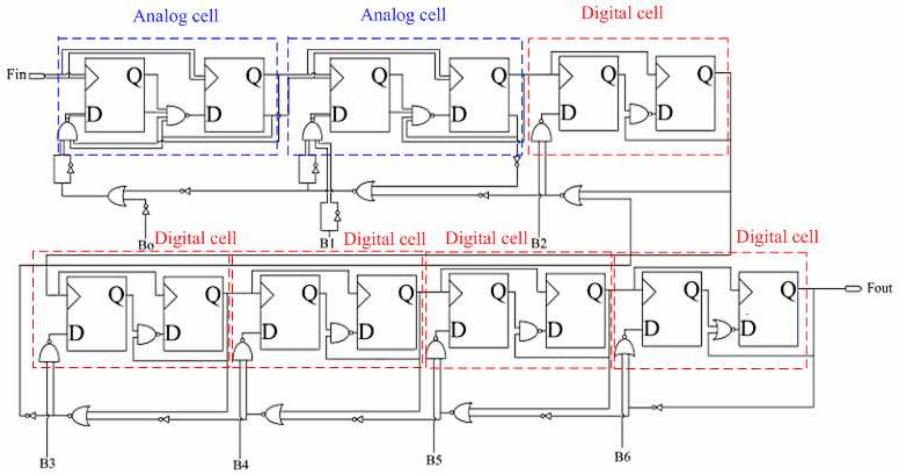


Figure 4: Fully programmable multi-modulus frequency divider.

We can vary the total division N by changing the input level of each block's control bit (B_0, B_1, $B_2 \ldots$) that are brought out from the register which will be introduced in next section, and the programmable dividing ratio is:

$$N = 2^7 + \sum_{n=0}^{6} b_n \cdot 2^n = 128 + \sum_{n=0}^{6} b_n \cdot 2^n$$

This divider structure provides high flexibility and the simple logic of the AND/OR gates assures that the modulus signals of the last stages are produced first and given to the next stage. Thus the delay time in the critical path, the feedback of the first stage, is minimized.

4. REGISTER

Owing to the great number of input signals a frequency synthesizer has, if we design a pad for every single input control signal and output signal, the chip size will be enlarged. A register composed of

seven cascaded D-type flip flops is shown in Fig. 5 due to seven control signals a single FPMMFD needs. The control signals are input from the node named Data, and by the clock ticking, the frequency divider can be accurately loaded. The register not only prevents long metal lines in layout which will lead to serious parasitic effects, it also scales the chip size down by reducing the pad numbers from seven to two.

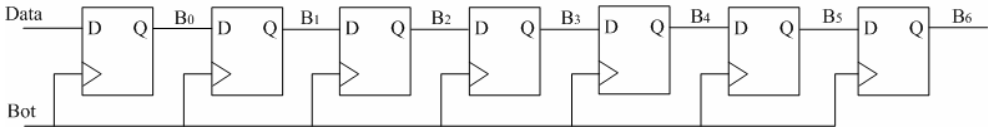


Figure 5: A seven stages cascaded register.

5. SIMULATION RESULTS

Figure 6 is the tuning range of the VCO which owns the varying range between 9.29 GHz and 10.9 GHz. Fig. 7 is the output signal through the buffer of a 10 GHz VCO and can be clearly distinguished that the output peak-to-peak voltage is about 1000 mV, and Fig. 8 also shows the output power spectrum, and the simulated result is −5.79 dBm at 10 GHz. Phase noise is −97 dBc/Hz at 1 MHz offset from the carrier signal.

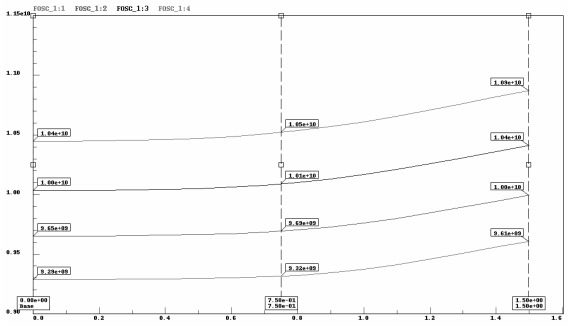


Figure 6: VCO tuning range varies from 9.29 GHz to 10.9 GHz.

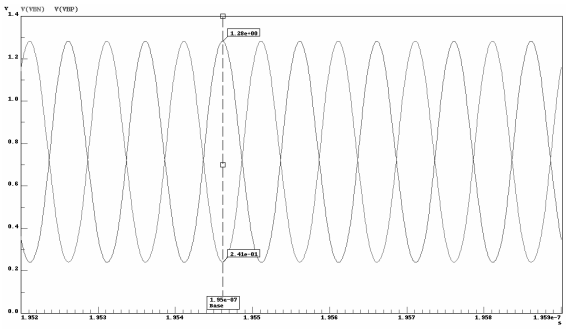


Figure 7: Output voltage of a 10 GHz VCO.

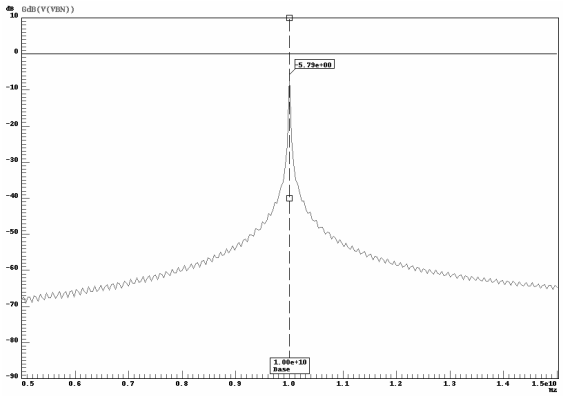


Figure 8: Output power is −5.79 dBm at 10 GHz.

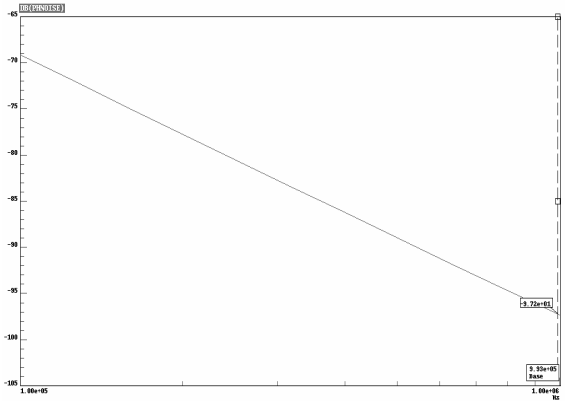


Figure 9: Phase noise.

The settling voltage and time simulation results are shown in Fig. 10 and Fig. 11 separately. In the figures we can find out that the settling voltage falls at 1 V at about 25 us and the oscillating frequency will maintain at 10 GHz.

Comparison between this work and the references are list in Tab. 1 and the corresponding chip layout of this study is shown in Fig. 12.

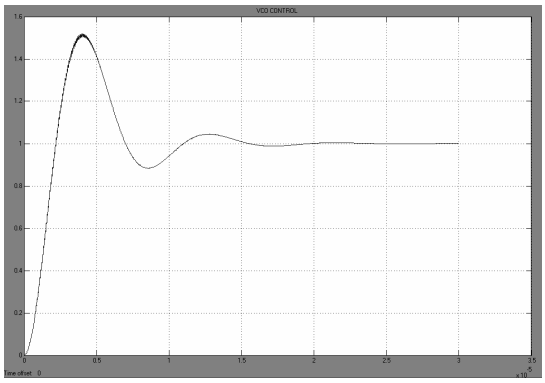


Figure 10: Settling voltage at about 1 V.

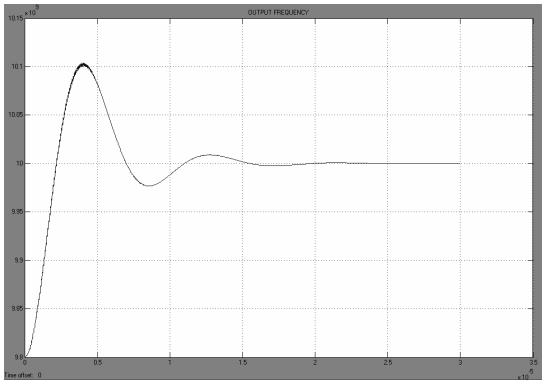


Figure 11: Locking frequency vs. settling time.

Specification	[6]2006	This work
Fabrication	0.13 m CMOS	0.18 m CMOS
Supply Voltage	1.5V	1.5V
Center Frequency	7.92GHz	10GHz
Tuning Range	7.84~9.41GHz	9.29~10.9GHz
Tuning Varactors	3	3
Settling Time	1.5ns	25us
Phase Noise @1MHz	-115dBc/Hz	-97.8dBc/Hz
Ref. Frequency	33MHz	16MHz
Die Area	1.3 X 0.7mm^2	0.998 X 0.7mm^2
Power Dissipation	62mW	23.55mW

Table 1: Comparison between the references and the simulation results.

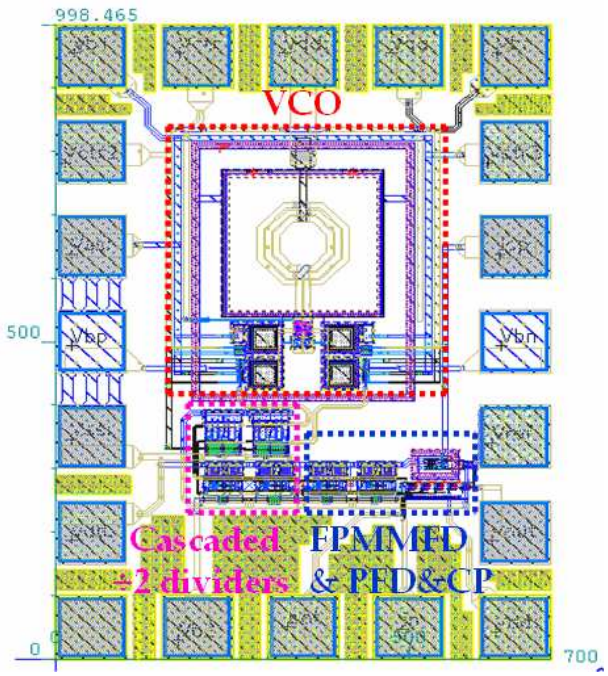


Figure 12: Full chip layout of the 10 GHz Synthesizer.

6. CONCLUSIONS

In this paper, an integer-N frequency synthesizer operated at 10 GHz around the MB-OFDM UWB system 5th band group with only 1.5 V supply voltage which owns a varying frequency ranging from 9.29 GHz to 10.9 GHz and delivers −5.9 dBm output power is shown. With a −97 dBc/Hz phase noise and 25 us settling time, we improve the traditional architecture can operates at high frequency such as 10 GHz under well EM simulations.

REFERENCES

1. Green, E. R. and S. Roy, "System architectures for high-rate ultra-wideband communication systems: a review of recent developments," Intel Labs 2111 NE 25th Ave. Hillsboro, OR97124.
2. Theill, D., C. Durdodt, A. Hanke, S. Heinen, S. van Waasen, D. Seippel, D. Pham-Stabner, and K. Schumacher, "A fully integrated cmos frequency synthesizer for bluetooth," *IEEE Radio Frequency Integrated Circuits (RFIC) Symposium*, 103–106, Phoenix, Arizona, USA, 20–22 May 2001.
3. Sia, C. B., K. W. Chan, C. Q. Geng, W. Yang, K. S. Yeo, M. A. Do, J. G. Ma, S. Chu, and K. W. Chew, "An accurate and scalable differential inductor design kit," 63–68, 22–25 March 2004.

4. Lam, C. and B. Razavi, "A 2.6-GHz/5.2-GHz frequency synthesizer in 0.4-µm CMOS technology," *IEEE Journal of Solid-State Circuits*, Vol. 35, Issue 5, 788–794, May 2000, Digital Object Identifier 10.1109/4.841508.
5. Foroudi, N. and T. A. Kwasniewski, "CMOS high-speed dual-modulus frequency divider for rf frequency synthesis," *IEEE Journal of Solid-State Circuits*, Vol. 30, Issue 2, 93–100, Feb. 1995, Digital Object Identifier 10.1109/4.341735.
6. Lee, J.-E., E.-C. Park, C.-Y. Cha, H.-S. Chae, C.-D. Suh, J. Koh, H. Lee, and H.-T. Kim, "A frequency synthesizer for UWB transceiver in 0.13/spl mu/m CMOS technology," *Silicon Monolithic Integrated Circuits in RF Systems*, Digest of Papers, *Topical Meeting on 18–20 Jan. 2006*, 4, 2006, Digital Object Identifier 10.1109/SMIC.2005.1587974.

Single-end Measurement of Polarization Mode Dispersion in Optical Fibers with Polarization-dependent Loss

H. Dong and P. Shum
Network Technology Research Centre, Nanyang Technological University, Singapore

Abstract— A new reflectometric technique is proposed to measure the polarization mode dispersion (PMD) in optical fiber links with polarization-dependent loss. Based on the PMD dynamical equation, some important relations can be obtained to make the backreflection measurement of PMD possible. Thus we can perform the single-end spectral resolved measurement of the differential group delay and the differential attenuation slope in both frequency and time domains simultaneously. In principle, this technique can be realized using continue-wave or pulsed probe lights and based on far-end Fresnel reflection or distributed Rayleigh backscattering. Some experimental results confirm the validity of the proposed method.

Most of the polarization mode dispersion (PMD) measurement techniques require both fiber ends to implement the tests, consequently, field tests on installed cables are time consuming and not simple [1]. Thus the backscattering technique has been proposed to perform the single-end PMD measurement [1–3]. In an optical fiber link with polarization-dependent loss (PDL), due to the round-trip effect, the round-trip complex PMD vector $\vec{W}_{\mathrm{B}} = \vec{\Omega}_{\mathrm{B}} + i\vec{\Lambda}_{\mathrm{B}}$, which can be measured directly in the test, is only explicitly related with the first two elements of the complex PMD vector $\vec{W} = \vec{\Omega} + i\vec{\Lambda}$ as [4]

$$\Omega_{\mathrm{B}}^2 - \Lambda_{\mathrm{B}}^2 = 4\left(\Omega_{\mathrm{L}}^2 - \Lambda_{\mathrm{L}}^2\right), \quad \vec{\Omega}_{\mathrm{B}} \cdot \vec{\Lambda}_{\mathrm{B}} = 4\vec{\Omega}_{\mathrm{L}} \cdot \vec{\Lambda}_{\mathrm{L}}. \tag{1}$$

where $\vec{\Omega}_{\mathrm{L}} = (\Omega_1,\, \Omega_2,\, 0)^{\mathrm{T}}$ and $\vec{\Lambda}_{\mathrm{L}} = (\Lambda_1,\, \Lambda_2,\, 0)^{\mathrm{T}}$ are the linear parts of the complex PMD vector, T denotes the matrix transpose. Moreover, due to the lack of equations, Eq. (1) can not be solved. As a consequence, the single-end PMD measurement in such optical fibers can not be realized up to now. In this paper, we propose a novel technique, which can perform the single-end spectral resolved PMD measurement in optical fibers with PDL. This technique is based on the PMD dynamical equation. Two equations expressing the third elements of the complex PMD vector, Ω_3 and Λ_3, can be obtained by the backreflection measurements carried simultaneously in both frequency and time domains. Thus, combined with Eq. (1), the differential group delay (DGD) and the differential attenuation slope (DAS) can be deterministically achieved. In principle, this technique is applicable for both far-end Fresnel reflection and distributed Rayleigh backscattering; on the other hand, it can be performed using continuous-wave or optical pulse.

For a fiber system with both birefringence and PDL, its Mueller matrix $\mathbf{M}$ meets the Lorentz transformation [5]. Thus, it has been demonstrated theoretically and experimentally that if we normalize $\mathbf{M}$ to make $\det \mathbf{M} = 1$, then [6]

$$\mathbf{B} = \frac{\mathrm{d}\mathbf{M}}{\mathrm{d}z}\mathbf{M}^{-1} = \begin{bmatrix} 0 & \alpha_1 & \alpha_2 & \alpha_3 \\ \alpha_1 & 0 & -\beta_3 & \beta_2 \\ \alpha_2 & \beta_3 & 0 & -\beta_1 \\ \alpha_3 & -\beta_2 & \beta_1 & 0 \end{bmatrix}, \quad \mathbf{P} = \frac{\mathrm{d}\mathbf{M}}{\mathrm{d}\omega}\mathbf{M}^{-1} = \begin{bmatrix} 0 & \Lambda_1 & \Lambda_2 & \Lambda_3 \\ \Lambda_1 & 0 & -\Omega_3 & \Omega_2 \\ \Lambda_2 & \Omega_3 & 0 & -\Omega_1 \\ \Lambda_3 & -\Omega_2 & \Omega_1 & 0 \end{bmatrix}. \tag{2}$$

where $\vec{\alpha} = (\alpha_1\, \alpha_2\, \alpha_3)^{\mathrm{T}}$ and $\vec{\beta} = (\beta_1\, \beta_2\, \beta_3)^{\mathrm{T}}$are the local PDL vector and local birefringence vector, respectively; $\vec{\Omega} = (\Omega_1\, \Omega_2\, \Omega_3)^{\mathrm{T}}$ and $\vec{\Lambda} = (\Lambda_1\, \Lambda_2\, \Lambda_3)^{\mathrm{T}}$ are the real and imaginary parts of the complex PMD vector. The round-trip Mueller matrix should be $\mathbf{M}_{\mathrm{B}} = \mathbf{R}\mathbf{M}^{\mathrm{T}}\mathbf{R}\mathbf{M}$ with $\mathbf{R} = \mathrm{diag}(1\,1\,1\,-1)$ [4] and it is also a Lorentz transformation. Then we have

$$\mathbf{B}_{\mathrm{B}} = \frac{\mathrm{d}\mathbf{M}_{\mathrm{B}}}{\mathrm{d}z}\mathbf{M}_{\mathrm{B}}^{-1} = \begin{bmatrix} 0 & \alpha_{\mathrm{B}1} & \alpha_{\mathrm{B}2} & \alpha_{\mathrm{B}3} \\ \alpha_{\mathrm{B}1} & 0 & -\beta_{\mathrm{B}3} & \beta_{\mathrm{B}2} \\ \alpha_{\mathrm{B}2} & \beta_{\mathrm{B}3} & 0 & -\beta_{\mathrm{B}1} \\ \alpha_{\mathrm{B}3} & -\beta_{\mathrm{B}2} & \beta_{\mathrm{B}1} & 0 \end{bmatrix}, \quad \mathbf{P}_{\mathrm{B}} = \frac{\mathrm{d}\mathbf{M}_{\mathrm{B}}}{\mathrm{d}\omega}\mathbf{M}_{\mathrm{B}}^{-1} = \begin{bmatrix} 0 & \Lambda_{\mathrm{B}1} & \Lambda_{\mathrm{B}2} & \Lambda_{\mathrm{B}3} \\ \Lambda_{\mathrm{B}1} & 0 & -\Omega_{\mathrm{B}3} & \Omega_{\mathrm{B}2} \\ \Lambda_{\mathrm{B}2} & \Omega_{\mathrm{B}3} & 0 & -\Omega_{\mathrm{B}1} \\ \Lambda_{\mathrm{B}3} & -\Omega_{\mathrm{B}2} & \Omega_{\mathrm{B}1} & 0 \end{bmatrix}. \tag{3}$$

where $\vec{\alpha}_{\mathrm{B}} = (\alpha_{\mathrm{B}1}\, \alpha_{\mathrm{B}2}\, \alpha_{\mathrm{B}3})^{\mathrm{T}}$ and $\vec{\beta}_{\mathrm{B}} = (\beta_{\mathrm{B}1}\, \beta_{\mathrm{B}2}\, \beta_{\mathrm{B}3})^{\mathrm{T}}$ are the local round-trip PDL vector and birefringence vector, respectively; $\vec{\Omega}_{\mathrm{B}} = (\Omega_{\mathrm{B}1}\, \Omega_{\mathrm{B}2}\, \Omega_{\mathrm{B}3})^{\mathrm{T}}$ and $\vec{\Lambda}_{\mathrm{B}} = (\Lambda_{\mathrm{B}1}\, \Lambda_{\mathrm{B}2}\, \Lambda_{\mathrm{B}3})^{\mathrm{T}}$ are the real and

imaginary parts of the round-trip complex PMD vector. Considering

$$\begin{cases} \dfrac{\partial \mathbf{P}_{\mathrm{B}}}{\partial z} = 2\mathbf{R}\mathbf{M}^{\mathrm{T}}\mathbf{R}\left(\mathbf{R}\mathbf{B}^{\mathrm{T}}\mathbf{R}\mathbf{P}_{\mathrm{L}} - \mathbf{P}_{\mathrm{L}}\mathbf{R}\mathbf{B}^{\mathrm{T}}\mathbf{R} + \dfrac{\partial \mathbf{P}_{\mathrm{L}}}{\partial z}\right)\left(\mathbf{R}\mathbf{M}^{\mathrm{T}}\mathbf{R}\right)^{-1} \\ \dfrac{\partial \mathbf{B}_{\mathrm{B}}}{\partial \omega} = 2\mathbf{R}\mathbf{M}^{\mathrm{T}}\mathbf{R}\left(\mathbf{R}\mathbf{P}^{\mathrm{T}}\mathbf{R}\mathbf{B}_{\mathrm{L}} - \mathbf{B}_{\mathrm{L}}\mathbf{R}\mathbf{P}^{\mathrm{T}}\mathbf{R} + \dfrac{\partial \mathbf{B}_{\mathrm{L}}}{\partial \omega}\right)\left(\mathbf{R}\mathbf{M}^{\mathrm{T}}\mathbf{R}\right)^{-1} \end{cases}. \tag{4}$$

Taking the sum of two equations in Eq. (4), we obtain

$$\frac{\partial \mathbf{P}_{\mathrm{B}}}{\partial z} + \frac{\partial \mathbf{B}_{\mathrm{B}}}{\partial \omega} = 2\mathbf{R}\mathbf{M}^{\mathrm{T}}\mathbf{R}\begin{bmatrix} 0 & C_1 & C_2 & 0 \\ C_1 & 0 & 0 & D_2 \\ C_2 & 0 & 0 & -D_1 \\ 0 & -D_2 & D_1 & 0 \end{bmatrix}\left(\mathbf{R}\mathbf{M}^{\mathrm{T}}\mathbf{R}\right)^{-1}. \tag{5}$$

where

$$\begin{aligned} C_1 &= \beta_3\Lambda_2 + \alpha_3\Omega_2 + \alpha_2\Omega_3 + \beta_2\Lambda_3 + \frac{\partial \alpha_1}{\partial \omega} + \frac{\partial \Lambda_1}{\partial z}, \\ C_2 &= -\beta_3\Lambda_1 - \alpha_3\Omega_1 - \alpha_1\Omega_3 - \beta_1\Lambda_3 + \frac{\partial \alpha_2}{\partial \omega} + \frac{\partial \Lambda_2}{\partial z}, \\ D_2 &= -\beta_3\Omega_1 + \alpha_3\Lambda_1 + \alpha_1\Lambda_3 - \beta_1\Omega_3 + \frac{\partial \beta_2}{\partial \omega} + \frac{\partial \Omega_2}{\partial z}, \\ D_1 &= \beta_3\Omega_2 - \alpha_3\Lambda_2 - \alpha_2\Lambda_3 + \beta_2\Omega_3 + \frac{\partial \beta_1}{\partial \omega} + \frac{\partial \Omega_1}{\partial z}. \end{aligned}$$

We can easily obtain the PMD dynamical equation as

$$\frac{\partial \vec{\Omega}}{\partial z} = \frac{\partial \vec{\beta}}{\partial \omega} + \vec{\beta} \times \vec{\Omega} - \vec{\alpha} \times \vec{\Lambda}, \quad \frac{\partial \vec{\Lambda}}{\partial z} = \frac{\partial \vec{\alpha}}{\partial \omega} + \vec{\beta} \times \vec{\Lambda} + \vec{\alpha} \times \vec{\Omega}. \tag{6}$$

In optical fibers and components used in optical communications, at least over the wavelength range of interest, we have [7]

$$\frac{\partial \vec{\alpha}}{\partial \omega} = 0 \quad \text{and} \quad \frac{\partial \vec{\beta}}{\partial \omega} = \frac{\vec{\beta}}{\omega}. \tag{7}$$

Using Eqs. (6) and (7), we have

$$\begin{aligned} C_1 &= 2(\alpha_2\Omega_3 + \beta_2\Lambda_3), & C_2 &= -2(\alpha_1\Omega_3 + \beta_1\Lambda_3), \\ D_1 &= 2\left(\frac{\beta_1}{\omega} + \beta_2\Omega_3 - \alpha_2\Lambda_3\right), & D_2 &= 2\left(\frac{\beta_2}{\omega} + \alpha_1\Lambda_3 - \beta_1\Omega_3\right). \end{aligned} \tag{8}$$

Because the right and left sides of Eq. (5) are similar matrices, they should have the same eigenvalues. Then

$$\begin{cases} 16\left(\vec{\alpha}_{\mathrm{L}}\cdot\vec{\beta}_{\mathrm{L}}\right)\left(\Omega_3^2-\Lambda_3^2\right)+16\left(\beta_{\mathrm{L}}^2-\alpha_{\mathrm{L}}^2\right)\Omega_3\Lambda_3+16\dfrac{\left|\vec{\alpha}_{\mathrm{L}}\times\vec{\beta}_{\mathrm{L}}\right|}{\omega}\Omega_3 = \left(\dfrac{\partial \vec{\Omega}_{\mathrm{B}}}{\partial z}+\dfrac{\partial \vec{\beta}_{\mathrm{B}}}{\partial \omega}\right)\cdot\left(\dfrac{\partial \vec{\Lambda}_{\mathrm{B}}}{\partial z}+\dfrac{\partial \vec{\alpha}_{\mathrm{B}}}{\partial \omega}\right) \\ 16\left(\beta_{\mathrm{L}}^2-\alpha_{\mathrm{L}}^2\right)\left(\Omega_3^2-\Lambda_3^2\right)-64\left(\vec{\alpha}_{\mathrm{L}}\cdot\vec{\beta}_{\mathrm{L}}\right)\Omega_3\Lambda_3+16\dfrac{\beta_{\mathrm{L}}^2}{\omega^2}+32\dfrac{\left|\vec{\alpha}_{\mathrm{L}}\times\vec{\beta}_{\mathrm{L}}\right|}{\omega}\Lambda_3=\left(\dfrac{\partial \vec{\Omega}_{\mathrm{B}}}{\partial z}+\dfrac{\partial \vec{\beta}_{\mathrm{B}}}{\partial \omega}\right)^2-\left(\dfrac{\partial \vec{\Lambda}_{\mathrm{B}}}{\partial z}+\dfrac{\partial \vec{\alpha}_{\mathrm{B}}}{\partial \omega}\right)^2 \end{cases}. \tag{9}$$

Due to $\beta_{\mathrm{B}}^2 - \alpha_{\mathrm{B}}^2 = 4(\beta_{\mathrm{L}}^2 - \alpha_{\mathrm{L}}^2)$ and $\vec{\beta}_{\mathrm{B}} \cdot \vec{\alpha}_{\mathrm{B}} = 4\vec{\beta}_{\mathrm{L}} \cdot \vec{\alpha}_{\mathrm{L}}$ [4], Eq. (9) becomes

$$\begin{cases} 4\left(\vec{\alpha}_{\mathrm{B}}\cdot\vec{\beta}_{\mathrm{B}}\right)\left(\Omega_3^2-\Lambda_3^2\right)+4\left(\beta_{\mathrm{B}}^2-\alpha_{\mathrm{B}}^2\right)\Omega_3\Lambda_3+16\dfrac{\left|\vec{\alpha}_{\mathrm{L}}\times\vec{\beta}_{\mathrm{L}}\right|}{\omega}\Omega_3 = \left(\dfrac{\partial \vec{\Omega}_{\mathrm{B}}}{\partial z}+\dfrac{\partial \vec{\beta}_{\mathrm{B}}}{\partial \omega}\right)\cdot\left(\dfrac{\partial \vec{\Lambda}_{\mathrm{B}}}{\partial z}+\dfrac{\partial \vec{\alpha}_{\mathrm{B}}}{\partial \omega}\right) \\ 4\left(\beta_{\mathrm{B}}^2-\alpha_{\mathrm{B}}^2\right)\left(\Omega_3^2-\Lambda_3^2\right)-16\left(\vec{\alpha}_{\mathrm{B}}\cdot\vec{\beta}_{\mathrm{B}}\right)\Omega_3\Lambda_3+16\dfrac{\beta_{\mathrm{L}}^2}{\omega^2}+32\dfrac{\left|\vec{\alpha}_{\mathrm{L}}\times\vec{\beta}_{\mathrm{L}}\right|}{\omega}\Lambda_3=\left(\dfrac{\partial \vec{\Omega}_{\mathrm{B}}}{\partial z}+\dfrac{\partial \vec{\beta}_{\mathrm{B}}}{\partial \omega}\right)^2-\left(\dfrac{\partial \vec{\Lambda}_{\mathrm{B}}}{\partial z}+\dfrac{\partial \vec{\alpha}_{\mathrm{B}}}{\partial \omega}\right)^2 \end{cases}. \tag{10}$$

$|\vec{\alpha}_{\rm L}\times\vec{\beta}_{\rm L}|/\omega$ and $\beta_{\rm L}^2/\omega^2$ can be neglected as long as Ω_3 and Λ_3 is far larger than 0.8 fs at 1550 nm. Finally

$$\begin{cases} 4\left(\vec{\alpha}_{\rm B}\cdot\vec{\beta}_{\rm B}\right)\left(\Omega_3^2-\Lambda_3^2\right)+4\left(\beta_{\rm B}^2-\alpha_{\rm B}^2\right)\Omega_3\Lambda_3=\left(\dfrac{\partial\vec{\Omega}_{\rm B}}{\partial z}+\dfrac{\partial\vec{\beta}_{\rm B}}{\partial\omega}\right)\cdot\left(\dfrac{\partial\vec{\Lambda}_{\rm B}}{\partial z}+\dfrac{\partial\vec{\alpha}_{\rm B}}{\partial\omega}\right) \\ 4\left(\beta_{\rm B}^2-\alpha_{\rm B}^2\right)\left(\Omega_3^2-\Lambda_3^2\right)-16\left(\vec{\alpha}_{\rm B}\cdot\vec{\beta}_{\rm B}\right)\Omega_3\Lambda_3=\left(\dfrac{\partial\vec{\Omega}_{\rm B}}{\partial z}+\dfrac{\partial\vec{\beta}_{\rm B}}{\partial\omega}\right)^2-\left(\dfrac{\partial\vec{\Lambda}_{\rm B}}{\partial z}+\dfrac{\partial\vec{\alpha}_{\rm B}}{\partial\omega}\right)^2 \end{cases}. \tag{11}$$

$(\Omega_3^2-\Lambda_3^2)$ and $\Omega_3\Lambda_3$ can be solved from Eq. (11). We already know that DGD and DAS can be expressed as [8]

$$\text{DGD}=\sqrt{\frac{1}{2}\left[\Omega^2-\Lambda^2+\sqrt{\left(\Omega^2-\Lambda^2\right)^2+4\left(\vec{\Omega}\cdot\vec{\Lambda}\right)^2}\right]},\quad \text{DAS}=\vec{\Omega}\cdot\vec{\Lambda}/\text{DGD}. \tag{12}$$

Based on Eqs. (1) and (11), we cam calculate $\Omega^2-\Lambda^2=(\Omega_{\rm L}^2-\Lambda_{\rm L}^2)+\Omega_3^2-\Lambda_3^2$ and $\vec{\Omega}\cdot\vec{\Lambda}=\vec{\Omega}_{\rm L}\cdot\vec{\Lambda}_{\rm L}+\Omega_3\Lambda_3$. It means DGD and DAS can be clearly achieved. In the telecommunication optical fiber links, optical fibers only have negligible PDL, and PDL mainly exists in some optical components. Then if the fiber section from z to $z+\Delta z$ has no PDL, Eq. (10) can be simplified as

$$\begin{cases} 4\beta_{\rm B}^2\Omega_3\Lambda_3=\left(\dfrac{\partial\vec{\Omega}_{\rm B}}{\partial z}+\dfrac{\partial\vec{\beta}_{\rm B}}{\partial\omega}\right)\cdot\dfrac{\partial\vec{\Lambda}_{\rm B}}{\partial z} \\ 4\beta_{\rm B}^2\left(\Omega_3^2-\Lambda_3^2\right)+\dfrac{4\beta_{\rm B}^2}{\omega^2}=\left(\dfrac{\partial\vec{\Omega}_{\rm B}}{\partial z}+\dfrac{\partial\vec{\beta}_{\rm B}}{\partial\omega}\right)^2-\left(\dfrac{\partial\vec{\Lambda}_{\rm B}}{\partial z}\right)^2 \end{cases}. \tag{13}$$

If there is no PDL in the whole optical fiber link under test, then we obtain

$$\Omega_3^2=\left[\left(\frac{\partial\vec{\Omega}_{\rm B}}{\partial z}+\frac{\partial\vec{\beta}_{\rm B}}{\partial\omega}\right)^2-\frac{4\beta_{\rm B}^2}{\omega^2}\right]/4\beta_{\rm B}^2. \tag{14}$$

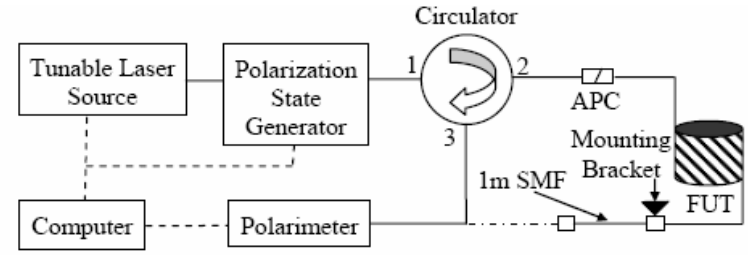


Figure 1: Experimental setup for single-end continuous-wave measurement of PMD.

To verify the Eq. (14), an experimental setup is shown in Fig. 1. The tunable laser source, polarization state generator and the polarimeter are controlled and synchronized by a computer. The polarization state at the port 1 of the circulator is tuned by the polarization state generator. At port 2, an APC connector is used to reduce the near-end Fresnel reflection. The FUT includes two 1 km-long single-mode fibers (SMF) and two 3 m-long Hi-bi fibers (Hibi), which are spliced together in the sequence of SMF-Hibi-SMF-Hibi. At the far-end of FUT, a mounting bracket is employed to guarantee the fiber far end is fixed in the measurement. And a 1 m-long SMF can be connected and disconnected with the FUT far end in order to induce the change of fiber length. Because the total fiber length is only approximately 2 km, the optical power contribution of Rayleigh backscattering can be neglected compared with the far-end Fresnel reflection [1].

To measure $\vec{\Omega}_{\rm B}$ and $\vec{\beta}_{\rm B}$, Mueller matrix method (MMM) is employed [9]. The test is divided into three steps. Firstly the polarimeter is connected to port 3 of the circulator and the 1 m-long SMF is connected with the far-end of the FUT. The wavelength is tuned from 1550 to 1555 nm with

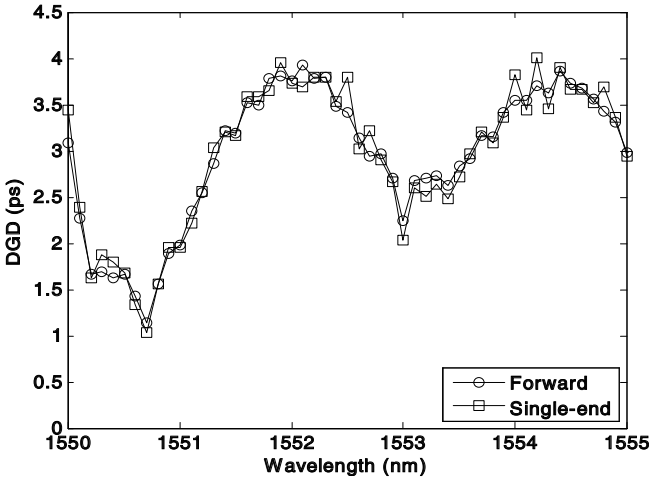


Figure 2: Comparison of DGD evolution in frequency domain measured using forward and single-end techniques.

0.1 nm step size and the polarimeter measures the evolution of polarization states at port 3 with respect to optical wavelength [9]. Secondly the 1m-long SMF is removed and then the measurement is repeated with the same parameters. Now DGD evolution with respect to optical wavelength can be calculated based on Eq. (14). Finally the polarimeter is connected to the far-end of FUT to repeat the measurement in the forward direction. Then the forward measurement presents the DGD value at every wavelength to compare with the one measured using single-end technique. The experimental results are shown in Fig. 2, and good agreement is observed. This confirms the validity our proposed single-end measurement method.

In the above measurement, two wavelength sweeping processes are performed and a short SMF is required to be connected and removed at the far-end of FUT. This is caused by the adoption of continuous-wave, which can make the experimental setup simpler. If a pulsed tunable laser source and an ultra-high speed polarization analyzer are employed, we can measure the polarization states of pulses reflected from two ends of the short SMF simultaneously, and then two above-mentioned issues can be overcome. But this will make the experimental setup more complex.

ACKNOWLEDGMENT

This work is partially supported by A*Star, Singapore under the project 042 101 0015 and OCLT, BUPT, P. R. China.

REFERENCES

1. Corsi, F., A. Galtarossa, L. Palmieri, M. Schiano, and T. Tambosso, "Continuous-wave backreflection measurement of polarization mode dispersion," *IEEE Photon. Technol. Lett.*, Vol. 11, 451–453, 1999.
2. Galtarossa, A. and L. Palmieri, "Spatially resolved PMD measurements," *IEEE J. Lightwave Technol.*, Vol. 22, 1103–1115, 2004.
3. Corsi, F., A. Galtarossa, and L. Palmieri, "Polarization mode dispersion characterization of single-mode optical fiber using backscattering technique," *IEEE J. Lightwave Technol.*, Vol. 16, 1832–1843, 1998.
4. Galtarossa, A. and L. Palmieri, "Theoretical analysis of reflectometric measurements in optical fiber link affected by polarization-dependent loss," *IEEE J. Lightwave Technol.*, Vol. 21, 1233–1241, 2003.
5. Barakat, R., "Theory of the coherency matrix for light of arbitrary spectral bandwidth," *J. Opt. Soc. Am.*, Vol. 53, 317–323, 1963.
6. Dong, H., P. Shum, M. Yan, G. Ning, Y. Gong, and C. Wu, "Generalized frequency dependence of output Stokes parameters in an optical fiber system with PMD and PDL/PDG," *Opt. Express*, Vol. 13, 8875–8881, 2005.
7. Gisin, N. and B. Huttner, "Combined effects of polarization mode dispersion and polarization dependent losses in optical fibers," *Opt. Comm.*, Vol. 142, 119–125, 1997.

8. Chen, L., O. Chen, S. Hadjifaradji, and X. Bao, "Polarization-mode dispersion measurement in a system with polarization-dependent loss or gain," *IEEE Photon. Technol. Lett.*, Vol. 16, 206–208, 2004.
9. Jopson, R. M., L. E. Nelson, and H. Kogelnik, "Measurement of second-order polarization-mode dispersion vectors in optical fibers," *IEEE Photon. Technol. Lett.*, Vol. 11, 1153–1155, 1999.

Advances in Signal Processing to Reduce Lift-off Noise in Eddy Current Tests

M. Cacciola[1], **A. Gasparics**[2], **F. C. Morabito**[1], **M. Versaci**[1], **and V. Barrile**[1]

[1]Universitá *Mediterranea* degli Studi di Reggio Calabria, Reggio Calabria, Italy
[2]Research Institute for Technical Physics and Materials Science, Budapest, Hungary

Abstract— Nowadays, Non Destructive Testing and Evaluation (NDT/NDE) are frequently used to evaluate integrity of manufactured articles in civil or industrial applications. In this framework, Eddy Current Test (ECT) technique has a primary role for inspection of conducting materials. Solution of many related inverse problems requires an accurate comprehension of measured data: it is one of the most challenging problem in this context. That's why a suitable signal pre-processing is necessary. In this paper, advanced signal processing techniques are evaluated in order to reduce the impact of lift-off effect on the eddy current data. Some hybrid approaches are also depicted, with encouraging results.

1. INTRODUCTION

Improving manufacturing quality and ensuring public safety are two of the most important targets in various applicative fields. Thus, components and structures are commonly inspected for early detection of defects or faults which may reduce their structural integrity. In experimental NDT/NDE, the available measurement data are explored, so that some clues may emerge in the inspection signal that are possibly representative of structural modification of the specimen, like cracks, flaws and corrosion. Practically solving of these inverse problems needs a suitable signal processing procedure, in order to mitigate the difficulties introduced by presence of noise into the measured data. In fact, all the measured experimental data is a composition of the informative pattern modified by non-informative variations, can be considered as the error or noise. The different noise sources can be classified as systematic (error due to the some know effects, like the lift-off effect) and non-systematic (uncorrelated or content independent sources, like the not-position but the time dependent noise during the data-acquisition period). In order to obtain suitable input data for the inverse problem solution, it is necessary to emphasize the informative pattern in the data set by filtering the noise. However, it is also critical to keep the information unchanged or unperturbed by the applied noise filtering techniques. Otherwise, the inverse problem solution will produce fraudulent result. Basically, the enhancement of the noise filtering can be characterized by a single parameter: the signal-to-noise ratio (SNR). However, it can reliable only if it is possible to assume that the informative content of the data set is untouched. In ECTs, one of the noise source is the so-called lift-off effect. The lift-off is defined as the distance between the specimen and the probe used for the inspection. During the scanning over the specimen surface, sensor lift-off is not constant. So, it non-linearly influences the measured data by introducing a typical low frequency and variable spatial distribution noise. Scientific literature proposes a lot of signal processing approaches exploited to reduce lift-off effect on ECTs (e.g., see [1–3] and references within); the main idea proposed in this work consists on an advanced signal processing approach based on the joint use of Independent Component Analysis (ICA) and Wavelet filtering techniques. In this way, it has been possible to improve the informative content of eddy current data, emphasizing useful signal and noteworthily increasing the SNR of measurements.

2. EXPERIMENTAL SET-UP AND COLLECTED DATA

In ECT, the probe allows for complex, time-varying voltage measurements and the presence of a flaw affects the formation of eddy currents. In our experimentations, probe is a single pancake exciting coil with Fluxset sensor® [4], provided by the Research Institute for Technical Physics and Materials Science (Fig. 1). Inspected specimen was a square plate $8 \times 8 \times 0.125$ cm made of INCONEL600 ($\sigma = 10^5$, $\mu = \mu_0$) material. It has a rectangular thin crack (9 mm in length and 0.2 mm in width) in the central region; depth of crack is 20% of the plate thickness. Defect is superficial (inner defect, ID) or hidden (outer defect, OD) according to the inspection side. The specimen was inspected by both ID and OD analysis, using a 110.7 mA (rms) exciting current at frequency 20 KHz. The amount of exciting coil turns is 93. The lift-off of the sensor core is

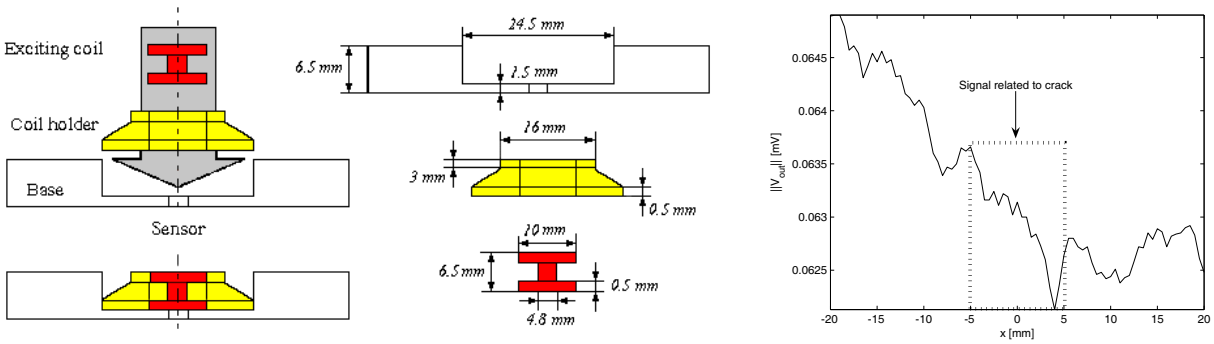


Figure 1: Layout of the Fluxset probe (left). $\|V_{\text{out}}\|$ measured at $Y = 0$, $-20 \leq X \leq 20$ in the ID analysis (right): it is possible to denote the lift-off noise into the signal trend.

about 0.75 mm. The measured quantity is the magnitude of output pick-up voltage ($\|V_{\text{out}}\|$). The offset signal was balanced out at the position: $X = 20\,\text{mm}$, $Y = 0\,\text{mm}$, i.e., the origin of used reference system. Scanned area was a $40 \times 40\,\text{mm}$ region, referred to the plate center, with 0.5 mm spacing along x and y direction ($-20\,\text{mm} \leq X \leq 20\,\text{mm}$, $-20\,\text{mm} \leq Y \leq 20\,\text{mm}$) [5]. A strong discontinuity in the homogeneity of the $\|V_{\text{out}}\|$ profile in a spatial location should evidence the presence of a defect in that zone of the volume. But, due to lift-off effect, irregularities caused by defects could not be clearly observable (Fig. 1). Therefore, a suitable signal pre-processing is necessary in order to remark the effective variation on $\|V_{\text{out}}\|$ caused by the defect.

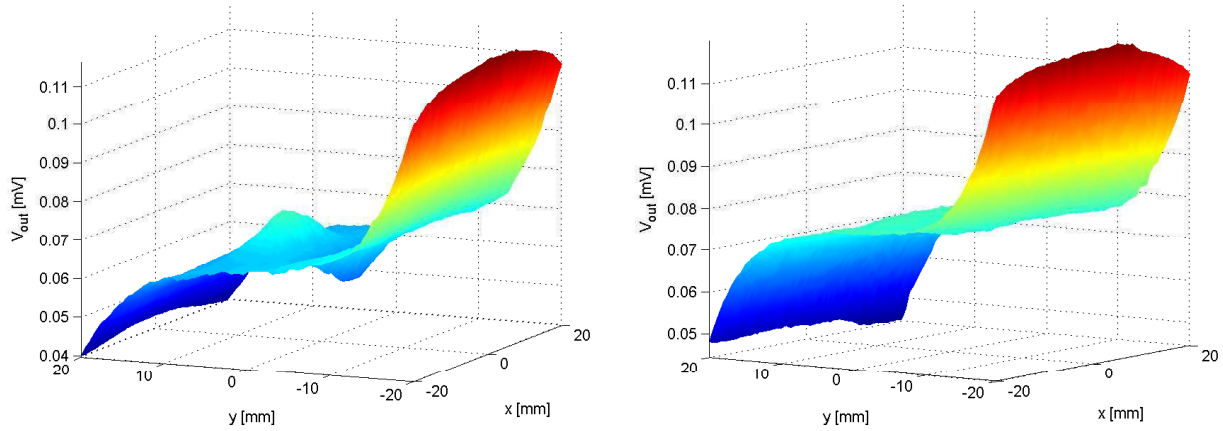


Figure 2: $\|V_{\text{out}}\|$ measured on the whole scanned area: ID analysis (at left), OD analysis (at right). Due to relatively small depth of defect, it is not easily visible by an OD analysis.

3. WT-BASED APPROACH FOR THE CASE-OF-STUDY

Wavelet approach measuring average fluctuations at different scales might prove less sensitive to noise than Fourier Transform. A suitable bank of discrete Wavelet filters can split low- and high-frequency contribute for a signal by a multi-resolution analysis in a Wavelet domain. In this way, global dynamics of signal $f(x)$ can be condensed in a series of Wavelet coefficients related to the low frequencies, i. e., the approximation coefficients at the higher multi-resolution level, A_M. On the other hand, local oscillations of $f(x)$ are depicted in a set of Wavelet details coefficients at different scales ($D_j, j = 1, 2, \ldots, M$), related to the high frequencies. In this way $f(x) = A_M(x) + \sum_{j=1}^{M} D_j(x)$. This wavelet decomposition enables the noise in the signal to be separated from the useful components. For the considered case-of-study, 10th-order Daubechies Wavelet [6] (Db10) has been used, decomposing the signal into five frequency levels by the Discrete WT (DWT). The high regularity of Db10 assures to restrict the lift-off noise into A_5. Thus, usage of the Inverse DWT (IDWT) excluding A_5 can reduce the lift-off effect on eddy current signal [1]: Fig. 3 shows both the WT-based denoising for $Y = 0$, $-20\,\text{mm} \leq X \leq 20\,\text{mm}$ of ID scanning, and the iteration

of the whole procedure for the ID and OD scannings. Due to small visibility of crack into OD data, also the related WT reconstruction is not recommendable for an useful experimentation and will be excluded in the prosecution of the work.

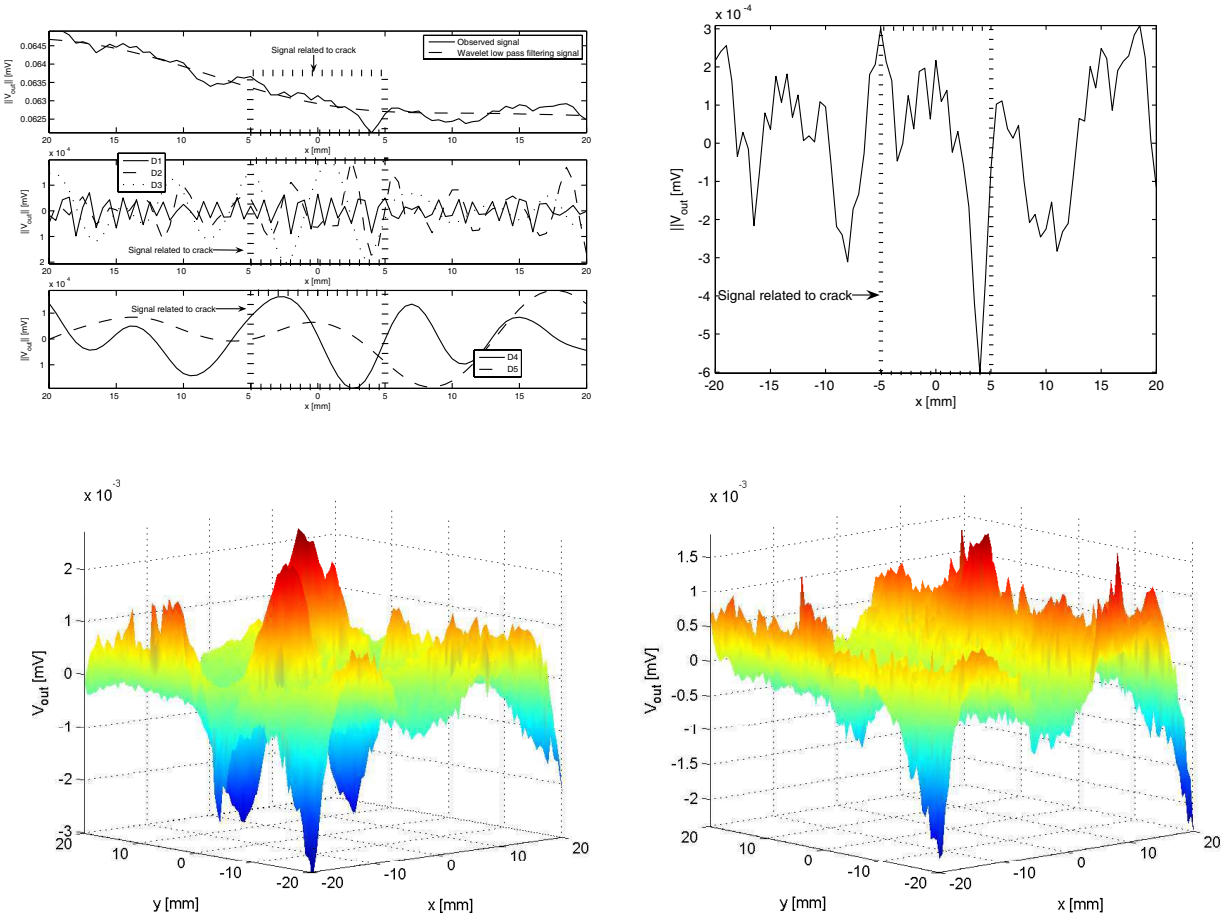


Figure 3: At top: Db10 decomposition of $\|V_{\text{out}}\|$ for the ID analysis (left); $\|V_{\text{out}}\|$ reconstruction without considering A_5 coefficients (right). At bottom: results of WT-based denoising on $\|V_{\text{out}}\|$ for ID (left) and OD (right) analysis.

4. ICA: THEORETICAL FUNDAMENTALS AND EXPLOITATIONS FOR THE CASE-OF-STUDY

Let us consider a signal $\mathbf{s}(t)$ in time-domain, which is registered by j receivers and is the result of mixing of j sources, such that $\mathbf{s}(t) = \{\mathbf{s}_1(t), \mathbf{s}_2(t), \ldots, \mathbf{s}_j(t)\}$, where $\mathbf{s}_k(t) = \sum_{h=1}^{j} a_{kh}\mathbf{x}_h(t)$. Under particular assumption, it is possible to recover the set of j sources by calculating a suitable mixing matrix $\mathbf{A}$, i. e., the matrix with elements a_{kh}. It is the aim of ICA, where each mixture $\mathbf{s}_k$ as well as each Independent Component (IC) $\mathbf{x}_h$ is a random variable. Mixtures must have zero mean and unitary variation (whitening process). Once $\mathbf{A}$ is calculated, it is possible to obtain its inverse $\mathbf{A}^{-1}$ and retrieve ICs having non-gaussian distributions [7]. In the case-of-study, signal $\mathbf{s}$ is considered as the mixing of $\|V_{\text{out}}\|$ for both ID and OD analysis. The FastICA software based on fixed-point algorithm [8] has been used, by exploiting the function $g = \tanh(a_1 \cdot u)$; $a_1 = 1.7$ to estimate the differential entropy (i. e., negentropy). Results of ICA-based pre-processing is shown in Fig. 4: it is possible to discriminate the component related to lift-off noise to the one representing the useful denoised $\|V_{\text{out}}\|$ signal. Compared with WT-based approach, the ICA method shows less micro-oscillations but higher lift-off at the top and bottom side of inspected area.

5. PROPOSED HYBRID WAVICA AND ICAWAV APPROACHES

The WT- and the ICA-based approaches have been subsequently mixed, in order to obtain an hybrid approach able to improve performances of exploited techniques avoiding the disadvantages of the two methods. In a first attempt, a 5th-level Db10 based Wavelet analysis has been initially taken into account to improve the SNR of both ID and OD ECTs; subsequently such denoised

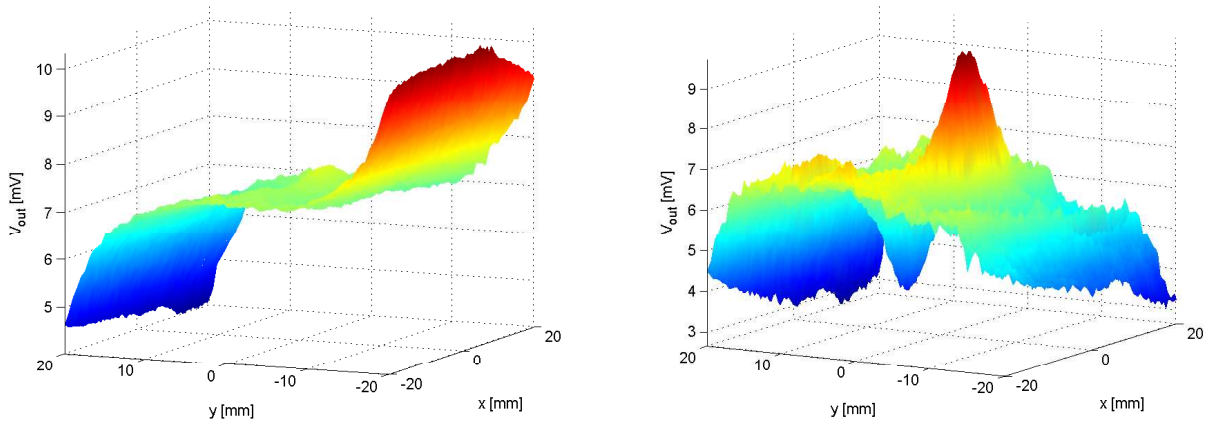


Figure 4: ICs of measured data: it is possible to denote how the component at right represents $\|V_{\text{out}}\|$ denoised by ICA.

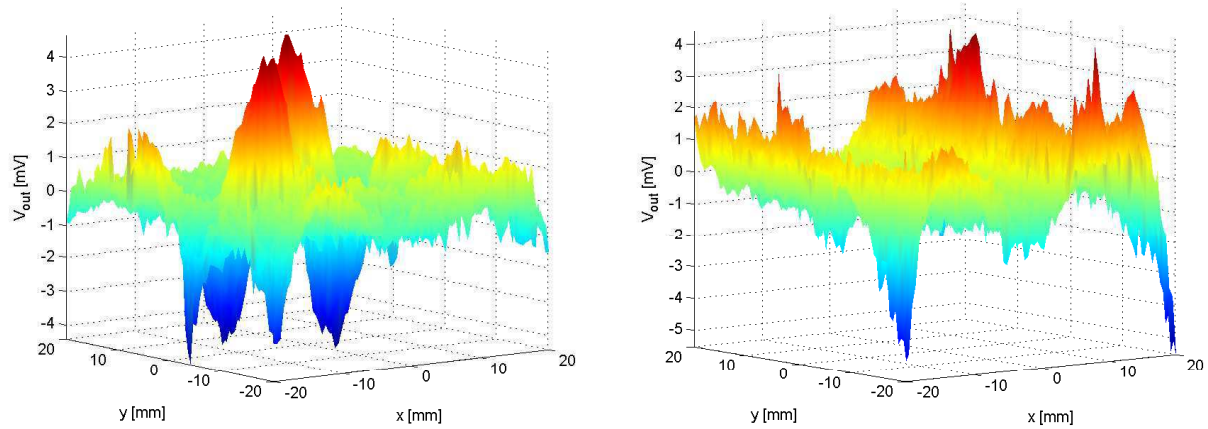


Figure 5: ICs of measured data: it is possible to denote how the component at left represents $\|V_{\text{out}}\|$ denoised by proposed WAVICA approach.

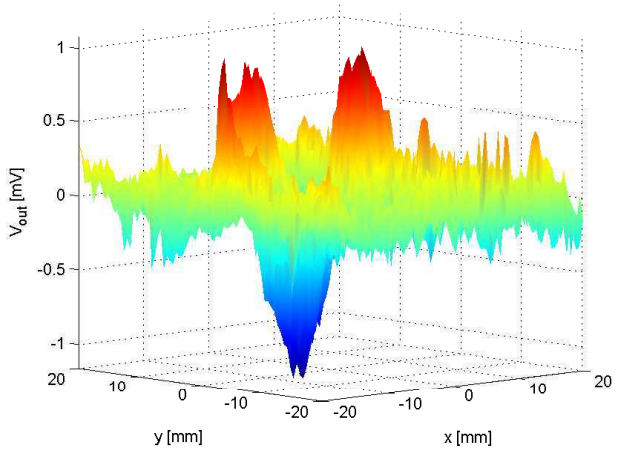


Figure 6: Results of ICAWAV-based denoising on $\|V_{\text{out}}\|$ for the ID analysis. The OD analysis has not been considered due to the previously presented results of the WT-based approach.

signals have been mixed in order to extract the lift-off component by using the ICA: it is the WAVICA approach. Results are shown in Fig. 5. In this way, it has been possible to improve the quality of ECT's measurement by reducing the amount of whole noise. Some peaks on $\|V_{\text{out}}\|$

are still present at the bottom-left corner of inspected area. Therefore, a further approach has been considered. In this case, the ICA has been firstly carried out on the mixture of ID and OD ECTs, and subsequently the useful component has been processed by means of Db10 WT at five multi-resolution scales (ICAWAV approach). So, it has been possible to obtain the best results (see Fig. 6), emphasizing the influence of defect in output pick-up voltage and strongly lowering the lift-off noise even at the borders (and corners) of examined area.

Table 1: SNR obtained by different exploited pre-processing techniques.

$\mathbf{SNR}_{ID}$	$\mathbf{SNR}_{WT}$	$\mathbf{SNR}_{ICA}$	$\mathbf{SNR}_{WAVICA}$	$\mathbf{SNR}_{ICAWAV}$
14.623	19.129	16.954	19.127	19.192

6. DISCUSSIONS ABOUT RESULTS AND CONCLUSIONS

In this paper, some advanced techniques have been evaluated to reduce the lift-off noise in eddy current NDT/NDE. Besides WT- and ICA-based procedures, hybrid WAVICA and ICAWAV approaches have been proposed, obtaining a noteworthily decrement of lift-off effect. In terms of SNR, the gain retrieved by proposed hybrid approaches is between 4.504 dB and 4.569 dB, if compared with SNR of the original ID signal (SNR_{ID}). Moreover, ICAWAV technique assures a raising on SNR of 0.063 dB and 2.238 dB compared to WT and ICA respectively. Table 1 resumes all obtained SNR values.

ACKNOWLEDGMENT

Authors are very pleased to thank all of the participants of the MANODET project, for providing on Web the dataset and information about Fluxset #6/6 used in this work [5].

REFERENCES

1. Chen, G., A. Yamaguchi, and K. Miya, "A novel signal processing technique for eddy-current testing of steam generator tubes," *IEEE Trans. on Magnetics*, Vol. 34, No. 3, 642–648, 1998.
2. Morabito, F. C., "Independent component analysis and feature extraction techniques for NDT/NDE data," *Materials Evaluation*, Vol. 58, No. 1, 85–92, 2000.
3. Fiori, S. and P. Burrascano, "ECT-data fusion by the independent component analysis for non-destructuve evaluation of metallic slabs," *Third International Conference on Independent Component Analysis and Signal Separation*, 323–327, San Diego, California, 2001.
4. Gasparics, A., C. S. Daroczi, G. Vertesy, and J. Pavo, "Improvement of ECT probes based on Fluxset-type magnetic field sensor," *Electromagnetic Nondestructive Evaluation*, No. 2, 146–151, 1998.
5. Gasparics, A., Probe_98/6 — Surface scan results within MANODET Project, available at http://alag3.mfa.kfki.hu/ect/manodet/results/FS06B/index.htm, October 13, 2006.
6. Daubechies, I., "Ten lectures on wavelets," *CBMS-NSF Series in Applied Mathematics*, 1992.
7. Hyvarinen, A. and E. Oja, "Independent component analysis: Algorithms and applications," *Neural Networks*, Vol. 13, No. 4–5, 411–430, 2000.
8. Hyvarinen, A., "Fast and robust fixed-point algorithms for independent component analysis," *IEEE Trans. on Neural Networks*, Vol. 10, No. 3, 626–634, 1999.

A Compact Printed Antenna with Band-stop Characteristic for UWB Application

Seokjin Hong, **Heejun Lee**, **Soonyong Lee**, **and Jaehoon Choi**

Department of Electrical and Computer Engineering, Hanyang University
17 Haengdang-Dong, Seongdong-Gu, Seoul, 133-791, Republic of Korea

Abstract— A Compact antenna with a band-stop characteristic for ultra-wideband application is proposed. To increase the impedance bandwidth of the proposed antenna, a tapered ground and circular slot are used. By attaching an inverted Y-shaped open-circuited stub in the circular ring, the band-stop characteristic is obtained. The designed antenna satisfies the voltage standing wave ratio requirement of less than 2.0 in the frequency between 3.12 GHz and 10.73 GHz while showing the band stop performance in the frequency band of 5.15 GHz to 5.78 GHz. This antenna shows an omnidirectional radiation pattern similar to that of a monopole antenna.

1. INTRODUCTION

With the advent of ultra-wideband (UWB) communication system operating from 3.1 GHz to 10.6 GHz, there has been increasing demand for antenna capable of operating at an extremely wide frequency range. Planar monopole antennas have received much attention due to their attractive merits such as wide impedance bandwidth, simple structure, and omnidirectional radiation pattern [1, 2]. However, due to the size limitation, additional efforts to miniaturize the total size are needed to incorporate an antenna inside portable terminals such as PDAs, laptop computer, and so on.

Since the existing wireless LAN service band (5.15 GHz–5.825 GHz) overlaps the UWB service band, UWB radio signal can interfere with those of WLAN band service. To overcome this problem, UWB antennas with good band-stop performance are necessary. Several UWB antennas with the band-stop characteristic have been proposed [3, 4].

In this paper, the design of a compact printed antenna with band stop characteristic for UWB application is presented. To achieve the wideband characteristic, a circular slot and tapered ground are used. Band-stop characteristic is achieved by attaching an inverted Y-shaped open-circuited stub in the circular ring. The design procedure is described, and simulated and measured results are presented.

2. ANTENNA DESIGN

The configuration of the proposed antenna is shown in Figure 1. The antenna is printed on the

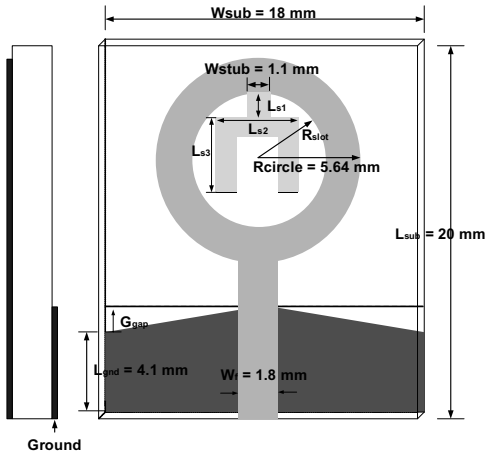


Figure 1: Geometry of the proposed antenna.

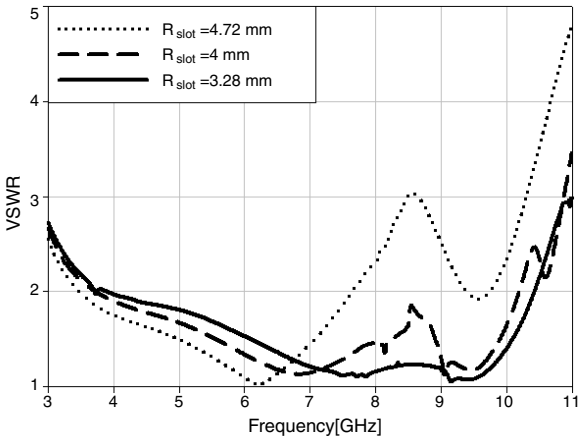


Figure 2: Simulated VSWR for different R_{slot}.

FR4 substrate with height(h) of 1.6 mm and relative permittivity of 4.4 and has a dimension of 18 mm × 20 mm ($W_{\text{sub}} \times L_{\text{sub}}$).

The antenna consists of annular ring patch, tapered ground and an inverted Y-shaped open-circuited stub. In this antenna design, a 50 Ω microstrip feed line is used to excite the annular ring

patch. The radius (R_{slot}) of an annular ring slot is a key factor to control the impedance bandwidth at higher band and the gap distance (G_{gap}) of a tapered ground is a prime element to decide the impedance bandwidth at lower band. To obtain the band-stop characteristic, an inverted Y-shaped open-circuited stub is attached to the circular ring. A band-stop characteristic is controlled by the length of an inverted Y-shaped open-circuited stub.

3. SIMULATIONS AND MEASUREMENTS ANALYSIS

The antenna is modeled numerically by using Ansoft's High-Frequency structure simulator(HFSS) [5]. Figure 2 shows the simulated VSWR characteristic of a proposed antenna for different circular slot radii (R_{slot}). In this design, the circular slot provides the mechanism to enhance the impedance bandwidth for the higher band. Figure 3 shows the VSWR characteristics for various value of G_{gap}. It is observed that the lower frequency is significantly affected by the variation in length G_{gap}. On the other hand, the higher band is insensitive to the change of G_{gap}.

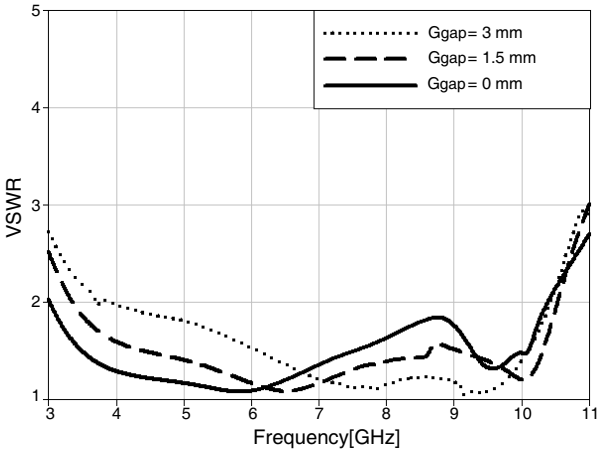


Figure 3: Simulated VSWR for different G_{gap}.

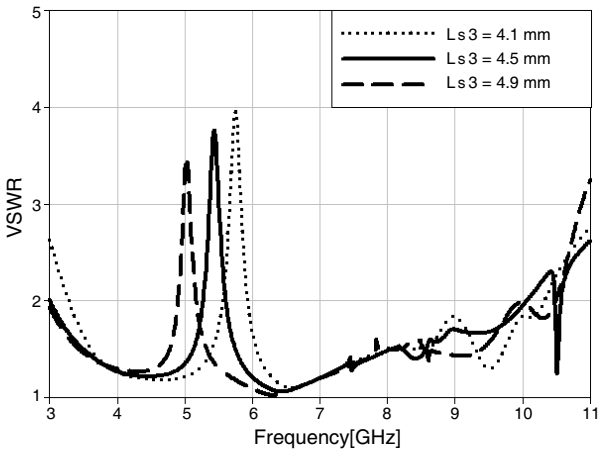


Figure 4: Simulated VSWR for various L_{s3}.

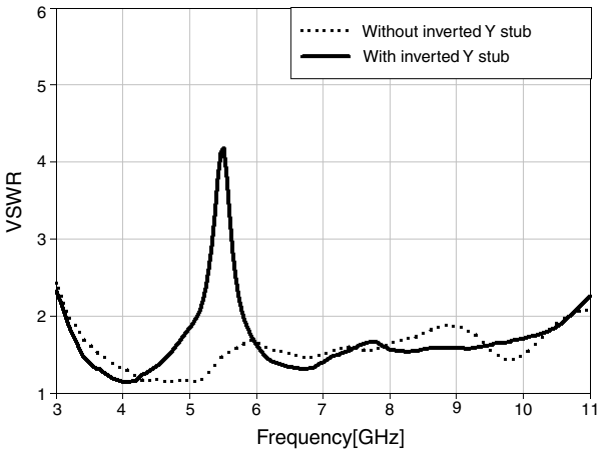


Figure 5: Measured VSWR for the proposed antenna.

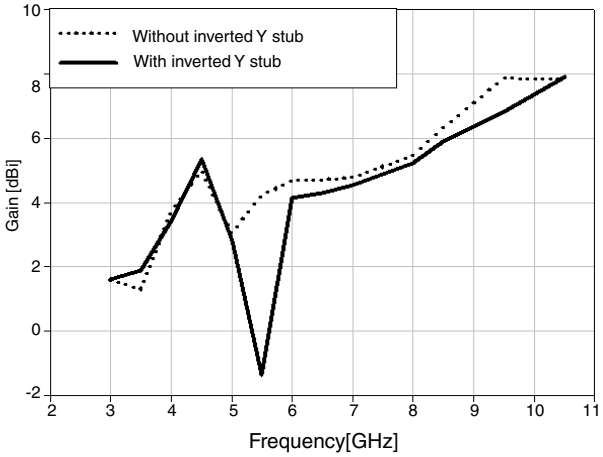
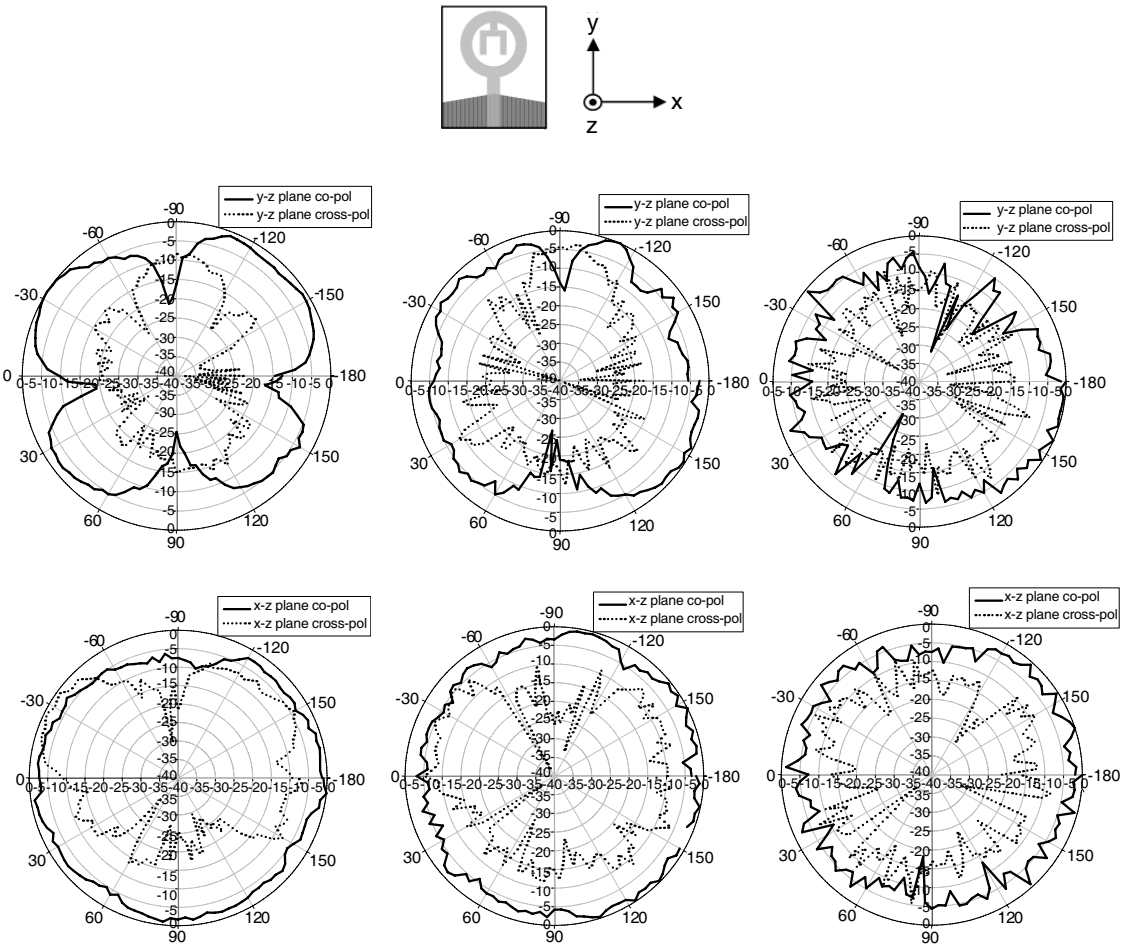


Figure 6: Measure antenna gain.

To further demonstrate the effect of an inverted Y-shaped open-circuited stub length on the band-stop frequency, an inverted Y-shaped stub is attached to the circular ring. Figure 4 shows the simulated results for the proposed antenna with various values of an inverted Y-shaped open-circuited stub length (L_{s3}). From the comparison of the results shown in Figure 4, one can conclude that the band-stop frequency for the proposed antenna is controlled by an inverted Y-shaped open-circuited stub length, which functions as a half-wavelength resonant structure. Figure 5 shows the measured VSWR characteristic with and without an inverted Y-shaped open-circuited stub. After adding a circular slot and tapered ground, the measured bandwidth of the antenna ranges from

3.12 GHz to 10.73 GHz for VSWR less than 2. This antenna has band-stop band of 5.15 GHz to 5.78 GHz after an inverted Y-shaped open-circuited stub is attached to the circular ring.

Figure 6 presents the measured antenna gain. As shown in Figure 6, gain decreases drastically at the band-stop frequency of 5.47 GHz. The measured radiation patterns in the y–z plane and x–z plane at 4 GHz, 6 GHz and 9 GHz are illustrated in Figures 7(a), (b) and (c), respectively. Dipole-like radiation patterns in the y–z plane and nearly omni-directional radiation pattern in the x–z plane are observed for copolarization.



(a) 4 GHz (b) 6 GHz (c) 9 GHz

Figure 7: Measured radiation patterns.

4. CONCLUSION

A compact antenna with band-stop characteristic has been proposed and implemented for UWB applications. The proposed antenna has a simple configuration and is easy to fabricate. To obtain the wide bandwidth characteristic, a circular slot and tapered ground are used. The dimensions of a circular slot and tapered ground were optimized by the parametric analysis and resulted in wide bandwidth performance from 3.12 GHz to 10.73 GHz. Furthermore, by attaching an inverted Y-shaped open-circuited stub into the circular slot, the band-stop characteristic is obtained in WLAN (5.15 GHz∼5.78 GHz) while the wide impedance bandwidth is maintained.

This research was supported by the MIC (Ministry of Information and Communication), Korea, under the ITRC (Information Technology Research Center) support program supervised by the IITA (Institute of Information Technology Assessment)(IITA-2005-(C1090-0502-0030)).

REFERENCES

1. Agrawall, N. P., G. Jumar, and K. P. Ray, "Wideband planar monopole antennas," *IEEE Trans. Antennas Propagat.*, 46, 294–295, 1998.

2. Ammann, M. J. and Z. N. Chen, "A wide-band shorted planar monopole with bevel," *IEEE Trans. Antennas Propagat.*, Vol. 51, 901–903, 2003.
3. Schantz, H. G., G. Wolenec, and E. M. Myszka, "Frequency-notched UWB antennas," *IEEE Conf Ultra-Wideband Syst. Technol.*, 214–218, Reston, VA, 2003.
4. Kerhoff, A. and H. Ling, "A parametric study of band-notched UWB planar monopole antenna," *IEEE Int. Symp. Antennas Propagat.*, 1768–1771, Monterey, CA.
5. "Ansoft High-frequency Structure Simulator (HFSS)," Ver. 10, Ansoft Corporation.

A Low-power Quadrature VCO Using Current-reused Technique and Back-gate Coupling

Yu-Ching Tsai, **Yi-Shing Shen**, **and Christina. F. Jou**
Chiao-Tung University, Taiwan, China

Abstract— A novel quadrature voltage controlled oscillator (QVCO) using current-reused technique and back-gate coupling is presented. The QVCO realized with LC-tank is demonstrated in a 0.18 um RF CMOS 1P6M process. Using the current-reused technique efficiently reduces the power dissipation. However, by stacking switching transistors in series like a cascade, the architecture cannot perform well in phase noise. Through an improved circuit schematic with back-gate coupling, the phase noise of the circuit can be lowered. As a result of reducing four transistors in the circuit, the total power dissipation can be cut down even more. The simulation shows the phase noise is around −107 dBc/Hz at 1 MHz offset and the output frequency tuning range of the fabricated QVCO is 600 kHz ranging from 5.15 to 5.75 GHz for 802.11a. The circuit draws only 1.5 mA from a 1.8-V supply. Compared with the recent works, the proposed topologies show a better phase noise performance and can be adopted for low-power applications.

1. INTRODUCTION

For modern low cost wireless systems such as WLAN 802.11a a high level of integration and low cost are significant criterions for the transceiver architecture. The development of single-chip CMOS solutions for the 5-GHz 802.11a wireless local area network standard is desirable to enable implementations at low cost. In today's complex radio architecture, it is often necessary for the local oscillator in a transceiver to produce two tones with a quadrature phase relationship. The voltage-controlled oscillator has been the subject of intense study. Recently, a CMOS VCO can be implemented by using a ring structure or an LC tank [1]. Both of them are beneficial and detrimental. By definition, the ring VCO is known for a wide tuning range but drawback of a high phase noise that it is disqualified for most structures in modern RF transceivers. Accordingly, most investigations are focused on the LC VCO design. Owing to the limited range of the variable capacitance of the varactor, the tuning range of CMOS LC VCOs is very small. Hence, Razavi [2] proposed that the varactor controlled by the DC potentials at both of its input terminals to enable the LC VCO yielding a wide tuning range. In recent years, several techniques exist to generate quadrature [1–4]. Each method has its own advantages and disadvantages. Different implementations of quadrature oscillators can be found in the literature:

1) Using four-delay stage ring oscillator to fulfill the quadrature phase. This can be done by using a polyphase filter to generate the signals [1].
2) A VCO followed by a passive RC complex filter. The main drawback of this solution is that a power-hungry buffer is needed between the VCO and the filter [3].
3) A VCO running at double frequency followed by a digital frequency divider based on flip-flop.
4) Two oscillators are injection locked in quadrature by coupling their second-order harmonic in anti-phase, and using a coupling network that exhibits high odd-mode and low even-mode impedance [4].
5) Two crossed-coupled VCO forced to run in quadrature by using coupling transistors. In [5] additional coupling transistors in parallel with the tank of two differential oscillators are used for quadrature generation (P-QVCO). Alternative solutions are the series-coupled QVCO (S-QVCO) topologies as proposed in [6]. That exhibits improvements in terms of phase robustness and phase noise compared to the P-QVCO.

With the demand for low cost and high integration of wireless transceiver building blocks, here, we present an alternative method to obtain quadrature oscillator based on the current-reused technique. By stacking switching transistors in series like a cascade, the structure reduces the power dissipation. Nevertheless, the stacking switching transistors in series suffer from an increase in phase noise. In this work, the original coupling transistors are modified by using back-gate coupling to improve the phase noise. A new QVCO using current-reused technique and back-gate coupling is demonstrated here.

2. CIRCUIT BLOCKS DESCRIPTION

Figure 1 shows the proposed QVCO topology. This VCO adopts a cross-coupled negative-gm configuration since the structure offers higher trans-conductance for a given current, which results in fast switching of the cross-coupled pair. The cross-coupled VCO operates as switches and the topology consists of two identical VCO in series. Each one is composed of switching transistors M_{sw}, varactors C_v and inductors L. In order to analyze this circuit, the following introduction is divided into two parts: (1) the current-reused technique. (2) The back-gate coupling.

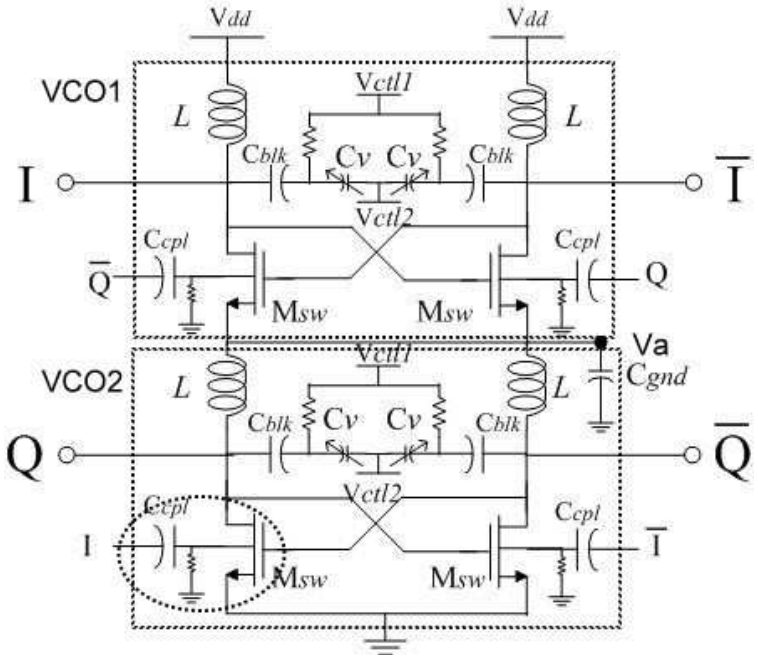


Figure 1: Schematic of the proposed quadrature VCO.

(1) The Current-reused Technique

The main advantage of current-reused technique is to lower the power dissipation [7]. However, a frequency tuning circuit is designed to control the two sides of the varactors since the dc voltages are different between two differential output nodes (I/Q). By way of adding capacitors C_{blk} for dc block and ac short, the voltage drop across the varactors will be identical. Moreover, we use two voltages (Vctl1 and Vctl2) to control the LC tank. The control voltage Vctl1 is adopted for tuning range and the second control voltage Vctl2 is adopted for adjusting the center of the linear region so as to maximize the tuning range. Therefore, the circuit is equivalent for two identical VCO in ac and two VCO in series like a cascade in dc. In general, a differential VCO uses a current source to start oscillators reliably and the phase noise of the VCO is increased. The node Va is operated at the double frequency of the oscillation frequency of the LC tank VCO, since the node Va is pulled up when each one of the differential NMOS turn on. As a result, the large capacitor Cgnd is adopted for ac ground to help suppress the injection of the high-frequency noise by canceling the large-amplitude harmonics generated at the node Va of the oscillator core. A VCO without a current source can operate with a low power supply. In the proposed QVCO, the current source is removed to avoid the flicker noise contributed by the current source.

The main issue for recent QVCO approach is to achieve a monolithic integration between low phase noise with a wide frequency tuning range and low power consumption at given operating frequencies. As mentioned above, although power consumption can be decreased to a half by reusing the dc current compared to the conventional QVCO topologies, it will increase the phase noise. Due to stacking switching transistors in series like a cascade, the architecture cannot perform well in phase noise. Hence, we use the back-gate coupling to improve the phase noise.

(2) The Back-gate Coupling

The useage of back-gates removes the additional noise contributions compared to the conventional coupling transistor based topology [8]. The coupling transistors are removed and the two differential VCOs are coupled through the back-gate of the core NMOS transistors. The capacitors

Ccpl are adopted for ac coupling and dc block. Fig. 3(a) shows the small-signal equivalent circuit for the circled part of the conventional QVCO circuit in Fig. 2(a). Fig. 3(b) shows the small-signal equivalent representation for the circled part of the conventional back-gate QVCO circuit in Fig. 2(b). Obviously, Vgs1 and Vgs2 in Fig. 3(a) are the counterparts of the Vgs3 and Vbs3 In Fig. 3(b). Similarly, gm1 and gm2 are the matching components of gm3 and gmb3. It can be seen that the two small-signal circuits shown in Fig. 3(a) and Fig. 3(b) are essentially the same. Therefore, the back-gate coupling can be seen as the coupling transistors. In the other hand, the conventional coupling transistors dissipate 30%∼100% of the power dissipated in the core transistors, which increases the total power dissipation. It goes without saying that by utilizing the back-gate, the additional noise source and the power dissipation can be avoided altogether.

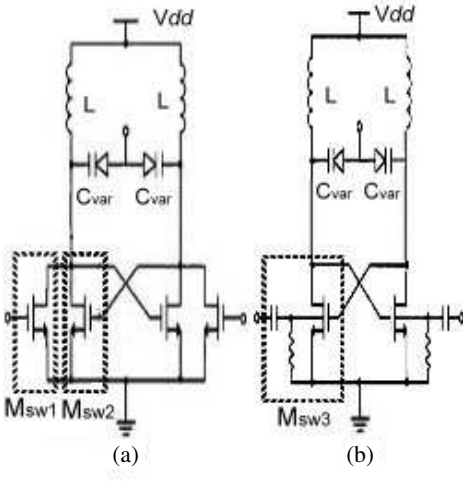


Figure 2: (a) The Conventional QVCO. (b) The conventional back-gate coupling.

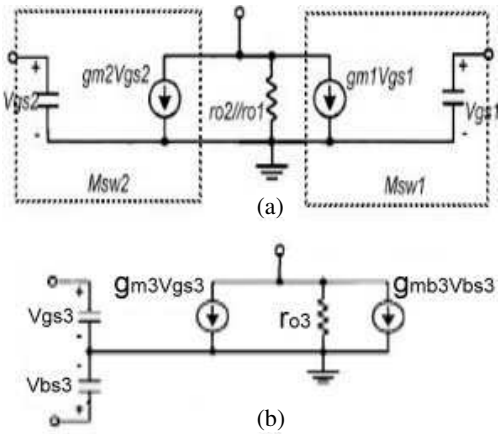


Figure 3: (a) Small-signal equivalent circuit in Fig. 2(a). (b) Small-signal equivalent circuit in Fig. 2(b).

3. SIMULATION RESULT

Figure 4 shows the layout of the proposed QVCO fabricated in TSMC 0.18 μm RF CMOS 1P6M process. The chip sizes are $1.05 \times 1.00\,\mathrm{mm}^2$ including the pads and Fig. 5 shows the quadrature output. According to Fig. 6, the output frequency tuning range of the fabricated QVCO is 600 kHz ranging from 5.15 GHz to 5.75 GHz from the center frequency of 5.25 GHz. Fig. 7 shows the phase noise of -107 dBc/Hz at 1 MHz offset. The core current consumed 1.5 mA from a 1.8 V supply. The figure-of merit (FOM) characteristic is expressed as [11].

Table 1: Comparison with the recent published papers about wideband VCOs.

QVCO	QVCO	Frequency [GHz]	Core Current [mA]	Phase Noise [dBc/Hz]	FOM
[7]	0.18	3.1	2.4	−102	−166
[8]	0.18	1.1	3	−120	−173
[9]	0.18	6	6.8	−106	−173
[10]	0.13	10	8	−95	−163
This Work	**0.18**	**5.25**	**1.5**	**−107**	**−177**

$$\mathrm{FoM} = 10\log\left[\left(\frac{f_0}{\Delta f}\right)^2 \frac{1}{L(\lambda)P_{\mathrm{supply}}}\right]$$

Here, f_o is the oscillation frequency of the VCO, Δf is the offset frequency, and $\mathrm{P_{supply}}$ is the power consumption. In order to compare with the recent published papers, the power-frequency-normalized figure of merit (FOM) is calculated. According to Table 1, the performance of the proposed QVCO is an excellent one for low-power topology compared with the reported results in the literature [7–10].

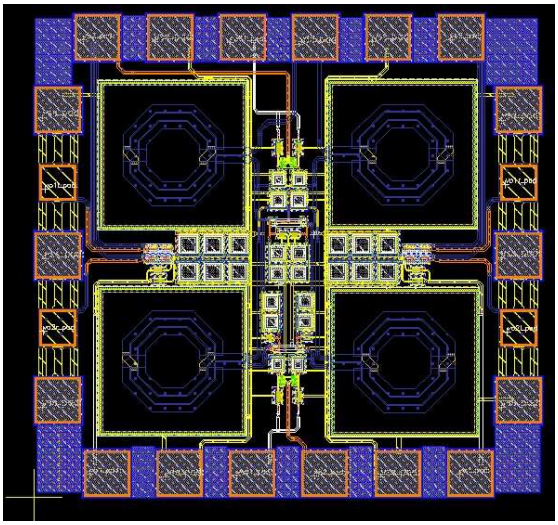


Figure 4: Layout of the QVCO.

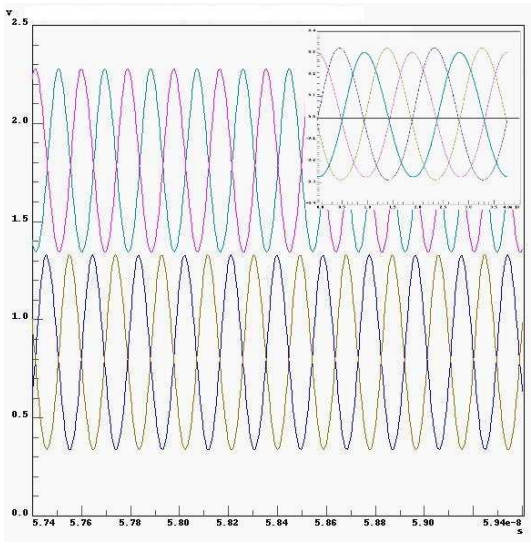


Figure 5: Quadrature output from core QVCO and buffer.

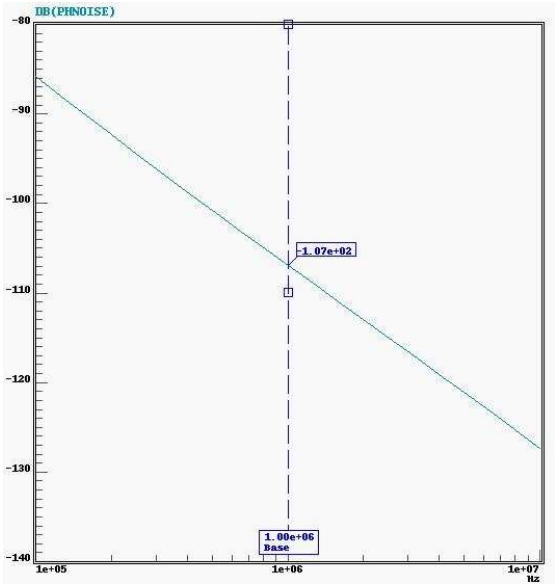


Figure 6: Simulation phase noise of this QVCO.

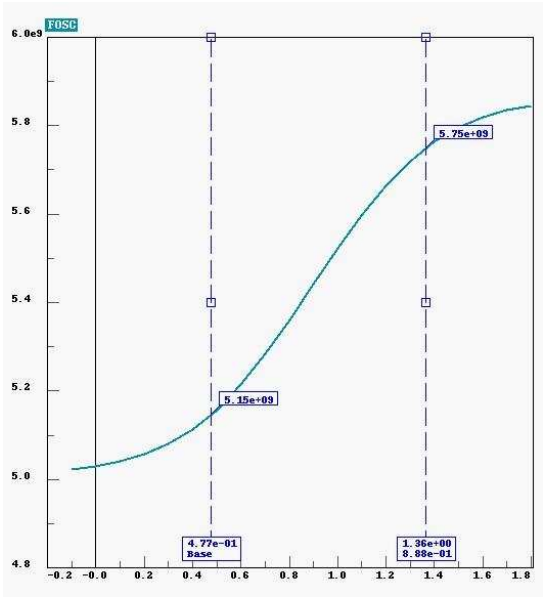


Figure 7: Turning range of this QVCO.

4. CONCLUSION

A low-power quadrature VCO using current-reused technique and back-gate coupling has been presented and fabricated by TSMC 0.18 μm RF CMOS 1P6M process. The tuning range of the QVCO is about 600 KHz for the center frequency of 5.25 GHz and the phase noise at 1 MHz offset is -107 dBc/Hz while dissipating only 1.5 mA for the QVCO from 1.8 supply. Due to stacking switching transistors in series like a cascade and utilizing the back-gate of the core transistors, this work improves the drawback in phase noise of the current-reused QVCO and shows a major advantage in power dissipation compared with overall performances of the conventional QVCOs.

REFERENCES

1. Vancorenland, P. and M. Steyaert, "A 1.57 GHz fully integrated very low phase noise quadrature VCO," *2001 Symposium on VLSI Circuits, Digest of Technical Paper*, 111–114, 2001.
2. Razavi, B., "A 1.8 V GHz CMOS voltage-controlled oscillator," *IEEE Int. Solid-state Circuit Conf. (ISSC) Dig. Tech. Papers*, 388–389, San Francisco, CA, Feb. 1997.
3. Crols, J. and M. Styaert, "A fully integrated 900 MHz CMOS double quadrature downcon-

verter," *IEEE Int. Solid-state Circuit Conf. Dig. Tech. Papers*, 136–137, San Francisco, CA, Feb. 1995.

4. Gierkink, S. L. J., S. Levantino, R. C. Frye, "A low-phase-noise 5-GHz CMOS quadrature VCO using superharmonic coupling," *IEEE Journal of Solid-state Circuits*, Vol. 38, No. 7, July 2003.
5. Rofougaran, A., J. Rael, M. Rofougaran, and A. Abidi, "A 900 MHz CMOS LC-oscillator with quadrature outputs," *IEEE JSSCC Dig. Tech. Papers*, 392–393, February 1996.
6. Andreani, P., "A low-phase-noise, low-phase-error 1.8 GHz quadrature CMOS VCO," *IEEE JSSCC Dig. Tech. Papers*, 290–291, February 1996.
7. Oh, N.-J. and S.-G. Lee, "Current reused LC VCOs," *IEEE Microwave and Wireless Component Letters*, Vol. 15, No. 11, November 2005.
8. Kim, H.-R., C.-Y. Cha, S.-M. Oh, M.-S. Yang, and S.-G. Lee, "A very low-power quadrature with back-gate coupling," *IEEE Journal of Solid-state Circuit*, Vol. 39, No. 6, June 2004.
9. Shin, H., Z. Xu, and M. F. Chang, "A1.8-V 6/9-GHz switchable dual-band quadrature LC VCO in SiGe BiCMOS technology," *IEEE Radio Frequency Integrated Circuits Symposium*, 71–74, 2–4 June 2002, DOI:10.1109/RFIC.2002.1011513.
10. Li, S., I. Kipnis, and M. Ismail, "A 10-GHz CMOS quadrature LC-VCO for multirate optical applications," *IEEE Journal of Solid-State Circuits*, 1626–1634, October 2003.
11. Kinget, P., *Integrated GHz Voltage Controlled Oscillator*, 353–381, Kluwer, Norwell, MA, 1999.

A Fully-integrated, Low Power, Fast-locking, Integer-N Frequency Synthesizer for MB-OFDM UWB System

Shih-Hao Tarng, Yu-Ching Tsai, and Christina. F. Jou
Chiao-Tung University, Taiwan, China

Abstract— A fast-locking frequency synthesizer for Band Group 1, 3 and 4 in MB-OFDM UWB system is designed using TSMC 0.18 μm RF CMOS 1P6M process. Based on the simulation results, the circuit consumes 46.35 mW. A wide tuning range of 6.28∼9.17 GHz is achieved. The best phase noise is −113.6 dBc/Hz at 1 MHz offset and the locking time is smaller than 300 nsec. Compared with the reference paper, the power dissipation is almost halved while all specifications are met.

1. INTRODUCTION

For the next generation of WPAN, MultiBand orthogonal frequency division multiplexing (MB-OFDM) UltraWideBand (UWB) is proposed to raise the data rate up to 480 Mbps. According to the proposal [1], the carrier frequencies are allocated in the 3.1∼10.6 GHz spectrum. Because the switch time is only several nano seconds in the proposal, the frequency synthesizers for UWB need some modifications. Here a topology is proposed . Two fast-settling frequency synthesizers are used to generate the desired signal by turns [2]. As proposed in [1], the symbol interval is 312.5 nsec and the guard time is 9.47 nsec. Therefore, a single PLL has to be locked within 322 nsec. This approach is practical for a conventional frequency synthesizer which is easy to be implemented. Here, in this paper, we present a frequency synthesizer to cover Band Group 1, 3, and 4, and the power dissipation is almost halved while all specifications are met when compared with [2].

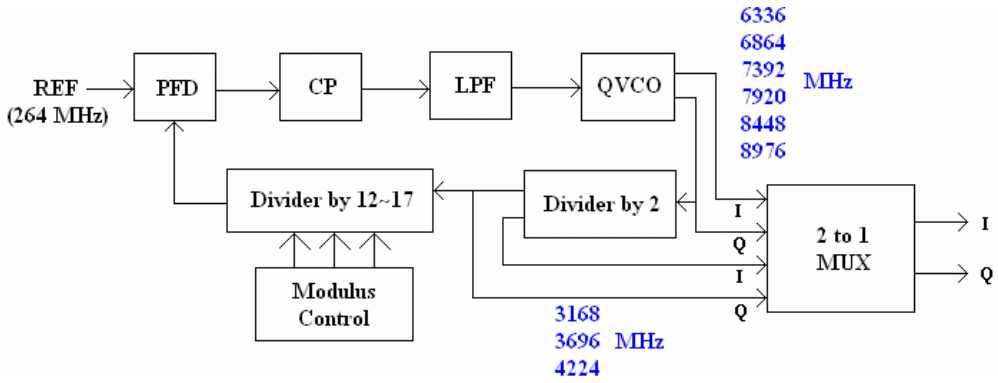


Figure 1: Proposed frequency synthesizer in this paper.

2. ARCHITECTURE

According to Figure 1, the proposed frequency synthesizer is based on an integer-N type phase-locked loop. It is composed of a quadrature voltage-controlled oscillator (QVCO), a multi-modulus frequency divider, a fast phase-frequency detector, a charge pump with variable current, and an on-chip third-order passive loop filter. There are seven digital input signals: four are to select the tuning range curves of the QVCO and the remainders are to control the division ratio in the multi-modulus divider.

2.1. Quadrature Voltage-Controlled Oscillator

As shown in Figure 2, the QVCO is in a complementary cross-coupled negative-gm configuration. The tail current source is removed to maximize the output swing. Two benefits also achieved due to the removal of the current source. First, the current source is the main contributor to the phase noise [3]. Second, when all transistors in the VCO core are put in GHz-switching bias condition, flicker noise will apparently be reduced by about 10 dB [4]. For a wide tuning range of 6∼9 GHz, the SCA (switched-capacitor-array) is used. SCA is composed of four pairs of binary-weighted MIM

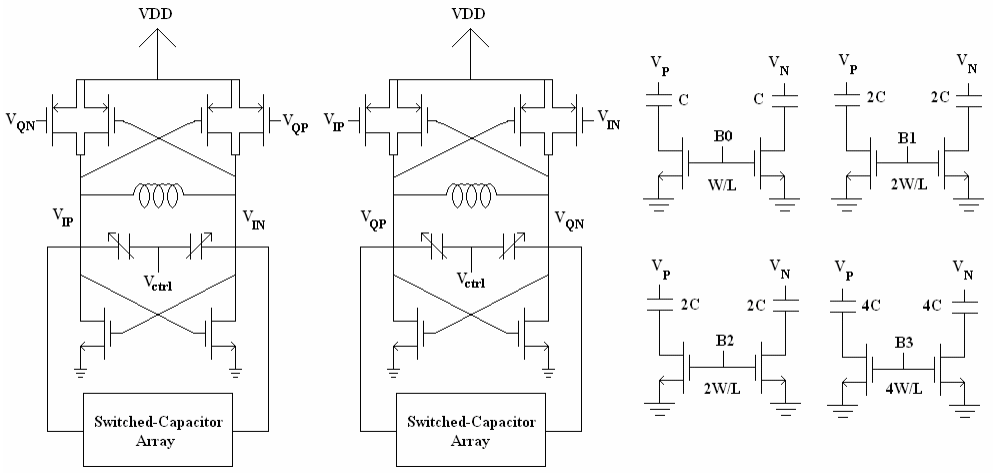


Figure 2: Schematic of the LC-QVCO and the switched-capacitor array.

capacitors and eight NMOS transistors as digital-control switches. The SCA decides the tuning range curve and then the varactors are for fine tuning. Therefore, no bulky varactors are required because the whole bandwidth is not covered only by the varacters.

2.2. Multi-modulus Frequency Divider

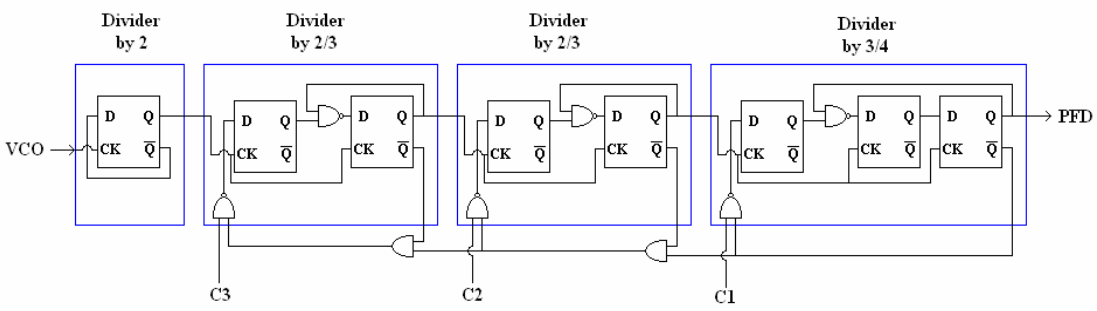


Figure 3: Schematic of the programmable divider.

For the wideband locking-range and high reference frequency consideration, current-mode logic (CML) is adopted in the whole programmable frequency divider. As shown in Figure 3, the multi-modulus frequency divider is implemented by cascaded a divider-by-2 circuit and three dual-modulus asynchronous frequency dividers. This design assures only the first divider works at the highest frequency and no pulse swallow counter or phase select state machine is needed. Moreover, the modulus-control signals of the last stage are produced first and given to the followed stage. Thus the delay in the critical path (the feedback of the first divider) is minimized [5]. In order to integrate the divider-by-2 into the loop, the division ratios are all even. The output frequency can be expressed by the following equation.

$$f_{\text{out}} = \frac{1}{2 \cdot (3 \cdot 2^2 + C1 \cdot 2^2 + C2 \cdot 2 + C3)} f_{\text{in}} \tag{1}$$

2.3. Phase-frequency Detector

The phase-frequency detector compares two inputs from the reference clock and the output at the last stage of the frequency divider. A conventional tri-state PFD is widely used for the wide comparable range of almost $\pm 2\pi$ radians. Due to the delay buffer in the reset path of the PFD which avoids the deadzone problem, the actually valid phase comparison range shrinks to $\pm|2\pi - \Delta|$. Δ can be found out by

$$\Delta = 2\pi \cdot t_{\text{delay}} \cdot f_{\text{REF}} \tag{2}$$

t_{delay} is the delay time in the rest path and f_{REF} is the reference frequency [6]. Since the reference is 264 MHz and Δ becomes considerable, the control signal will not monotonically approach to lock-in range.

In order to solve such a problem, the precharged PFD in [2] is designed and shown in Figure 4. When the phase error is close to $\pm 2\pi$, the next leading signal is coming before the reset signal triggers the DFF and the next leading will be ignored in the conventional PFD. But with the inserted delay in the precharged PFD, the falling edge of the reset signal can produce the correct comparison. Therefore no signal is lost.

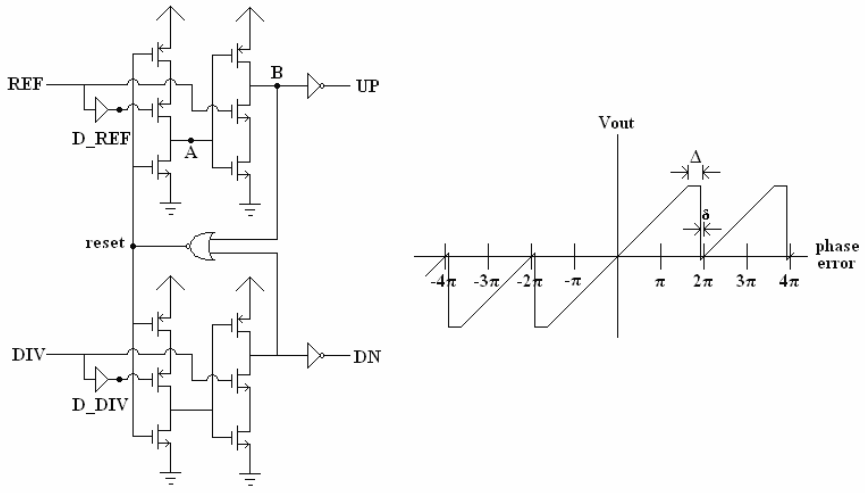


Figure 4: Precharge PFD.

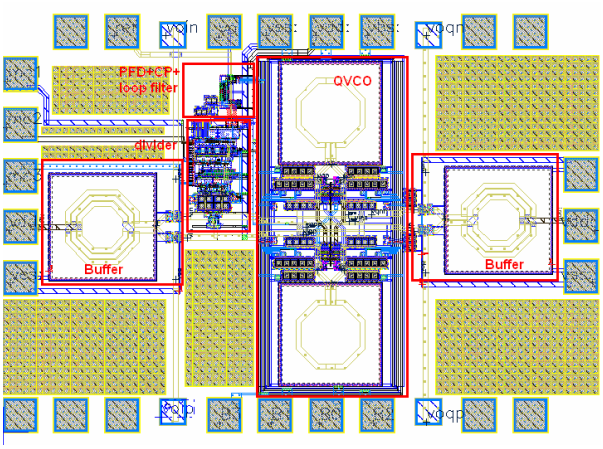


Figure 5: Circuit layout.

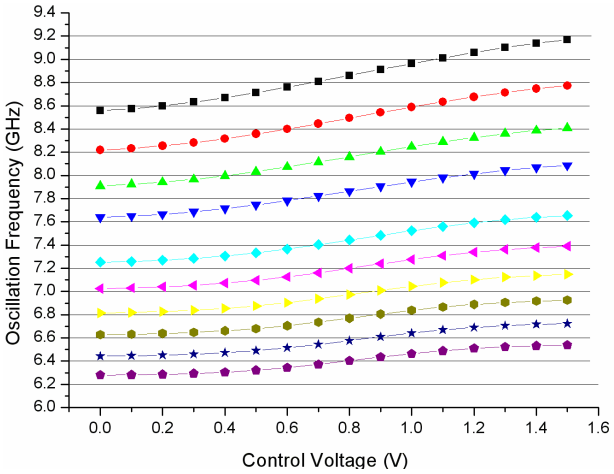


Figure 6: Tuning range curves with different banks.

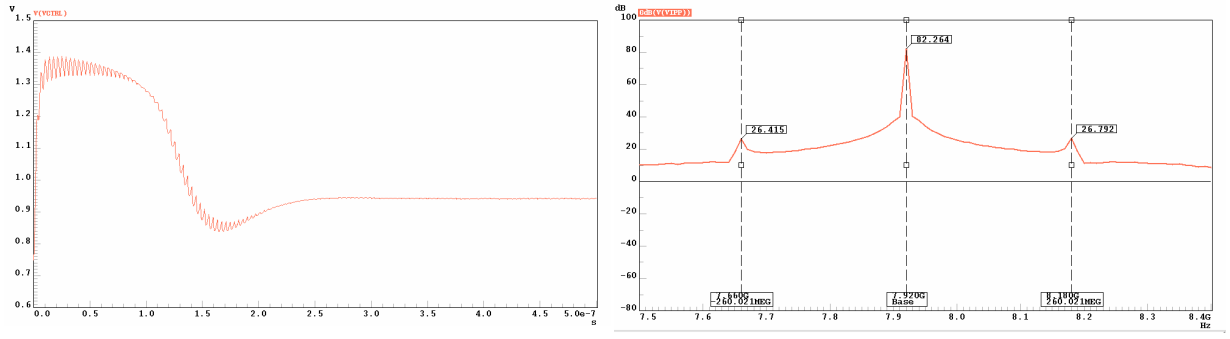


Figure 7: Transient waveform and spectrum at 7.92 GHz.

3. SIMULATION RESULTS

Figure 5 shows the layout of this circuit and the size is $1.75 \times 1.23\,\text{mm}^2$. According to Figure 6, the 10 tuning range simulated curves cover each other and signal from 6.12∼9.15 GHz can all be generated. This range includes the whole required bandwidth. Following the tuning range, the simulated transient waveform and spectrum at 7.92 GHz are shown in Figure 7. The performance of all carrier frequencies is listed in Table 1. The locking time is less than 300 nsec and the phase noise is lower than −109.8 dBc/Hz at 1 MHz offset. The spurious tone is smaller than −34.2 dBc at 264 MHz offset while it actually has no effect on the channel at all. Moreover, the simulated power-frequency-normalized figure-of-merit (FOM) is higher than 178.6. The total power dissipation is only 46.35 mW.

Table 1: Summary of output power and phase noise performance at six frequencies.

Carrier Frequency	**Output Power**	**Phase Noise @ 1 MHz offset**	**Spurious Tone**	**FOM**
6.336 GHz	1.08 dBm	−110.50 dBc/Hz	−34.5 dBc	178.6
6.864 GHz	2.17 dBm	−109.83 dBc/Hz	−34.2 dBc	178.6
7.392 GHz	4.10 dBm	−110.62 dBc/Hz	−37.9 dBc	180.1
7.920 GHz	5.44 dBm	−113.63 dBc/Hz	−55.5 dBc	183.7
8.448 GHz	6.41 dBm	−111.11 dBc/Hz	−40.5 dBc	181.7
8.976 GHz	6.86 dBm	−110.34 dBc/Hz	−35.6 dBc	181.5

Table 2: Comparison with the reference paper.

	JSSC, Aug. 2005 [2]	**This work**
Technology	0.18 μm CMOS	0.18 μm CMOS
Voltage Supply	1.8 V	1.5 V
Reference	528 MHz	264 MHz
Frequency Tuning Range	6.17∼9.11 GHz	6.28∼9.17 GHz
Average Phase Noise (dBc/Hz)	−109.48 @ 1 MHz	−111.01 @ 1 MHz
FOM	178.53	180.7
Sidebands	$<$ −52 dBc @ 528 MHz	$<$ −34.2 dBc @ 264 MHz
Settling Time	$<$ 150 ns	$<$ 300 ns
Power Dissipation		
QVCO	14.4 mW	12.46 mW
Divider	31.5 mW	21.42 mW
PFD + CP	5.04 mW	4.85 mW
Buffers	34.2 mW	7.62 mW
Total	85.14 mW	46.35 mW

4. CONCLUSION

A fast-locking and low power frequency synthesizer for MB-OFDM UWB has been designed using TSMC 0.18 μm RF CMOS 1P6M process. The tuning range is about 2.9 GHz and no gap occurs between each band. Therefore, it can generate six carrier frequencies from 6.336 to 8.976 GHz in a step of 528 MHz. The phase noise is $-109.8 \sim -113.6$ dBc/Hz at 1 MHz offset. The total power dissipation is 46.35 mW from 1.5 V power supply. The performance of this work is listed and compared with Ref. [6] in Table 2, the power dissipation is only 54.4% of the synthesizer in [2] and all specifications are met.

REFERENCES

1. Batra, A., et al., "Multi-band OFDM physical layer proposal for IEEE 802.15 Task Group 3a," *IEEE*, Piscataway, NJ, IEEE P802.15-03/268r3-TG3a, Mar. 2004.
2. Geum-Young, T., et al., "A 6.3-9-GHz CMOS fast settling PLL for MB-OFDM UWB applications," *IEEE J. Solid-state Circuits*, Vol. 40, 1671–1679, Aug. 2005.
3. Rael, J. J. and A. A. Abidi, "Physical processes of phase noise in differential LC-oscillators," *IEEE Custom Integrated Circuits Conf.*, 569–572, May 2000.
4. Klumperink, E. A. M., et al., "Reducing MOSFET 1/f noise and power consumption by switched biasing," *IEEE J. Solid-state Circuits*, Vol. 35, 994–1001, July 2000.
5. Theil, D., et al., "A fully integrated CMOS frequency synthesizer for bluetooth," *IEEE Radio Frequency Integrated Circuits Symposium*, 103–106, May 2001.
6. Mansuri, M., D. Liu, and C.-K. K. Yang, "Fast frequency acquisition phase-frequency detectors for GSamples/s phase-locked loops," *IEEE J. Solid-state Circuits*, Vol. 37, 1331–1334, Oct. 2002.
7. Best, R., *Phase-locked Loops: Design, Simulation, and Applications*, Mcgraw-Hill Companies, 2003.

A 0.7 V Transformer-feedback CMOS Low-noise Amplifier for 5-GHz Wireless LAN

H. I. Wu, **R. S. Fan**, **and C. F. Jou**
Institute of Communication Engineering, Chiao Tung University, Hsinchu, Taiwan, China

Abstract— In this paper, we present a 0.7 V low-noise amplifier(LNA) using monolithic transformer feedback to neutralize the gate-drain overlap capacitance of a field-effect transistor(FET). It is a single-ended amplifier implemented in 0.18-μm CMOS technology designing for 5-GHz wireless local-area networks (LANs). This LNA achieves a simulated power gain of 10.84 dB, noise figure(NF) of 2.4 dB, and input referred 1 dB compression point (P1 dB) of −8 dBm at 5.8 GHz. Operating from a 0.7-V supply, the power consumptions for the low noise amplifier(LNA) are 4.8 mW.

1. INTRODUCTION

The market for the modern wireless communication systems demands low-noise amplifiers (LNAs) operating at low dc bias conditions. The use of supply voltage at or below 1 V becomes even more important when considering analog/RF and digital circuitry on the same die, from both cost and packing considerations. The noise figure (NF) should not exceed 3 dB when it has a gain more than 10 dB. [1]. The most popular topology of LNA is cascode LNA. It can obtain high gain, low noise and sufficient isolation between input and output, but this topology is not suitable for low voltage supply. Some paper published so far represents new topology for low-voltage application [2, 3]. In this paper, we use one Mos to be suitable for low voltage application, but it can not maintain good isolation from input and output. Because as operating frequencies increase, the field-effect transistor (FET) gate-drain overlap capacitance C_{gd} plays an more important role in performance. Since its magnitude is comparable to the gate-source capacitance in deep-submicron CMOS technology. In this paper we use the inductive-feedback to cancel the signal from output to input through C_{gd}. Since the effect of the Miller capacitor C_{gd} is canceled, the isolation between input and output can be improved. Because we did not use the conventional cascode LNA, the power supply V_{dd} can be lower to be suitable for low voltage application.

2. CIRCUIT IMPLEMENTATION

In this paper we use inductive-feedback, which introduces magnetic coupling between source and drain inductor, as shown in Fig. 1. Because feeding back a portion of output signal can cancel the signal feedback from output to input through C_{gd} the isolation between input and output can be improved when only single active device (i.e., MOS) is used. The transformer h parameters are given by $h_{11} = s(1-k^2)L_{11} \approx 0$, $h_{12} = \frac{-k}{n} \approx \frac{-1}{n}$, $h_{12} = \frac{k}{n} \approx \frac{1}{n}$, $h_{22} = \frac{1}{sL_{22}} \approx 0$. K represents the transformer coupling coefficient $k = M/(L_{11}L_{22})$ and the transformer turn ratio is given by $n = \sqrt{L_{22}/L_{11}}$. The small signal equivalent circuit is shown in Fig. 2. The input impedance, derived from the small signal circuit, $Z_{in} = Z_{gs} + \frac{\beta(s)(Z_L \| Z_{gd})}{n} + \frac{(\beta(s)+1)(Z_L \| Z_{gd})}{n^2}$ [4]. The input impedance is different from the conventional cascode LNA with source degeneration inductor, as shown in Fig. 3. Its input impedance is $Z_{in} = sL_s + \frac{1}{sC_{gs}} + \frac{g_m L_s}{C_{gs}}$. The conventional topology uses the inductor L_g to obtain noise and input conjugate match simultaneously [5]. The impedance of this design shows that we can not use a L_g to achieve noise and input conjugate match simultaneously, so we use the series-L shunt-C network to obtain balance between noise and input matching. In this design we use the inductive-feedback to cancel the signal feedback from output to input through the Miller capacitor C_{gd}. Since the effect of the Miller capacitor C_{gd} is canceled, the isolation between input and output can be improved. To achieve the neutralization, we design $\frac{n}{k} \approx \frac{C_{gs}}{C_{gd}}$ and its value equals to about 3 for the 0.18 um CMOS technology [4]. In order to lower the noise of the circuit, we choose the channel length of MOS $W_{opt} \approx \frac{1}{3\omega C_{ox} R_s L}$ [6], and the source inductor L_s to make sure that the circuit is stable after the inductive-feedback. The series gate inductor L_g and shunt C_g network is for the input impedance and it would not deteriorate the noise performance too much. The L_d and coupling coefficient k are designed properly to achieve $\frac{n}{k} \approx 3$. The shunt capacitor with drain inductor L_d is to be resonant at operating frequency to achieve output matching.

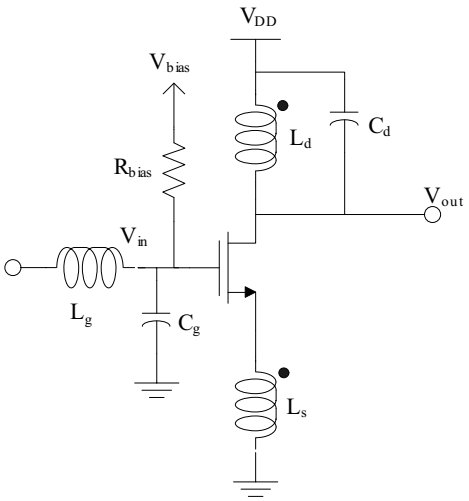


Figure 1: Transformer-feedback single ended LNA.

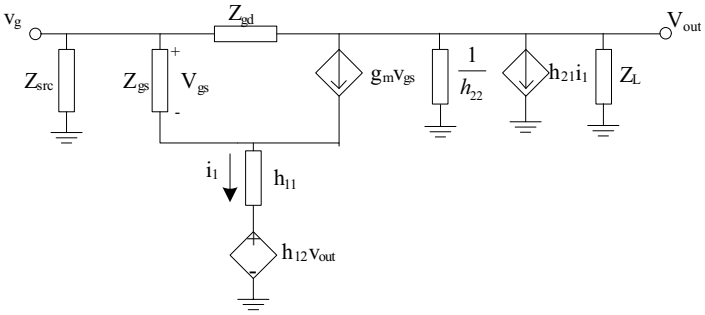


Figure 2: Small-signal equivalent circuit of transformer-feedback LNA [4].

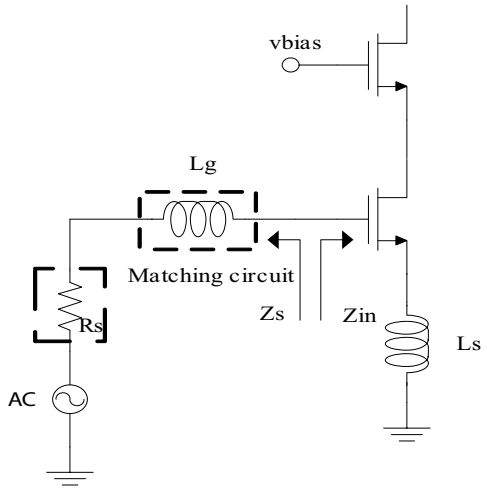


Figure 3: Schematic of a cascade LNA topology [5].

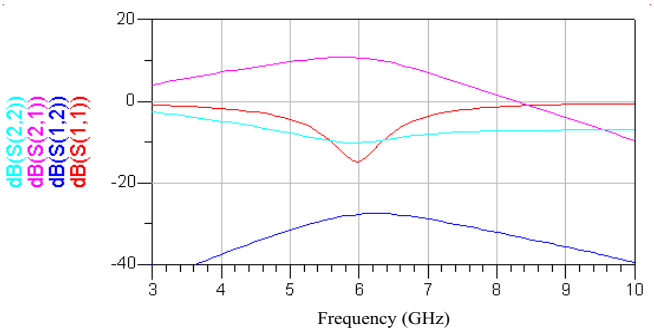


Figure 4: S parameter.

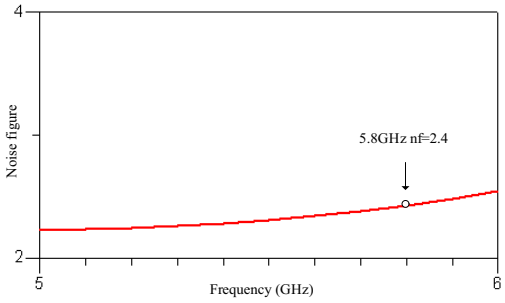


Figure 5: Noise figure.

3. SIMULATION FIGURE AND CONCLUSION

In this design, ADS is used to simulate circuit and Sonnet is used to simulate the transformer. The simulation result shows that the input and output return loss S11 is 13.2 dB and S22 is 10.1 dB at 5.8 GHz. The isolation S21 is 27.684 dB, which shows that the inductive-feedback can improve the isolation effectively, as shown in Fig. 4 and the noise figure is 2.4 dB as shown in Fig. 5. The P1 dB of this design is -8.6 dBm, as shown in Fig. 6. Total consumed power is 6.86 mA$\times$0.7 V$=$ 4.8 mW. In this design we use the inductive feedback to improve the isolation between input and output and we did not use the conventional cascode LNA, and the power supply V_{dd} can be lower to achieve low voltage application.

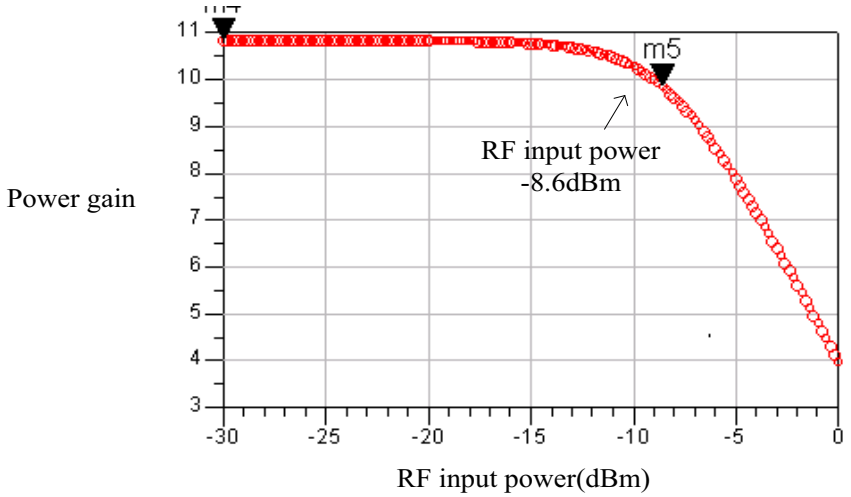


Figure 6: P1 dB.

REFERENCES

1. Shaeffer, D. K. and T. H. Lee, "A 1.5 V, 1.5 GHz CMOS low noise amplifier," *IEEE Journal of Solid-state Circuit*, Vol. 32, No. 5, May 1997.
2. Manku, T., et al., "A low-voltage design technique for RF integrated circuit," *IEEE Transaction on Circuits and Systems_II: Analog and Digital Signal Processing*, Vol. 45, No. 10, October 1998.
3. Wang, Y., M. Z. Khan, and K. Iniewski, "A 0.65 V, 1.9 mW CMOS low-noise amplifier at 5 GHz," *Proceedings of the 9th International Database Engineering & Application Symposium*, 2005.
4. Cassan, D. J. and J. R. Long, "A 1-V transformer-feedback low-noise amplifier for 5-GHz wireless LAN in 0.18-µm CMOS," *IEEE Journal of Solid-state Circuits*, Vol. 38, No. 3, March 2003.
5. Nguyen, T.-K., et al., "CMOS low-noise amplifier design optimization techniques," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 52, No. 5, May 2004.
6. Lee, T. H., "The design of CMOS radio-frequency integrated circuits," Cambridge University Press, Cambridge, UK, 1998.

A 0.75 V CMOS Low-noise Amplifier for Ultra Wide-band Wireless Receiver

Hui-I. Wu, **Zi Hao Hsiung**, **and Christina F. Jou**

Microwave and Millimeter wave Laboratory, Institute of Communication Engineering
Chiao Tung University, Hsinchu, Taiwan, China

Abstract— An ultra-wideband (UWB) CMOS low-noise amplifier(LNA) topology that combines a common-gate stage for input wideband matching with a shunt-peaked folded-cascode configuration for wideband amplifying stage is presented in this paper. The proposed UWB LNA achieves simulated power gain > 10 dB from 3.3 to 10.1 GHz with only 0.75 V supply using 0.18 µm CMOS process. Its broadband matching is better than -10 dB for S11 and S22 from 3.1 to 10.6 GHz, and an average noise figure is about 4 dB, while consuming 11.5 mw with output buffer amplifier.

1. INTRODUCTION

A low-noise amplifier (LNA) is the first stage, after antenna and low-pass filter in the receiver block of a communication system. It is widely used in the front-end of narrow band communication system. For UWB applications, the criteria to judge its performances are slightly different. The UWB LNA face several challenges, such as broadband input matching, sufficient power gain through entire bandwidth to suppress the noise of next stage, low noise figure and low power consumption. There are several existing structures for wideband amplifiers. The well-known distributed amplifiers [4] can achieve high gain through extremely broad bandwidth, but it consumes large power and area, because it cascades several stages and uses many inductors in circuit. Many circuits use the 3rd-order Chebyshev filter [2, 3] to achieve input broadband matching. Although it can achieve good return loss, it still consumes large area because of using several inductors. Among the demonstrated UWB LNA in the published papers [3], the lowest supply voltage is 1-V. Here, we propose an UWB LNA design (shown in Fig. 1) in which a common-gate and a folded-cascode architecture are employed to achieve good input matching and enough gain from 3.1 to 10.6 GHz with only 0.75 V supply.

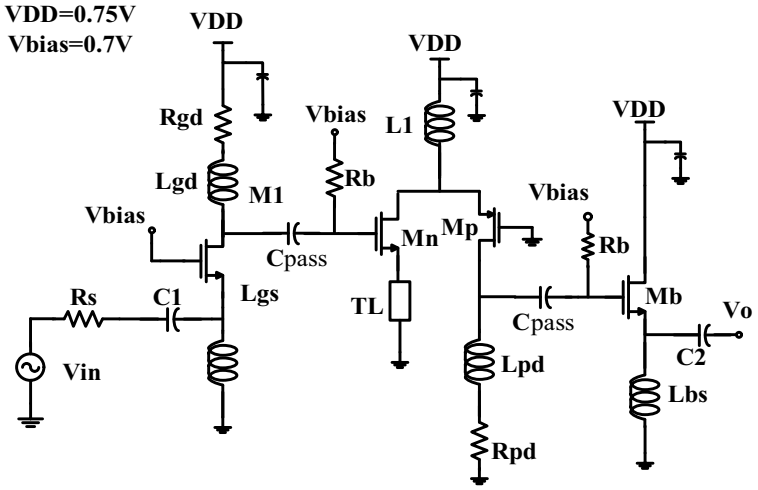


Figure 1: The proposed low-V low-noise amplifier.

2. CIRCUIT DESIGN

The LNA circuit can be divided into three block — input matching stage, amplifying stage, and output buffer stage. The whole circuit of UWB LNA is shown in Fig. 1, and the supply voltage is only 0.75 V. The DC blocking capacitors are C1 and C2. For input stage, a common-gate (M1) is used to match to 50 Ω, its small signal equivalent circuit is shown in Fig. 2. C_{gs} and C_{gd}

respectively are the gate-to-source and the gate-to-drain parasitic capacitances; r_o is the channel length modulation resistor of MOS transistor. The input impendance of the common-gate can be derived as [1]

$$Z_{in} = \frac{1}{g_m + \dfrac{1}{Z_s(\omega)} + \dfrac{1 - g_{m1} Z_o(\omega)}{r_o + Z_o(\omega)}} \tag{1}$$

where

$$Z_s(\omega) = j\omega L_s // \frac{1}{j\omega C_{gs}} \tag{2}$$

$$Z_o(\omega) = \frac{1}{j\omega C_{gd}} // Z_L // Z_{in2} \tag{3}$$

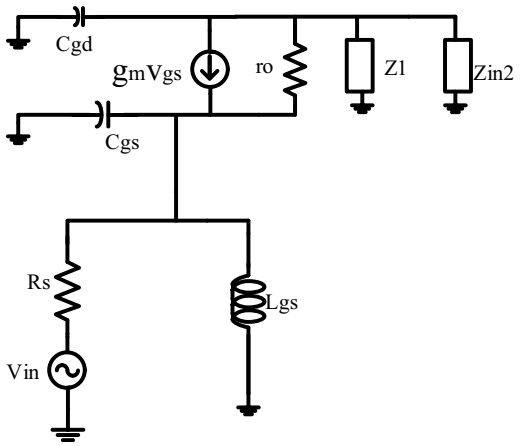


Figure 2: Common-gate small signal equivalent circuit.

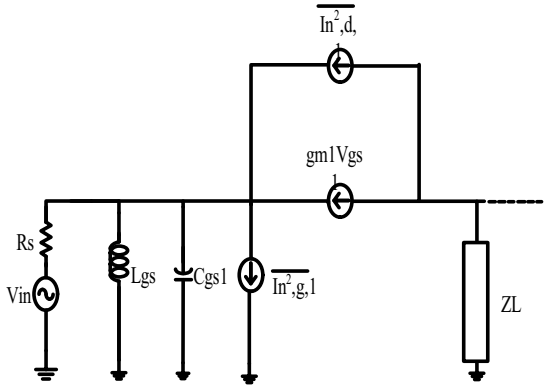


Figure 3: Common-gate noise model.

L_s and C_{gs} should resonate at the center of operating band, so that the second term disappear at the center frequency. The third term in denominator is induced by channel length modulation of a MOS. In order to decrease the effect of ro on input impendance, the transconductance of M1 is set to be slightly greater than 20 ms [1]. The first stage noise model is shown in Fig. 3, where $i_{n,d}$ and $i_{n,g}$ are respectively channel thermal noise and induced gate noise. The noise factor of common-gate stage can be derived as [1].

$$F{=}1{+}\frac{\alpha\delta\left(1{-}|C|^2\right)\omega^2 C_{gs_1} R_s}{5g_{m1}} + \frac{\gamma}{\alpha R_s g_{m1}}\left(1{+}\frac{R_s^2}{|Z_s(\omega)|^2}\right){+}\frac{\alpha\delta|c|^2\omega^2 C_{gs1}^2 R_s}{5g_{m1}} + \frac{2|c|\sqrt{\dfrac{\delta\gamma}{5}}\omega C_{gs1} R_s}{j g_{m1} Z_s(\omega)} \tag{4}$$

In Eq. (4), because $Z_s(\omega)$ is in the denominator, L_s and C_{gs} should resonate at the center of operating band to obtain minimal noise factor. The transconductance (g_{m1}) of M1 should be increased to decrease noise factor. But the input matching will be worse than -10 dB when g_{m1} increases to 35 mS. For a common-gate circuit, there is trade-off between input matching and noise factor. For the proposed circuit, g_{m1} is set about 33 ms to decrease noise factor, but input return loss still better than -10 dB. For the amplifying stage, a folded-cascode with shunt-peaked topology [8], (Mn, Mp, L1, Lpd and Rpd) is implemented for reducing Miller effects, achieving higher gain and better isolation between output and input port. The inductor peaking impedance increases with the increasing of the operating frequency, and load resistor rpd is added to increase the low frequency gain, so this topology can achieve enough power gain through entire bandwidth by only 0.75 V bias. TL is a transmission line which can stabilize folded-cascode circuit. rgd can increase the flatness of the power gain.

3. SIMULATION RESULTS AND LAYOUT CONSIDERTIONS

This work is simulated by using TSMC 0.18 μm mixed-signal/RF CMOS 1P6M technology under typical-typical corner with a low supply voltage of 0.75 V. The post-layout circuit simulation is

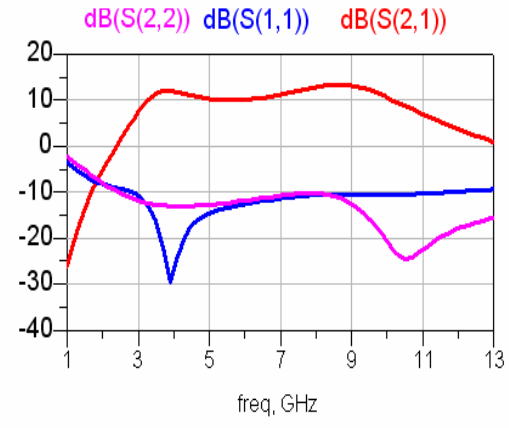


Figure 4: S parameter.

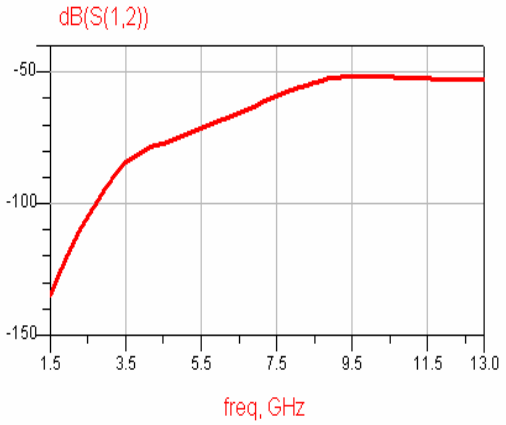


Figure 5: Reverse isolation.

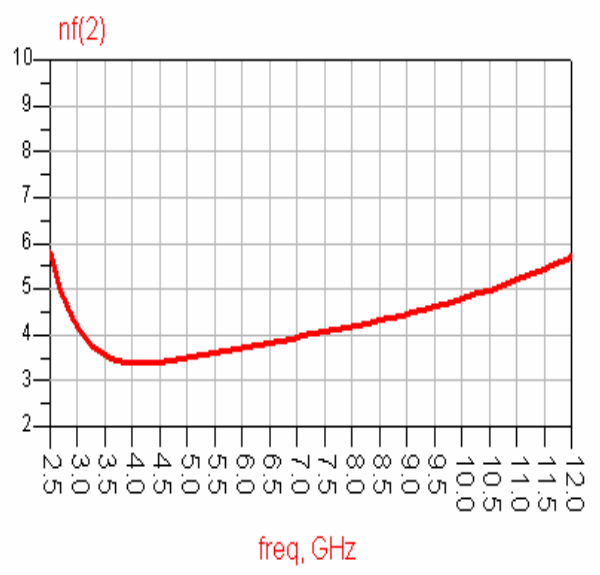


Figure 6: Noise figure.

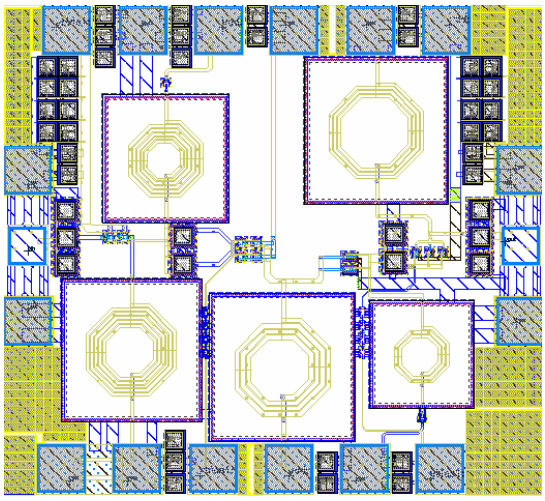


Figure 7: Layout of the proposed circuit.

performed using Agilent ADS and Sonnet to include parasitic effects of the actual devices in the CMOS process. The proposed LNA consumes 11.5 mw with 0.75 V supply voltage. The total width of M1 is set to be 100 µm, and Lgs is chosen to be 7.5 nH to optimize simultaneously input return loss and noise factor. The S-parameter is shown in Fig. 4 and Fig. 5. The simulated S11 and S22 are better than −10 dB over 3.1–10.6 GHz, and power gain (S21) is $>$ 10 dB through 3.3–10.5 Ghz. The simulated minimum noise figure is 3.4 at 4.1 GHz, and noise figure is from 3.4 to 4.75 through 3.1–10.1 GHz, as shown in Fig. 6. Discussion about the topic of nonlinearity, one-tone test of 3.1 GHz, 6.5 GHz and 10.1 GHz is performed to obtain the input referred 1dB compression point. The simulated P1db are −19.4 dBm, −13.8 dBm, −12.4 dBm at 3.1 GHz, 6.5 GHz and 10.1 GHz.

	Tech.	BW (GHz)	S11 (dB)	S22 (db)	Power Gain (dB)	NFmin (dB)	VDD (V)	Power (mW)
This work	**0.18 µmCMOS**	**3.3–10.1**	**<-10**	**<-10**	**10.1–13.1**	**3.4**	**0.75**	**11.5**
[1]	0.18 µm CMOS	3.1–10.6	<-9	<-13.0	15.9–17.5	3.1	1.8	33.2
[2]	0.18 µm CMOS	2.4–9.5	<-9.4	<-8.0	7.4–10.4	4.2	1.8	18
[3]	0.18 µm CMOS	3–8.2	<-10.0	<-10.0	10	3.68	1	16.1
[7]	0.18 µm CMOS	0.2–22	<-8	N/A	8.1	4.3	1.3	52
[8]	0.18 µm CMOS	3–10	<-7.5	<-11	11.3–13.1	2.8	1.5	27

The layout of the proposed UWB LNA is presented in Fig. 7. The die area including the pads is 0.97 mm×1.07 mm.

4. CONCLUSIONS

A CMOS UWB LNA is designed with a common-gate stage and folded cascade shunt-peaked amplifier. The simulation results show that the proposed LNA can achieve enough gain ,low noise and good matching from 3.1 GHz to 10.6 GHz while consuming 11.5 mW with only 0.75 V supply. For post-simulation the transmission line high frequency effect is considered. Since the design does not require any off-chip components, it can be easily integrated as part of a complete low-voltage transceiver.

REFERENCES

1. Lu, Y., K. S. Yeo, A. Cabuk, J. Ma, M. A. Do, and Z. Lu, "A novel CMOS low-noise amplifier design for 3.1-to 10.6-GHz ultra-wide-band wireless receivers," *IEEE Transactions on Circuits and Systems-I: Regular Papers*, Vol. 53, No. 8, August 2006.
2. Bevilacqua, A. and A. M. Niknejad, "An ultra-wide-band CMOS lownoise amplifier for 3.1–10.6-GHz wireless receivers," *IEEE J. Solid-State Circuits*, Vol. 39, No. 12, 2259–2268, Dec. 2004.
3. Chang, B.-Y. and C. F. Jou, "Design of a 3.1–10.6 GHz low-voltage, low-power CMOS low-noise amplifier forultra-wideband receivers," *IEEE APMC2005 Proceedings*, 2005.
4. Liu, R.-C., K.-L. Deng, and H. Wang, "A 0.6–22 GHz broadband CMOS distributed amplifier," *IEEERadio Frequency Integrated Circuits Symp. Dig. Papers*, 103–106, 2002.
5. Chen, S.-C., R.-L. Wang, M.-L. Kung, and H.-C. Hung, "An integrated CMOS low noise amplifier for ultra wide band applications," *International Conference on Wireless Networks, Communications and Mobile Computing*, 2005.

Adaptive Selection of Sampling Interval in Frequency Domain for Estimating the Poles of Scatterers

Shaogang Wang, **Xinpu Guan**, **Dangwei Wang**, **Xingyi Ma**, **and Yi Su**
School of Electronic Science and Technology
National University of Defense Technology, Changsha 410073, China

Abstract— In this paper, the matrix pencil (MP) method has been utilized for estimating the complex natural resonances from IFFT of calculated frequency responses. The novelty is that frequency sampling interval is investigated. The responses are usually computed by the accurate method of moments (MoM) in frequency domain, which is time-consuming. So, the frequency sampling interval's effect on results should be analyzed, and a scheme for adaptive choice of frequency sampling interval is proposed in this paper. The scheme is validated by a special example, a perfect conducting sphere. Lastly, poles of a complex radar target, a missile model, are calculated by the proposed method and compared with other available data.

1. INTRODUCTION

It is known that extracting complex natural resonance frequencies (poles) of the target from its transient response is desired in target identification. The target poles contain the information of decay factors and the resonant frequencies that are only determined by target's structure and material. So, they should be estimated with high accuracy. Of many methods, the matrix pencil (MP) is an effective one for extracting complex natural resonance frequencies of scatterers from transient responses [1–3].

A method in time domain, finite difference time domain (FDTD) or MoM in time domain (TD-MoM), can be used to calculate transient responses, from late time part of which poles are estimated. But that poles are not accurate as those extracted from impulse responses, which can be obtained by the deconvolution technology in time domain. Noteworthily, noise will be inflicted in deconvolution.

The method of moments (MoM) in frequency domain is widely used, accurate and chosen to calculate the frequency response in this paper. The impulse response can be obtained by using Inverse Fast Fourier Transform (IFFT) technique from the frequency response.

As all known, frequency sampling interval is less and more accurate results are obtained. On the other way, MoM is time-consuming. The sampling interval must be investigated for results with enough accuracy in a short period of time. An algorithm of adaptive choice of frequency sampling interval is proposed in this paper.

The paper is organized as follows. In the next section, the mathematical formulation of MP is briefly presented. In Section 3, a scheme for adaptive choice of frequency sampling interval is proposed and validated. Finally, in Section 4, the proposed scheme is validated through a perfect conducting sphere and an application is presented. Numerical results of a complex radar target, a missile model, are computed for poles and compared with available data.

2. FORMULATIONS OF THE MP METHOD

If the frequency response calculated by MoM is $[\mathbf{Y}(f_i)]$ $(i = 0, 1, \ldots, N' - 1)$. The sampling interval is f_s. Then, The impulse response, $[\mathbf{y}(kT_s)]$ $(k = 0, 1, \ldots)$, can be obtained through IFFT. T_s is the sampling period, and $T_s = 1/(2N'f_s)$. By the singularity expansion method (SEM), the time sequence can be rewritten as

$$y(kT_s) \approx \sum_{i=1}^{M} R_i z_i^k \quad \text{for} \quad k = 0, 1, \ldots, N-1 \tag{1}$$

and

$$z_i = e^{s_i T_s} = e^{(-\alpha_i + j\omega_i)T_s} \quad \text{for} \quad i = 1, 2, \ldots, M \tag{2}$$

where, s_i and R_i are poles and residues. N is the length of impulse response. M is the number of poles.

Two matrices $[\mathbf{Y_1}]$ and $[\mathbf{Y_2}]$ are considered, which defined as

$$[\mathbf{Y_1}] = \begin{bmatrix} y(0) & y(1) & \cdots & y(L-1) \\ y(1) & y(2) & \cdots & y(L) \\ \vdots & \vdots & & \vdots \\ y(N-L-1) & y(N-L) & \cdots & y(N-2) \end{bmatrix}_{(N-L)\times L} \tag{3}$$

$$[\mathbf{Y_2}] = \begin{bmatrix} y(1) & y(2) & \cdots & y(L) \\ y(2) & y(3) & \cdots & y(L+1) \\ \vdots & \vdots & & \vdots \\ y(N-L) & y(N-L+1) & \cdots & y(N-1) \end{bmatrix}_{(N-L)\times L} \tag{4}$$

The pencil parameter, L, is very useful in eliminating some effects of noise in the data.

The parameters z_i may be found as the generalized eigenvalues of the matrix pair $\{[\mathbf{Y}_2]; [\mathbf{Y}_1]\}$. Equivalently, the problem of solving for z_i can be cast as an ordinary eigenvalue problem,

$$\left\{[\mathbf{Y_1}]^+[\mathbf{Y_2}] - \lambda[\mathbf{I}]\right\} \tag{5}$$

where $[\mathbf{I}]$ is identify matrix, and $[\mathbf{Y_1}]^+$ is the Moore-Penrose pseudoinverse of $[\mathbf{Y_1}]$.

The residues, R_i, are solved for from the following least-squares problem:

$$\begin{bmatrix} y(0) \\ y(1) \\ \vdots \\ y(N-1) \end{bmatrix} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ Z_1 & Z_2 & \cdots & Z_M \\ \vdots & \vdots & & \vdots \\ Z_1^{N-1} & Z_2^{N-1} & . & Z_M^{N-1} \end{bmatrix} \begin{bmatrix} R_1 \\ R_2 \\ \vdots \\ R_M \end{bmatrix} \tag{6}$$

3. ADAPTIVE CHOICE OF FREQUENCY SAMPLING INTERVAL

It is time-consuming to calculate frequency responses by MoM, so fewer frequency sampling points are expected. On the other way, it needs more frequency sampling points to obtain more accurate results. A scheme of adaptive choice of frequency sampling interval is proposed in this part.

Although a scatterer has many poles, only several of them (about five at most) can be numerically obtained and useful to characterize the target [4]. It is important to accurately estimate those dominant poles. So, an average relative error, err, is given in advance and defined as

$$err = \frac{\sum_{i=1}^{M'} |s_i - s_i'|}{\sum_{i=1}^{M'} |s_i|} \tag{7}$$

where M' is the number of dominant poles. $s_i^{'}$ are estimated poles from frequency response at an interval of f_s, and s_i are those at an interval of $0.5f_s$. If the value of (7) is less than the given reference error ε, s_i are the expected results. Otherwise, the frequency interval is chosen as half of that last time, and poles are calculated again until the average relative error is less than ε.

The procedure of the algorithm is following:

(1) Give expected error ε, frequency step f_s, the number of frequency samples N'.
(2) Calculate frequency response, poles $s_i^{'}$ and estimate the number of dominant poles M'.
(3) Let $f_s = 0.5f_s$, calculate frequency response, and poles s_i.
(4) Compute the relative error err.
(5) If $err > \varepsilon$, let $s'_i = s_i$, and go to (3).
(6) If $err < \varepsilon$, stop it.

4. NUMERICAL EXAMPLES

For the scheme validated, a conducting sphere, the radius of which is 0.06 m ($a = 0.06$), is specifically considered. The incident wave is plane wave of unit amplitude. A frequency range of 0 Hz (excluded) to 5.12GHz is chosen as the excitation frequencies. At the observation 1000 m far from the center

of the sphere, the backscattering response data, which can be calculated by MoM or Mie here, are used to extract poles.

In Table 1, poles are valued as the form of sa/c. $err1$ are the average relative errors between numerical results and theoretical values. $err2$ are the average relative errors between numerical results and those of last time. If given $\varepsilon = 0.01$, from Table 1 we can see that $f_s = 40\,\text{MHz}$ and only 128 frequency points are enough for accurate poles. The errors indicate the proposed scheme is correct.

Table 1: Poles and relative errors under different frequency steps.

	Theory [5]	$f_s = 160\,\text{MHz}$ N'=32	$f_s = 80\,\text{MHz}$ N'=64	$f_s = 40\,\text{MHz}$ N'=128	$f_s = 20\,\text{MHz}$ N'=256	$f_s = 10\,\text{MHz}$ N'=512
1	−0.500+j0.866	−0.500+j0.854	−0.496+j0.858	−0.498+j0.862	−0.500+j0.864	−0.500+j0.865
2	−0.702+j1.807	−0.668+j1.768	−0.700+j1.791	−0.702+j1.798	−0.700+j1.803	−0.702+j1.804
3	−0.843+j2.758	−0.846+j2.636	−0.836+j2.736	−0.839+j2.747	−0.843+j2.752	−0.844+j2.754
4	−0.954+j3.715	−1.096+j3.368	−0.948+j3.682	−0.951+j3.697	−0.953+j3.706	−0.954+j3.711
5	−1.048+j4.676	−1.003+j3.995	−1.047+j4.640	−1.049+j4.654	−1.048+j4.668	−1.048+j4.671
6	−1.129+j5.642	−0.900+j4.919	−1.122+j5.615	−1.109+j5.597	−1.114+j5.644	−1.123+j5.643
		$err1 = 0.1$	$err1 = 0.007$	$err1 = 0.006$	$err1 = 0.002$	$err1 = 0.001$
			$err2 = 0.09$	$err2 = 0.004$	$err2 = 0.004$	$err2 = 0.001$

As an application, poles of a common complex radar target are calculated in this part. A kind of missile model is considered as shown in Fig. 1(a) and Fig. 1(b). the unit of labels is meter. The main part of the model is a circle cylinder. The front part is half a spheroids, and the tail is half a sphere.

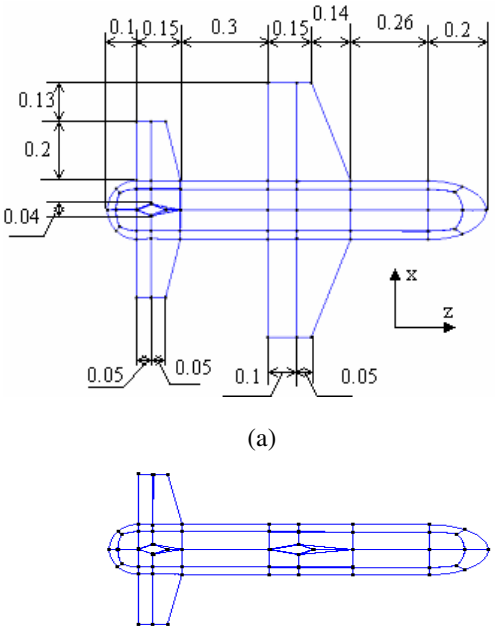


Figure 1: (a) The planform of the missile model, (b) The side view of the missile model.

With the incident plane wave of unit amplitude z-polarized and x-traveling, the frequency responses at the observation 1000 m far from the center of the half sphere are calculated by a fast MoM method [6] in the range from 0 Hz (excluded) to 307.2 MHz. That obtained frequency responses are shown in Fig. 2. The cut impulse with 100 points ($N = 100$) is used. Appling the MP method with $L = 49$, we only obtain two stable dominant poles shown in Table 2. The poles

Table 2: Poles and relative errors under different frequency steps of the missile.

	MoM	$f_s = 2.4\,\text{MHz}$ N'=128	$f_s = 1.2\,\text{MHz}$ N'=256	$f_s = 0.6\,\text{MHz}$ N'=512
1	−0.45+j2.37	−0.465+j2.318	−0.466+j2.355	−0.466+j2.373
2	−0.34+j4.30	−0.355+j4.256	−0.356+j4.296	−0.356+j4.317
		$err1 = 0.015$	$err1 = 0.0057$	$err1 = 0.006$
			$err2 = 0.01$	$err2 = 0.005$

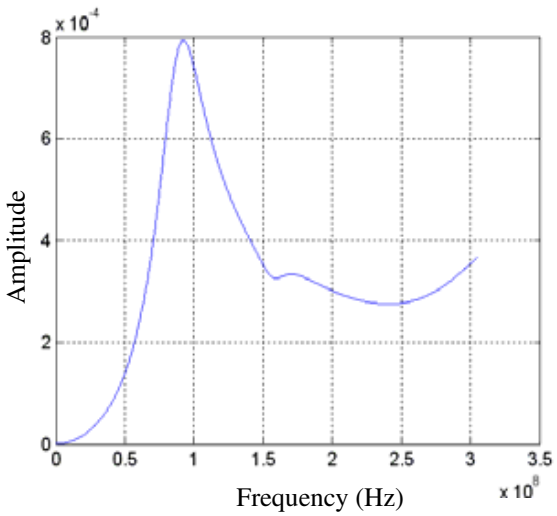


Figure 2: The frequency response of the model.

are valued in the form as sL'/c, where $L' = 1.3\,\text{m}$ is the length of the model. The compared data are obtained using the approach as in [7]. The required relative error is also 0.01 ($\varepsilon = 0.01$). The meanings of $err1$ and $err2$ are the same as those in Table 1.

The errors in Table 2 show that the poles we calculated are correct at least frequency sampling points, and 512 points is enough.

5. CONCLUSION

In application of the MP method for estimating the complex natural resonances of scatterers from frequency response data, a algorithm of adaptive choice of frequency sampling interval is proposed in this paper, because the responses are usually calculated by time-consuming MoM in frequency domain. Numerical examples show our method's effect on the results with required precision from frequency points at least.

REFERENCES

1. Hua, Y. and T. K. Sarkar, "Generalized pencil-of-function method for extracting poles of an EM system from its transient response," *IEEE Trans. Antennas Propagat.*, Vol. 37, No. 2, 229–234, 1989.
2. Sarkar, T. K. and O. Pereira, "Using the matrix pencil method to estimate the parameters of a sum of complex exponentials," *IEEE Antennas and Propagation Magazine*, Vol. 37, No. 1, 48–55, Feb. 1995.
3. Sarkar, T. K., S. Park, J. Koh, and S. M. Rao, "Application of the matrix pencil method for estimating the SEM (singularity expansion method) poles of source-free transient responses from multiple look directions," *IEEE Trans. Antennas Propagat*, Vol. 48, No. 4, 612–618, 2000.
4. Chauveau, J., N. de Beaucoudrey, and J. Saillard, "Characterization of radar targets in resonance domain with a reduced number of natural poles," *Radar Conference*, EURAD, 85–88, European, Oct. 6, 2005.

5. Chen, K.-M. and D. Westmoreland, "Impulse response of a conducting sphere based on singularity expansion method," *Proceeding of the IEEE*, Vol. 69, No. 6, 747–750, June 1981.
6. Wang, S., X. Guan, X. Ma, D. Wang, and Y. Su, "A fast method of moment to calculate monostatic RCS of arbitrary shaped conducting objects," *CIE International Conference on Radar*, 2006.
7. Erez, E. and Y. Leviatan, "Analysis of natural frequencies of cavities and scatterers using an impulsive current model," *IEEE Trans. Antennas Propagat*, Vol. 38, No. 4, 534–540, 1990.

Bandwidth Enhancement Technique in Microstrip Antenna for Wireless Applications

RSA Raja Abdullah, D. Yoharaaj, and Alyani Ismail

Universiti Putra Malaysia (UPM), Malaysia

Abstract— Low profile, light weight, easily mounted and broad bandwidth are the key characteristics for antennae designed for wireless applications. The microstrip antenna (MSA) suits the features mentioned except for its narrow bandwidth. The bandwidth of the microstrip antenna usually ranges from less than 1% to several percent. This paper experimentally investigates an alternative approach in enhancing the bandwidth of the microstrip antenna for the *Wireless Local Area Network (WLAN)* application operating at a frequency of 2.45 GHz. The bandwidth enhancement technique which is studied is the *Identical Dual Patch Microstrip Antenna with Air-Gap (IDMA)*. A bandwidth enhancement of more than 11% was achieved.

1. INTRODUCTION

Wireless applications have undergone rapid development in recent years. One particular wireless application that has experienced this trend is the Wireless Local Area Network (WLAN). The wireless application that is selected to be studied is the 2.4 GHz to 2.5 GHz frequency band which is based on the 802.11 b WLAN standard. This frequency band is very popular due to its low cost.

WLAN antenna requires being low profile, light weight and broad bandwidth. The microstrip antenna suits the features very well except for its narrow bandwidth. The WLAN antenna should have a minimum bandwidth of 100 MHz to fully utilize the WLAN band based on the 802.11b standard. The conventional microstrip antenna could not fulfill this requirement as its bandwidth usually ranges from less than 1% to several percent [1]. Although the required operating frequency range is from 2.4 GHz to 2.5 GHz, at least double the bandwidth is required to avoid expensive tuning operations and to cause uncritical manufacturing. Therefore, there is a need to enhance the bandwidth of the microstrip antenna for WLAN applications.

This paper investigates a technique which can enhance the bandwidth of the microstrip antenna without increasing the lateral size and the complexity of the microstrip antenna too much. The Identical Dual Patch Microstrip Antenna with Air-Gap (IDMA) bandwidth enhancement technique takes the advantage of using the air gap to lower the effective permittivity and increase the total thickness of the microstrip antenna which is essential for bandwidth enhancement. This bandwidth enhanced microstrip antenna can be deployed for the WLAN application operating at a frequency of 2.45 GHz.

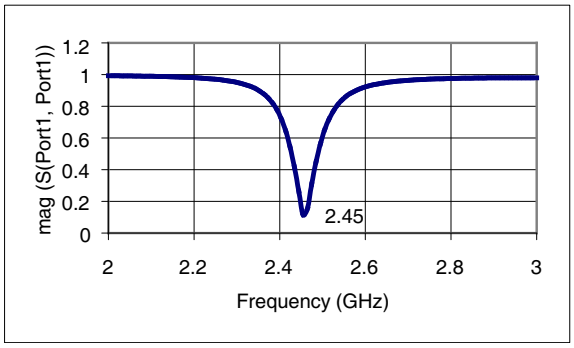
2. SIMULATED AND MEASURED RESULTS OF THE SINGLE-LAYER MICROSTRIP ANTENNA

The basic single-layer microstrip antenna is designed and fabricated to serve as a benchmark for the design of the bandwidth enhanced microstrip antenna. The rectangular probe-fed patch was selected due to its ease of analysis and it is commonly used. The patch is fed with the 50 Ohms-SMA connector. A low permittivity dielectric substrate was used, namely the 1.575 mm thick RT/Duroid 5880. The copper cladding for this substrate is the 35 μm thick rolled copper. A full-wave analysis simulation tool by Ansoft was used. The fabricated microstrip antenna was measured with the Agilent E8362B Network Analyzer.

The antenna dimensions were calculated and detail of the calculation could be found in [2]. These calculated results serve as the starting parameters for the simulation process. Numerous iterative simulations were done to obtain the optimum configuration of the microstrip antenna. Once the desired operating frequency and input impedance are obtained, the bandwidth is taken at VSWR$<$ 2. After the desired operating frequency and input impedance were obtained, the simulated bandwidth was recorded. The optimized configuration of this design is then used to fabricate the microstrip antenna. Table 1, Figure 1, Figure 2 and Figure 3 show the comparison between the simulated and measured results of the single layer fabricated antenna respectively.

Table 1: Simulated and measured results of the single-layer MSA.

Dimensions	***Simulated***	***Measured***
Width	50 mm	same
Length	39.5 mm	same
Location of the Probe	13.25 mm	same
Operating frequency	2.45 GHz	2.425 GHz
Input Impedance	50.69 Ohms	54.64 Ohms
Bandwidth	36 MHz (1.469%)	46.71 MHz (1.699%)



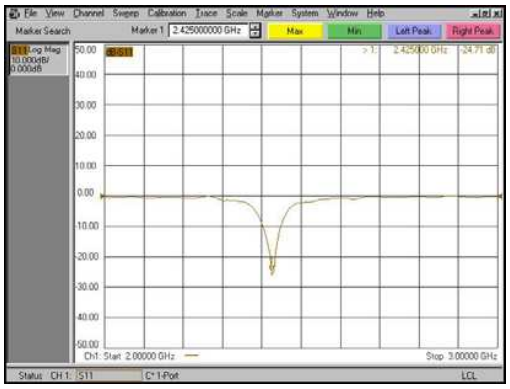
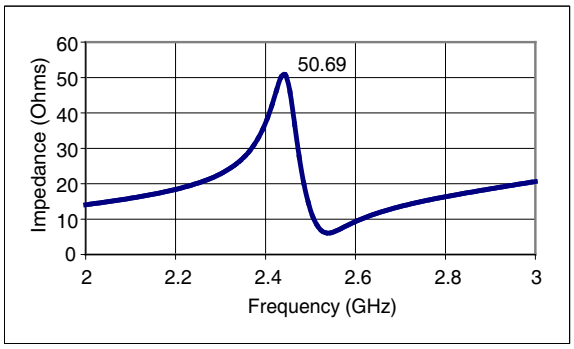


Figure 1: Simulated and measured operating frequency of the single-layer microstrip antenna.



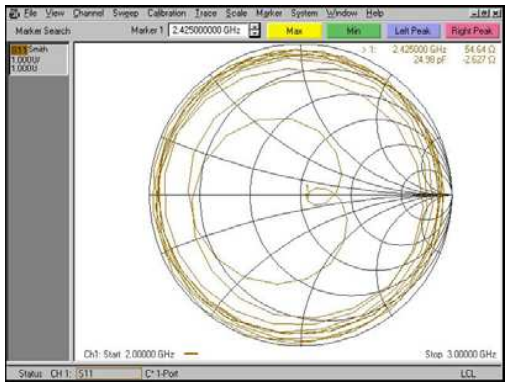
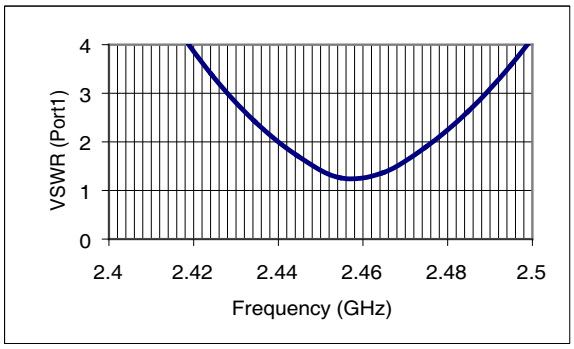


Figure 2: Simulated and measured input impedance of the single-layer microstrip antenna.



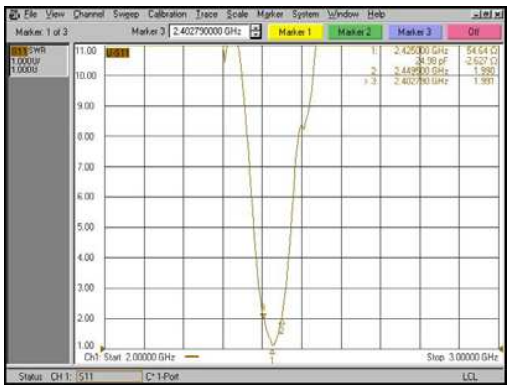


Figure 3: Simulated bandwidth and measured of the single-layer microstrip antenna.

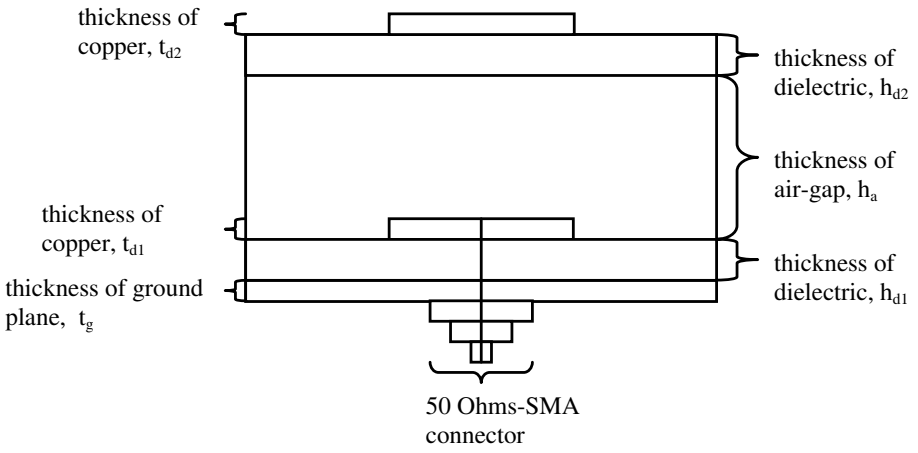


Figure 4: Structure of the IDMA bandwidth enhancement technique.

3. SIMULATED AND MEASURED RESULTS OF IDENTICAL DUAL-PATCH MICROSTRIP ANTENNA WITH AIR-GAP (IDMA)

The Identical Dual-patch Microstrip Antenna with Air-gap (IDMA) takes the advantage of the air gap which lowers the effective permittivity and increase the total thickness of the microstrip antenna which is an essential factor for bandwidth enhancement. The operating frequency of the fabricated microstrip antenna can easily be tuned without the need of a new design by just varying the size of the air gap. The structure of this bandwidth enhanced microstrip antenna is shown in Figure 4. Similar to the design of the single-layer microstrip antenna, the same type of substrate, RT/Duroid 5880 was used and the patch is fed by a 50 Ohms-SMA connector. The stacked patch is identical to the probe-fed patch whereby both these patches are rectangular in shape. The main focus in this design is the effective permittivity of this three-layer microstrip antenna. This parameter and the total thickness of this microstrip antenna influence the bandwidth. The equations that were used in this design are [3–5]:

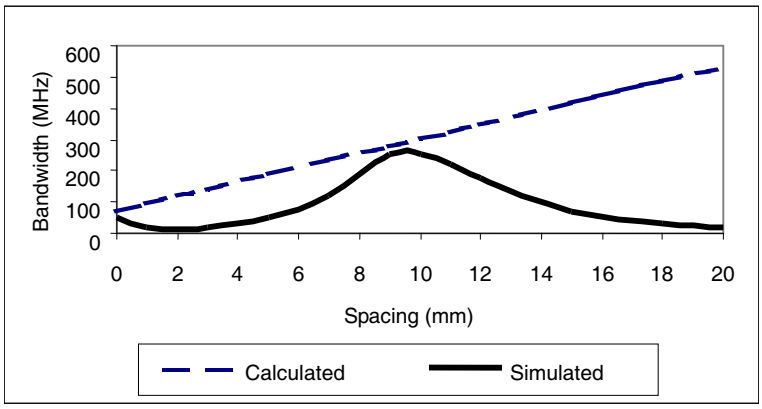


Figure 5: Comparison between the calculated and simulated bandwidth of IDMA.

$$\varepsilon_{av} = \left(\frac{(\varepsilon_r hd1 + \varepsilon_r h_a + \varepsilon_r h_{d2})}{\left(\frac{h_t}{3}\right)} \right) \tag{1}$$

where $h_t = h_{d1} + h_a + h_{d2}$

$$BW = \frac{\sqrt{2}p}{45\pi}\left(1 - \frac{1}{\varepsilon_{av}} + \frac{2}{5\varepsilon_{av}^2}\right)\left(\frac{1}{\varepsilon_{av}}\right)\left(\frac{h_t}{\lambda}\right)\left(\frac{W}{L}\right) \tag{2}$$

where

$$p = 1 + \frac{a_2}{20}(k_0 w)^2 + a_4\left(\frac{3}{560}\right)(k_0 w)^4 + b_2\left(\frac{1}{10}\right)(k_0 L)$$

where $a_2 = -0.16605$, $a_4 = 0.00761$, $b_2 = -0.09142$ and $k_0 = 2\pi/\lambda$

Once the desired operating frequency, input impedance and bandwidth were obtained in the simulation, the optimized configuration of the microstrip antenna was taken. The calculated and the simulated results are compared in order to determine the exact size of the air gap. The results of these comparisons are shown in Figure 5.

Once the optimum bandwidth is obtained which is ~ 250 MHz, the spacing between the probe-fed patch and the stacked patch can be determined. To further increase the accuracy of the simulated results, fine-tuning is done and the maximum achievable simulated bandwidth is ~ 270 MHz. The final configuration of the microstrip antenna is shown in Table 2 and Figure 6 through Figure 8 show the simulated results.

Table 2: Simulated results of IDMA.

Dimensions	***Simulated***
Width	51 mm
Length	41 mm
Location of the Probe	3.5 mm
Spacing (air-gap)	9 mm
Operating frequency	2.45 GHz
Input Impedance	50.15 Ohms
Bandwidth	270 MHz (11.020%)

Table 3: Measured results of IDMA.

Parameter	***Measured results***
Operating Frequency	2.44 GHz
Impedance	61.63 Ohms
Bandwidth	287.77 MHz (11.794%)

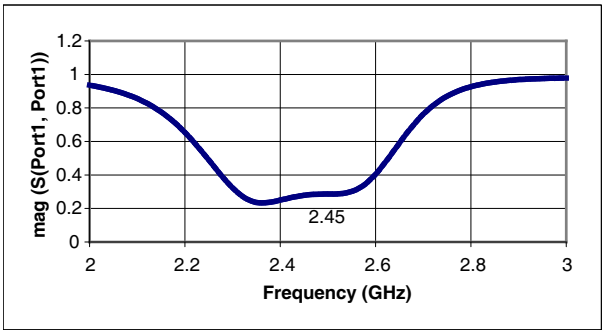


Figure 6: Simulated operating frequency of IDMA.

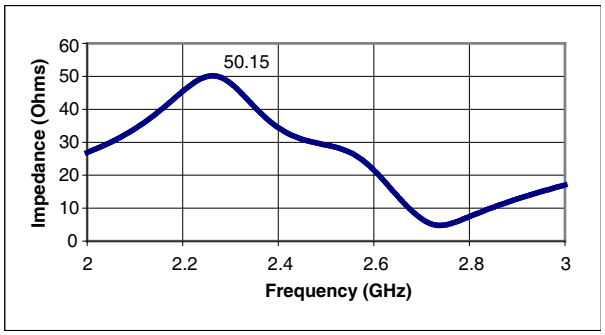


Figure 7: Simulated input impedance of IDMA.

Using the optimized configuration of the microstrip antenna in the simulation results, this design is fabricated. The measured results of this bandwidth enhanced microstrip antenna are shown in Table 3 and Figure 9 through Figure 11 respectively.

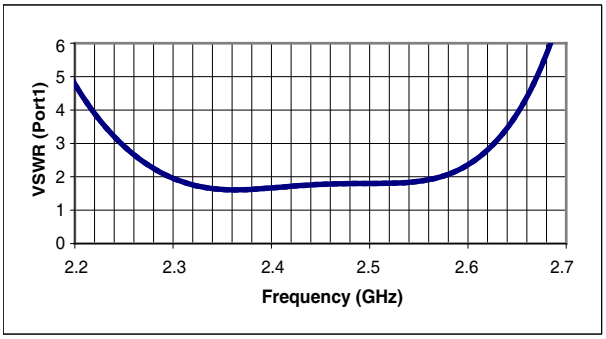


Figure 8: Simulated bandwidth of IDMA.

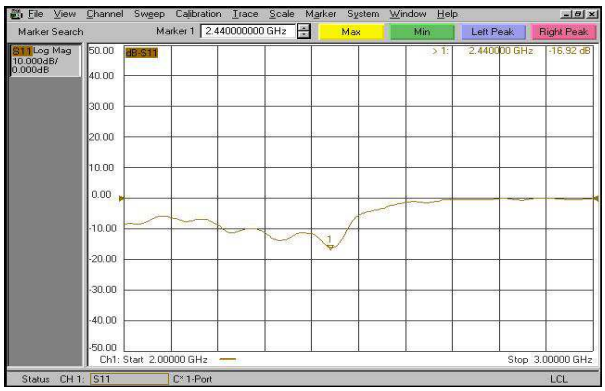


Figure 9: Measured operating frequency of IDMA.

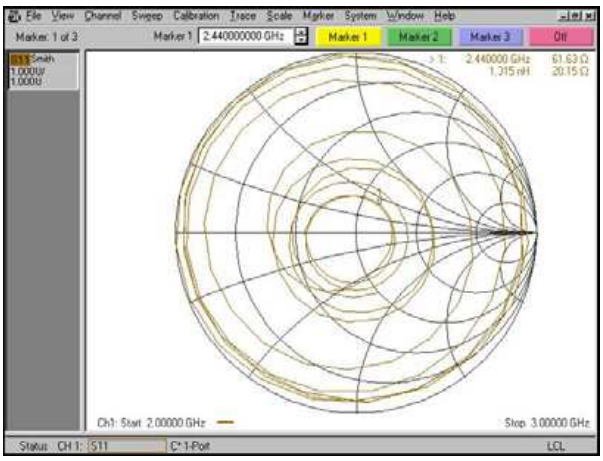


Figure 10: Measured input impedance of IDMA.

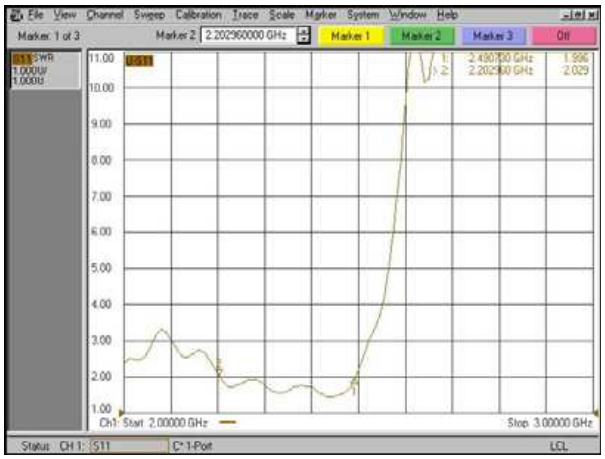


Figure 11: Measured bandwidth of IDMA.

4. DISCUSSION

Based on the results from the simulations and measurements, it is found that the single-layer microstrip antenna has a very narrow bandwidth that is less than 2% which is not sufficient to fully cover the WLAN band based on the 802.11b standard. There is a need to use a bandwidth enhancement technique in this microstrip antenna and the IDMA is deployed. Using this technique, both the simulated and measured results give a bandwidth enhancement at more than 11%. Furthermore, there is a very good agreement between the simulated and measured results. The comparison between the simulated and measured bandwidths of the single-layer microstrip antenna and IDMA are shown in Figure 12 and Figure 13 respectively.

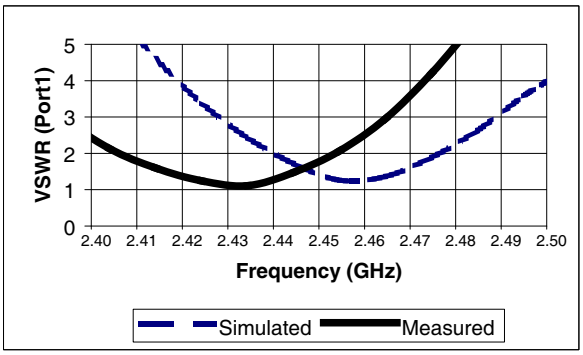


Figure 12: Simulated and measured bandwidth of the single-layer microstrip antenna.

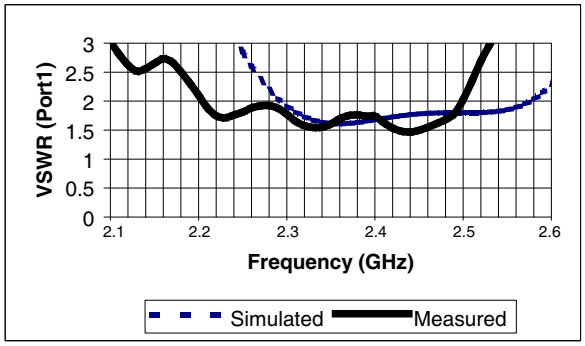


Figure 13: Comparison between the simulated and measured bandwidth of IDMA.

The comparison between the simulated bandwidth of the single-layer microstrip antenna and the IDMA is shown in Figure 14. Table 4 shows the summary of this experimental investigation. It is shown clearly that the simulated and measured results are very similar. Figure 15 shows the fabricated microstrip antennas.

Table 4: Comparison between the simulated and measured results of the single-layer microstrip antenna and IDMA.

Parameters	***Single-Layer Microstrip Antenna***		***IDMA***	
	Simulated	***Measured***	***Simulated***	***Measured***
Operating Frequency (GHz)	2.45	2.4248	2.45	2.44
Impedance (Ω)	50.69	50.15	51.238	61.63
Bandwidth (MHz)	35 (1.429%)	41 (1.4699%)	**270 (11.020%)**	**287.77 (11.794%)**

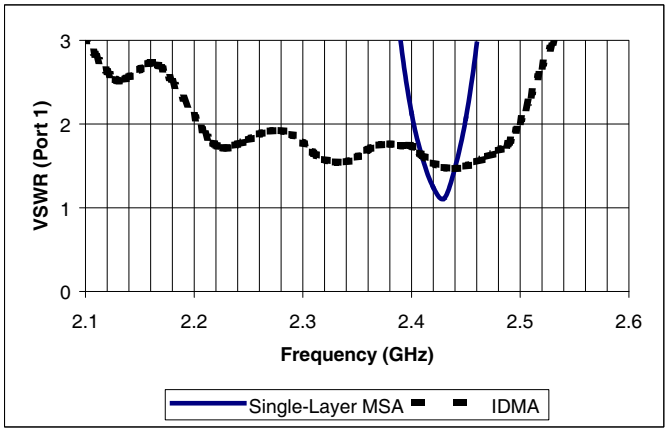


Figure 14: Comparison of the simulated bandwidth.

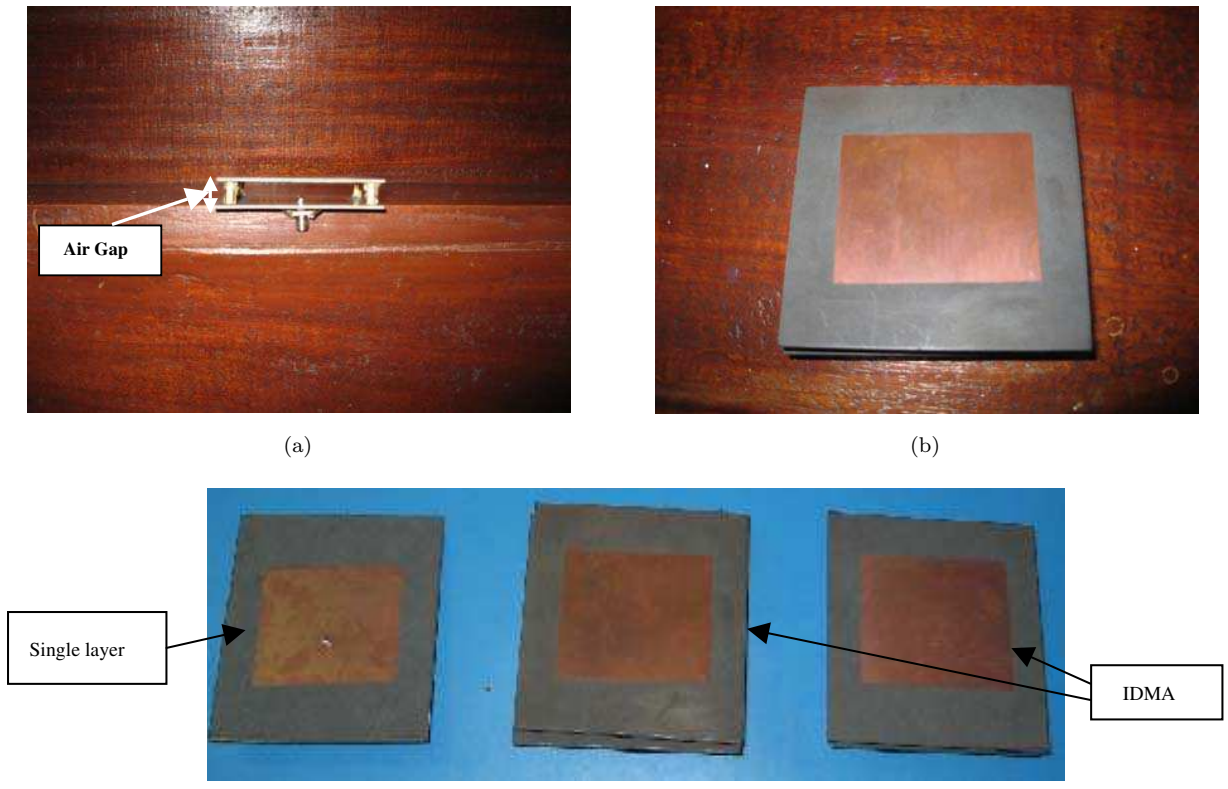


Figure 15: The fabricated microstrip antenna (a) side view of IDMA, (b) top view of IDMA and (c) single-layer microstrip antenna and IDMA.

5. CONCLUSION

The technique for enhancing bandwidth of the Microstrip antenna has been shown and it can be used for WLAN applications as it fully utilizes the entire 2.4 GHz–2.5 GHz band. As mentioned, this technique has its advantages such as it does not increase the lateral size of the microstrip antenna and disadvantages such as it increases the height of the microstrip antenna. Therefore, trade-off issues need to be considered in this design.

REFERENCES

1. Ramesh, G., B. Prakash, B. Inder, and I. Apisak, *Microstrip Antenna Design Handbook*, Artech House, Inc, USA, 2001.
2. Yoharaaj, D., R. S. Azmir, and A. Ismail, "A new approach for bandwidth enhancement technique in microstrip antenna for wireless applications," *International RF and Microwave Conference*, RFM06, Putrajaya, Malaysia, 12–14 September 2006.
3. Nirun, K., N. Chalermpol, and T. Taspong, "Parametric study of the rectangular microstrip antenna with an air gap," *R&D Journal*, Vol. 24, No. 2, King Mongkut's University of Technology Thonburi, 131-142, Bangkok, May–August 2001.
4. Jackson, D. R. and N. Alexopoulos, "Simple approximate formulas for input resistance, bandwidth and efficiency of a resonant rectangular patch," *IEEE Transactions on Antennas and Propagation*, Vol. 39, 409, March 1991.
5. Balanis, C. A., *Antenna Theory: Analysis and Design*, John Wiley & Sons, Inc, USA, 2005.

Design and Analysis of the Novel Test Tube Magnet for Portable NMR Device

Jizhong Chen[1,2], **Yiming Zhang**[2], **and Jijun Xiao**[1]

[1]Graduate University of the Chinese Academy of Sciences
P. O. Box 2703, Beijing CN 100080, China
[2]Institute of Electrical Engineering, Chinese Academy of Sciences
P. O. Box 2703, Beijing CN 100080, China

Abstract— The paper presents a novel test tube magnet (TTM) for Portable NMR device, which is composed of a dipole cylinder magnet and a hemi-cylinder or a hemi-sphere magnet. The hemi-cylinder or the hemi-sphere magnet is attached to one end of the dipole cylinder magnet, so that the whole magnet is shaped like a test tube. TTM generates a wide diameter spherical volume (DSV) with a high magnetic field homogeneity, which makes the device more applicable for portable NMR. The two configurations were simulated with three-dimensional finite-element methods. Flux density elements are effectively corrected in the DSV. The homogeneity increases, the magnet volumes and mass are markedly decreased, and the construction is very reasonable. Taken together the new designs have improved characteristics for portable NMR.

1. INTRODUCTION

In the past few years, portable NMR devices have been rapidly developed, and several important applications have been suggested and realized. The main problem in designing portable NMR device is not so much the production of high magnetic fields, but rather high and homogeneous fields in a large and accessible volume at acceptable weights. Permanent magnets, because of their advantages in terms of cost and no maintenance, are receiving increased attention for the construction of portable NMR devices. Halbach arrays [1] of permanent magnets offer a unique and elegant solution to strong and homogenous magnets. A direct and simple solution is a Halbach arrangement of magnets on a cylinder magnet forming a dipolar magnetic field with the transverse fields direction along the cylinder axis.

For NMR in moderately homogeneous fields, only the inner dipole is of practical interest. This field is perfectly homogeneous for a perfect Halbach geometry and an infinitely long cylinder. However since in reality the cylinder must have a finite length, the strength of magnetic field gradually decreased from the center to the edge of the cylinder.

In order to get an enough homogeneous magnetic field region, axis length of the cylinder magnet must be long enough. In fact, the region is very small in the whole cylinder magnet, in which the homogeneity is satisfied.

A novel test tube magnet (TTM) has been designed for portable NMR device. TTM is composed of a dipole cylinder magnet and a hemi-sphere magnet or a hemi-sphere magnet at one end, which gives the whole arrangement the shape of a test tube. Such a design compensates for the drop of magnetic field at the lower end of the magnet and also increases the homogeneous region. Furthermore, the size and weight of the whole magnet are decreased, which are relative to a cylinder with the same homogeneity. Therefore, the compensated magnetic field in TTM is much more suitable for measuring the sample in the test tube.

2. THEORY

Easy Axis Rotation Theorem is the theoretical foundation of Halbach magnets. If in a 2D, soft-steel free, system all easy axes are rotated by some degree, then all magnetic fields outside the system rotate by the same angle on the contra-side without a change in amplitude.

If the space between the two circles r_{ci} and r_{co}, inner and outer radii of cylinder, is filled with rare earth magnetic material, with a remanence of B_r and easy axis rotated with $\beta(\varphi) = 2\varphi$. Magnet field [1] in the space $r_c < r_{ci}$ is

$$B\left(r_c, \varphi\right) = B_r \ln\left(r_{ci}/r_{co}\right) \tag{1}$$

c being the cylinder magnet. The magnetic field outside the magnet is exactly zero.

If the space between the two spheres r_{si} and r_{so}, inner and outer radii of sphere, is filled with rare earth magnetic material, and the remnant field $\overrightarrow{B_r}$ distribution can be expressed as $\overrightarrow{B_r} = B_r\left(\overrightarrow{e_r}\cos\theta + \overrightarrow{e_\theta}\sin\theta\right)$ ($\overrightarrow{e_r}$, $\overrightarrow{e_\varphi}$, and $\overrightarrow{e_\theta}$ are the unit vectors of spherical coordinates, respectively.), the magnetic field inside the sphere magnet is

$$\overrightarrow{B_s} = \frac{4}{3}B_r \ln\left(r_{so}/r_{si}\right)\overrightarrow{k} \tag{2}$$

s is the sphere magnet.

An infinite dipole magnet is cut by the x-y plane at $z = 0$ (see Fig. 1). If $V_c^{z+}(x, y, z)$ is the scalar potential produced by the dipole magnet when $z > 0$, then the scalar potential produced by the dipole magnet at $z < 0$ must be $V_c^{z-}(x, y, z)$. If $V_c(x, y)$ is the scalar potential inside the infinite dipole, obviously, the following must hold:

$$V_c^{z+}(x, y, z) + V_c^{z-}(x, y, z) = V_c(x, y) \tag{3}$$

Applying the appropriate operator to (5), we get the field quantity $B_c(x, y, z)$ of interest,

$$B_c^{z+}(x, y, z) + B_c^{z-}(x, y, z) = B_c(x, y) \tag{4}$$

One obtains

$$By_c^{z+}(x, y, 0) = By_c^{z-}(x, y, 0) \tag{5}$$

3. DESIGN AND ANALYSIS

Many designs and application of cylindrical magnets [2–4] with Halbach arrays have been done. For the compensation of the transverse field along the cylinder axis, some design principles [4, 5] have been given.

For portable NMR, the sample is always measured in a test tube. Typically, the diameter of a test tube is about one inch. Therefore, the field of the cylinder magnet, with $r_{ci} = 30\,\text{mm}$ and $r_{co} = 35\,\text{mm}$, is simulated by three-dimensional finite-element methods.

There are two methods to get enough DSV of the homogeneous magnetic field region for portable NMR, which are prolonging cylinder magnet axis length and compensating the field on the end.

Table 1: Homogeneity and DSV/L of the Halbach Cylinder.

L(mm)	45	67.5	90	135	180
Homogeneity (ppm)	7150	4064	1928	418	108
DSV/L	0.4	0.27	1928	0.13	0.10

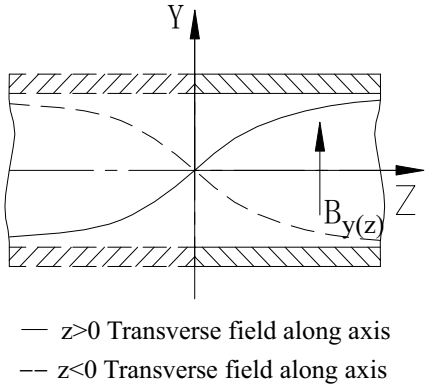


Figure 1: Along axis, the magnetic field distribution of Halbach magnet.

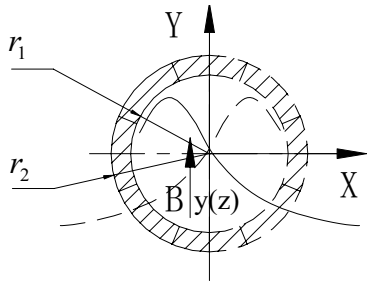


Figure 2: Along X coordinate the magnetic field distribution of Halbach magnet.

Prolonging the length of the cylinder magnet (L) along its axis is one of the traditional methods. In the conditions of DSV=18 mm and $Br = 1.23\,\text{T}$, Table 1 lists the homogeneity and DSV/L produced by some different length cylinders.

One can see from Table 1 that the homogeneity is better, but DSV/L ratio is worse with prolonging the length of the magnet. We can found out the method exhausted more magnet size and weight to obtain more DSV, so that prolonging the magnet axis length is not acceptable for portable NMR.

Next, we consider the method of compensating the field on the end. Based on the field distribution as described in Fig. 1 and the requirement of design and manufacture, two structures are proposed. Hemi-cylinder and hemi-sphere respectively to compensate the decay of the field towards one end of such a cylindrical magnet and hence increase the DSV.

Figure 2 shows a transverse cut through the infinite dipolar magnet in the y-z plane at $x = 0$. The direction of magnetic field parallels Y coordinate. If $B_c^{x+}(x, y, z)$ is the field quantity of interest by the dipole magnet at $x > 0$, then the field quantity of interest at $x < 0$ must be $B_c^{x-}(x, y, z)$. We get the field quantity $B_c(y, z)$ of interest,

$$B_c^{x+}(x, y, z) + B_c^{x-}(x, y, z) = B_c(y, z) \tag{6}$$

$$By_c^{x+}(0, y, z) = By_c^{x-}(0, y, z) \tag{7}$$

Being the same radius r_{ci} and r_{co}, the flux density on the edge of cylinder magnet equals to the field on the section of hemi-cylinder magnet, and both directions are the same on the axis.

$$By_c^{x+}(0, y, z) = By_c^{z+}(x, y, 0) \tag{8}$$

One obtains

$$By_c^{x-}(x, y, z) + By_c^{z+}(x, y, z) = By_c(x, y, z) \tag{9}$$

According to the same principle,

$$By_s(x, y, 0) = By_c(0, y, z) \tag{10}$$

Using (1), (2) in (10), one obtains

$$\frac{4}{3} \cdot Br \ln(r_{so}/r_{si}) = B_r \ln(r_{co}/r_{ci}) \tag{11}$$

$$(r_{so}/r_{si})^{4\beta} = r_{co}/r_{ci} \tag{12}$$

We get,

$$By_s^{x+}(x, y, z) + By_c^{z-}(x, y, z) = By_c(x, y, z) \tag{13}$$

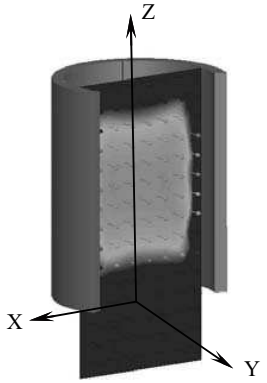


Figure 3: The 3D magnetic field distribution of Halbach magnet.

From the (9), (13), we know that the hemi-cylinder and hemi-sphere can compensate the flux density at the end of cylindrical magnet. The field distribution of a cylindrical magnet is simulated with three-dimensional finite-element methods, as described in Fig. 3. The field distribution of hemi-cylinder magnet and hemi-spherical magnet is simulated, as described in Figs. 4(a), (b).

Along Z-coordinate, the flux density distribution of cylinder magnet can be seen as in Fig. 5 (dashed line) on $z > 0$.The field is relatively uniformity in the center of cylinder magnet, in which

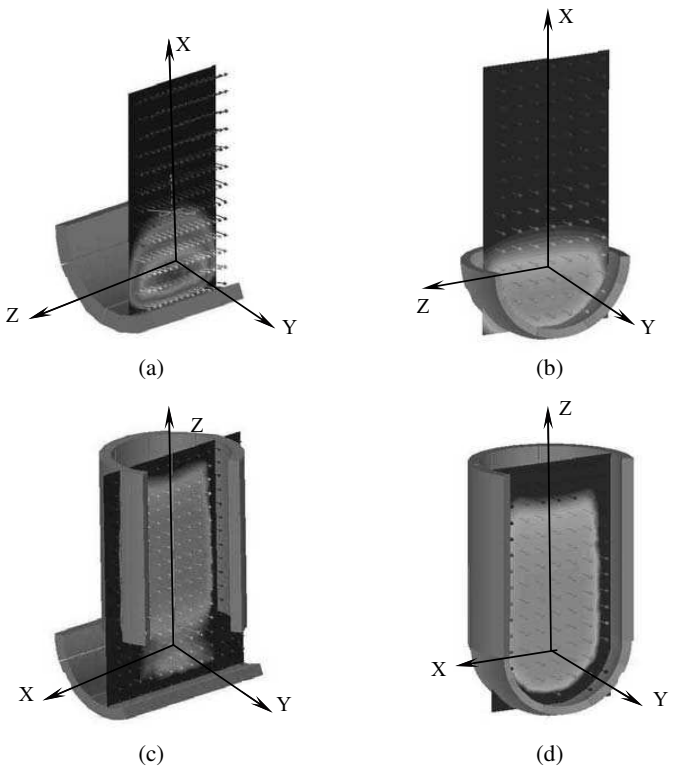


Figure 4: The 3D magnetic filed distribution (a) Hemi-cylinder magnet, (b) hemi-sphere magnet, (c) Hemi-cylinder TTM, (d) hemi-sphere TTM.

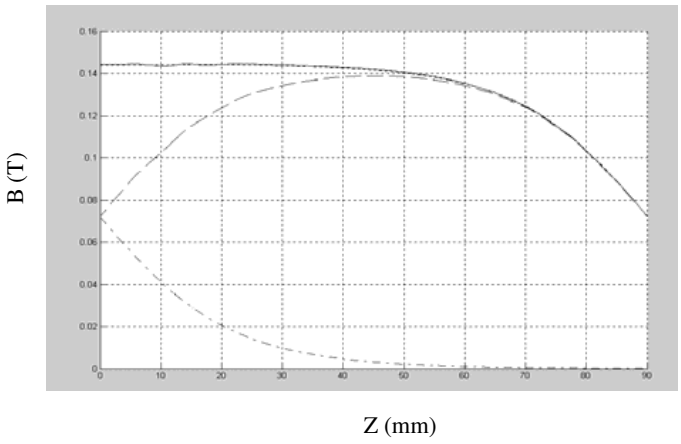


Figure 5: Along the axis, the field distributions of cylinder magnet (dashed line), hemi-cylinder or hemi-sphere magnet (center line), hemi-cylinder TTM (solid line), hemi-sphere TTM (point line).

the samples are detected. Along X-coordinate, on $x > 0$ the flux density distribution of Hemi-cylinder or hemi-sphere magnet can be seen as in Fig. 5 (center line).

The hemi-cylinder or the hemi-sphere magnet is attached to one end of the cylindrical magnet to construct the TTM. If cylinder magnet and hemi-cylinder magnet both have same r_{ci} and r_{co}, and the radius r_{si} and r_{so} of the hemi-sphere magnet is according to (12), the field distribution in the TTM is simulated with three-dimensional finite-element methods, as described in Figs. 4(c), (d).

The distribution of the flux components Bx, By, and Bz as in Figs. 6 (a), (b), (c). The Bx, By, and Bz components of the hemi-cylindrical magnet and hemi-spherical magnet just compensate the Bx, By, and Bz elements of cylinder magnet, so that the Bx and Bz component of TTM are almost zero and By is very uniform in the central space of the TTM.

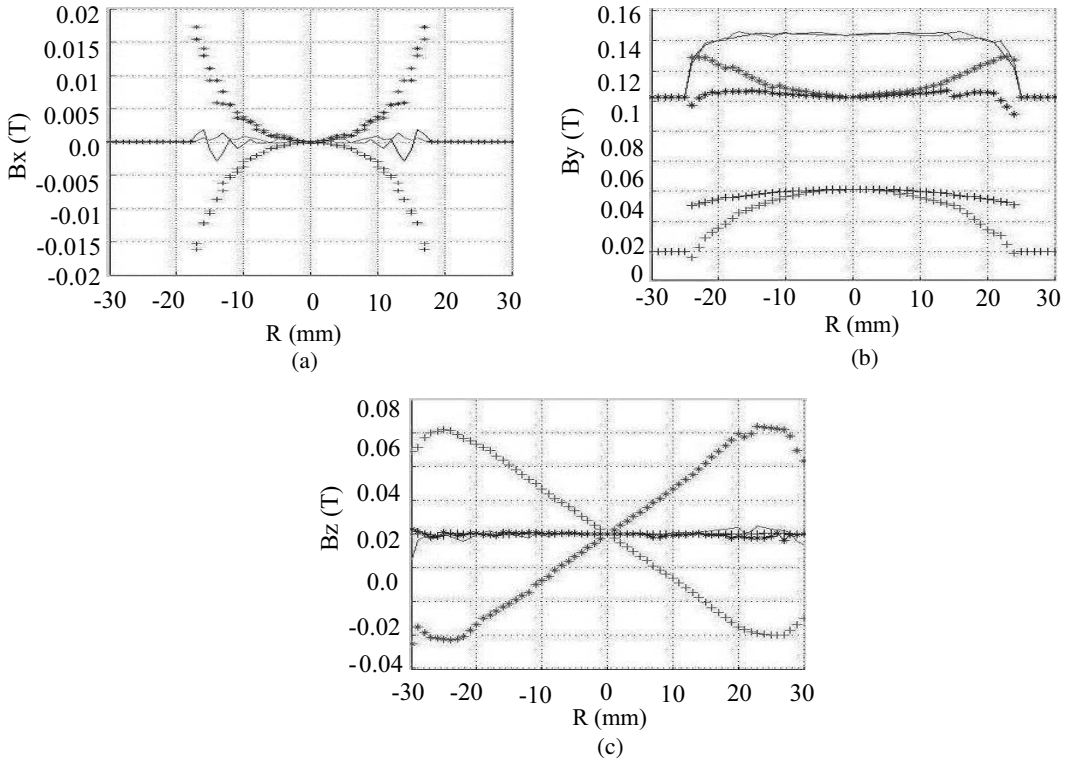


Figure 6: At Z =10 mm, on x-coordinate as gray and y-coordinate as black, the distribution of Bx, By, and Bz in the Cylinder magnet as "+", hemi-cylinder or hemi-sphere magnet as "*", and TTM as "−". (a) Bx element, (b) By element, (c) Bz element.

4. CONCLUSIONS

The obtained results are available in the project discussion of the TTM for a portable NMR device.

- The components are effectively corrected in the region of homogeneous magnetic field.
- As the homogeneity increases markedly, the magnet volumes and mass decreased accordingly, which make their construction more desirable for sample measurement in the test tube.

Taken together the new designs have improved characteristics for portable NMR.

ACKNOWLEDGMENT

This work was supported in part by the Chinese National Foundation Sciences, Project No. 50377043.

REFERENCES

1. Halbach, K., "Design of permanent multipole magnets with oriented rare earth cobalt materials," *Nuclear Instruments and Methods*, Vol. 169, 1–10, 1980.
2. Furlani, E. P., S. Reznik, and W. Janson, "A three-dimensional field solution for bipolar cylinders," *IEEE Trans. Magn.*, Vol. 30, No. 5, Sept. 1994.
3. Furlani, E. P., S. Reznik, and A. Kroll, "A three-dimensional field solution for radically polarized cylinders," *IEEE Trans. Magn.*, Vol. 31, No. 1, Jan. 1995.
4. Raich, H. and P. Blumler, "Design and construction of a Dipolar Halbach array with a uniformity field from identical bar magnets," *NMR Mandhalas Concepts in Magnetic Resonance Part B(Magnetic Resonance Engineering)*, Vol. 23B, No. 1, 16–25, 2004.
5. Hu, Z., "The research of improvement permanent magic ring," *The Master Paper of IEE*, CAS.

Optimization Method for Passive Pole Pieces Design

Le Wang[1], **Haitao Zhu**[1], **Xiaoyu Song**[1], **Zhe Jin**[2], **Xin Tang**[1], **and Weiming Wang**[2]

[1]Institute of Heavy Ion Physics, and Beijing Key Laboratory of Medical Physics and Engineering
Physical School, Peking University, Beijing 100871, China
[2]Key Laboratory for Quantum Information and Measurements, Ministry of Education
School of Electronics Engineering & Computer Science, Peking University, Beijing 100871, China

Abstract— The pole pieces of permanent magnet require precise design since the highly uniform magnetic field are needed for nuclear magnetic resonance imaging. A new method has been introduced in this paper to optimize the shape of passive pole pieces. Based on Filter Equation [1, 3] a field alteration function is proposed in this method to indicate the magnetic field change according to the pole pieces variation. And this function is optimized with BFGS method to determine the shape of pole pieces which produce the most uniform magnetic field. Simulation indicates that this method is much better than the optimization methods in previous published papers with respect to the speed and field uniformity.

In recent years, C-shaped Magnetic Resonance Imaging (MRI) system is being widely used in clinics. The MRI device requires high field uniformity in a large imaging region. In order to generate the highly uniform field, the pole pieces of C-shaped permanent magnet should be designed accurately [4, 5]. In the published papers, there are mainly two ways to calculate the shape of the pole piece. First, many intelligent optimization algorithms, such as Genetic Algorithm [6] and Tabu Search Scheme [7], are used to optimize the shape of the passive shimming pole pieces. Second, Abele points out the magnetic potential in the imaging region can be expanded to specific series, and introduces the filter equations and hybrid pole pieces to totally get rid of low-order harmonics [1, 3]. In this paper a new method for symmetric passive pole pieces design is put forward and some insufficiencies in filter equation method are also modified.

1. METHODS

1.1. Insufficiencies in Filter Equation

According to the reference paper [1], if the pole pieces are composed of materials of infinite permeability and the magnet constitution are approximately axisymmetric, the magnetostatic potential within the reference cylinder with radius r_i, which includes the imaging area, can be expanded as

$$\phi(r,z) = \phi_2 + \frac{z_0+z}{2z_0}\phi_0 + \sum_{n=1,2,3\cdots} a_n I_0\left(n\pi\frac{r}{2z_0}\right)\sin\left(n\pi\frac{z_0+z}{2z_0}\right) \tag{1}$$

where $2z_0$ is the distance between two pole pieces, I_0 is the first kind of modified Bessel function, Φ_0 is potential difference between two pole pieces, and Φ_2 is the potential of lower pole as showed in Fig. 1.

The author also points out some low-order harmonics in Eq. (1) can be eliminated by changing potential on the flat pole pieces. The Filer Eq. (2) is introduced to establish the relation between the potential shift $\delta\Phi$ and harmonics of the field.

$$\frac{\pi^2}{8z_0^2}(-1)^n n\int_{r_i}^{r_e} rK_0\left(n\pi\frac{r}{2z_0}\right)\times\left[\delta\phi_+(r)-(-1)^n\delta\phi_-(r)\right]dr = \Delta a_n \tag{2}$$

K_0 is the third kind of modified Bessel function, and $\delta\Phi_+$ and $\delta\Phi_-$ are potential shift on two pieces respectively.

However, from Eq. (1), the magnetic induction intensity along z axis is

$$Bz(r,z) = -\mu_0\frac{\partial\phi(r,z)}{\partial z} = -\frac{\mu_0}{2z_0}\phi_0 - \sum_{n=1,2,3\cdots} a_n\frac{n\pi}{2z_0}I_0\left(n\pi\frac{r}{2z_0}\right)\cos\left(n\pi\frac{z_0+z}{2z_0}\right) \tag{3}$$

$$Bz(0,0) = -\mu_0\left.\frac{\partial\phi(r,z)}{\partial z}\right|_{r=z=0} = -\frac{\mu_0}{2z_0}\phi_0 - \sum_{n=2,4,6\cdots} a_n\frac{\mu_0 n\pi}{2z_0}(-1)^{n/2} \tag{4}$$

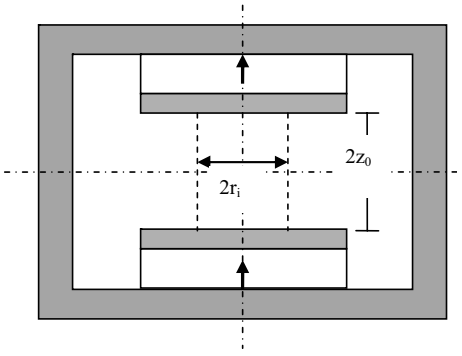


Figure 1: Main magnet structure of the magnetic resonance imaging system. The shadow region are pole pieces and yoke, the arrow indicate the direction of magnetization in permanent magnet.

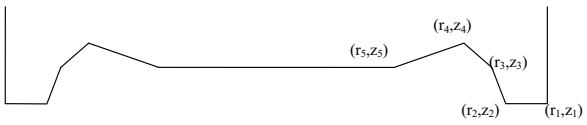


Figure 2: Shape of double shimmed passive pole piece.

Eq. (4) indicates the harmonics is not only the distortion of the field they also contribute to the central field, namely the uniform field. So what we need to do is using some low-order harmonics to counteract the influence from high order harmonics in the imaging area instead of totally eliminating some low-order harmonics.

Passive pole pieces design which consists of several convex or concave annuluses on the pole is widely used in products to improve the field uniformity [4, 5]. Since the convex or concave degree is nonlinear with the field, some intelligent optimization methods are used with 2D FEM to optimize the shape of the pole pieces. As for these methods, the optimization process is time-consuming because FEM must be used to calculate the field when the shape of pole pieces changes in every step.

1.2. Field Alteration Function

In order to solve the problems mentioned above, a method is brought forward which is mainly composed of three steps. First, the filter equation is used to establish a connection between the change of shape of pole pieces and the change of the harmonics. So a Field Alteration Function is established whose variable is no longer the shape of the pole but the magnitude of low-order harmonics. Second, the Field Alteration Function is optimized in order to generate the most uniform field in the imaging area. Third, shape of pole pieces is calculated to get the objective low-order harmonics.

The influence of a convex or concave annulus to the field is equivalent with a specific potential change $\delta\Phi$ on a flat pole piece. As for a double shimmed pole pieces in Fig. 2, five points which are connected with straight lines determine the shape of the pole pieces. Only the apexes of each annulus can move in z direction, which means only z_1, z_2 and z_4 can change and $z_1 = z_2$. For each annulus, when the apex changes we suppose the change of $\delta\Phi$ approximately depends only on the z value of the apex. So when z_1 changes from $z_{1,0}$ to $z_{1,0} + \Delta z_1 (z1 = z2)$, the equivalent potential change shifts from $\delta\Phi_1(r)$ to $\sigma\Phi_1'(r)$, which have relation:

$$\delta\phi_1'(r) = A(\Delta z_1)\delta\phi_1(r) \tag{5}$$

Similarly, when z_4 changes from $z_{4,0}$ to $z_{4,0} + \Delta z_4$, the equivalent potential change shifts from $\delta\Phi_2(r)$ to $\sigma\Phi_2'(r)$, which have relation:

$$\delta\phi_2'(r) = B(\Delta z_4)\delta\phi_2(r) \tag{6}$$

where, $\boldsymbol{A}$ and $\boldsymbol{B}$ are two functions we don't know. Then the magnetic field alteration after the change of z_1 and z_4 is:

$$\begin{bmatrix} D_{1,1} & D_{1,2} \\ D_{2,1} & D_{2,2} \\ D_{3,1} & D_{3,2} \\ \vdots & \vdots \end{bmatrix} \times \begin{bmatrix} A(\Delta z_1) - 1 \\ B(\Delta z_4) - 1 \end{bmatrix} = \begin{bmatrix} \Delta a_1 \\ \Delta a_2 \\ \Delta a_3 \\ \vdots \end{bmatrix} \tag{7}$$

The coefficient $D_{n,m}$ are given by

$$\begin{aligned} D_{n,1} &= \frac{\pi^2}{2y_0^2} n \int_{r_1}^{r_3} rK_0\left(n\pi\frac{r}{z_0}\right) \times \delta\phi_1(r)dr \\ D_{n,2} &= \frac{\pi^2}{2y_0^2} n \int_{r_3}^{r_5} rK_0\left(n\pi\frac{r}{z_0}\right) \times \delta\phi_2(r)dr \end{aligned} \tag{8}$$

$\Delta a_1, \Delta a_2, \Delta a_3, \ldots$ indicate the harmonics' alteration.

The change of harmonics Δa_n can be got through calculating magnetic field with 2D FEM. So we will get the value of

$$\begin{bmatrix} D_{1,1} \\ D_{2,1} \\ D_{3,1} \\ \vdots \end{bmatrix} \times (A(\Delta z_1) - 1) \tag{9}$$

when z_1 increase from $z_{1,0}$ to $z_{1,0} + \Delta z_1$ without changing $z_{4,0}$.

In the same way, we can also get

$$\begin{bmatrix} D_{1,2} \\ D_{2,2} \\ D_{3,2} \\ \vdots \end{bmatrix} \times (B(\Delta z_4) - 1) \tag{10}$$

Let

$$\begin{bmatrix} [A(\Delta z_1) - 1] D_{1,1} & [B(\Delta z_4) - 1] D_{1,2} \\ [A(\Delta z_1) - 1] D_{2,1} & [B(\Delta z_4) - 1] D_{2,2} \\ [A(\Delta z_1) - 1] D_{3,1} & [B(\Delta z_4) - 1] D_{3,2} \\ \vdots & \vdots \end{bmatrix} = \begin{bmatrix} E_{1,1} & E_{1,2} \\ E_{2,1} & E_{2,2} \\ E_{3,1} & E_{3,2} \\ \vdots & \vdots \end{bmatrix} \tag{11}$$

If there is arbitrary change of z_1 and z_4, such as $z_{1,0} + \Delta z_1'$ and $z_{4,0} + \Delta z_4'$, the harmonics change will be

$$\begin{bmatrix} E_{1,1} & E_{1,2} \\ E_{2,1} & E_{2,2} \\ E_{3,1} & E_{3,2} \\ \vdots & \vdots \end{bmatrix} \times \begin{bmatrix} A(\Delta z_1') - 1 \Big/ A(\Delta z_1) - 1 \\ B(\Delta z_4') - 1 \Big/ B(\Delta z_4) - 1 \end{bmatrix} = \begin{bmatrix} \Delta a_1 \\ \Delta a_2 \\ \Delta a_3 \\ \vdots \end{bmatrix} \tag{12}$$

From Eq. (12)

$$\begin{bmatrix} \Delta a_3 \\ \Delta a_4 \\ \vdots \end{bmatrix} = \begin{bmatrix} E_{3,1} & E_{3,2} \\ E_{4,1} & E_{4,2} \\ \vdots & \vdots \end{bmatrix} \times \begin{bmatrix} A(\Delta z_1') - 1 \Big/ A(\Delta z_1) - 1 \\ B(\Delta z_4') - 1 \Big/ B(\Delta z_4) - 1 \end{bmatrix} = \begin{bmatrix} E_{3,1} & E_{3,2} \\ E_{4,1} & E_{4,2} \\ \vdots & \vdots \end{bmatrix} \times \begin{bmatrix} E_{1,1} & E_{1,2} \\ E_{1,2} & E_{2,2} \end{bmatrix}^{-1} \times \begin{bmatrix} \Delta a_1 \\ \Delta a_2 \end{bmatrix} \tag{13}$$

Although we don't know $\boldsymbol{A}(\Delta z_1')$ and $\boldsymbol{B}(\Delta z_4')$ when apexes change arbitrarily, we can utilize the alteration of a_1 and a_2 to indicate the change of harmonics, which means the field alteration can be expressed according to a_1 and a_2

$$\begin{aligned} F(a_1, a_2, r, z) \;=\; & \phi_2 + \frac{z_0 + z}{2z_0}\phi_0 + \sum_{n=1}^{2} a_n I_0\left(n\pi\frac{r}{z_0}\right) \sin\left(n\pi\frac{z_0 + z}{z_0}\right) \\ & + \left[I_0\left(3\pi\frac{r}{z_0}\right) \sin\left(3\pi\frac{z_0 + z}{z_0}\right) I_0\left(4\pi\frac{r}{z_0}\right) \sin\left(4\pi\frac{z_0 + z}{z_0}\right) \cdots \right] \\ & \times \left(\begin{bmatrix} E_{3,1} & E_{3,2} \\ E_{4,1} & E_{4,2} \\ \vdots & \vdots \end{bmatrix} \times \begin{bmatrix} E_{1,1} & E_{1,2} \\ E_{2,1} & E_{2,2} \end{bmatrix}^{-1} \times \begin{bmatrix} a_1 - a_{1,0} \\ a_2 - a_{2,0} \end{bmatrix} + \begin{bmatrix} a_{3,0} \\ a_{4,0} \\ \vdots \end{bmatrix} \right) \end{aligned} \tag{14}$$

where $a_{n,0}$ correspond to the harmonics when $z_1 = z_{1,0}$ $z_4 = z_{4,0}$. Eq. (14) is the Field Alteration Function we needed.

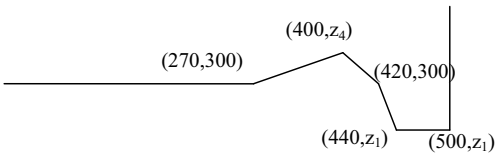


Figure 3: Shape of specific passive pole pieces with some given points.

1.3. Optimize Field Alteration Function

To produce the most uniform field for magnetic resonance imaging, the objective function should be defined as

$$f(a_1, a_2) = \max\left\{-\frac{\partial F(a_1, a_2, r, z)}{\partial z}\bigg|_{(r,z)\in\Omega}\right\} - \min\left\{-\frac{\partial F(a_1, a_2, r, z)}{\partial z}\bigg|_{(r,z)\in\Omega}\right\} \tag{15}$$

where Ω is the imaging area. Usually Ω is a sphere or an ellipsoid. As an example showed in Fig. 3, the diameter of pole pieces is 1000 mm, the distance between the centers of two poles is 600 mm and Ω is a sphere with radius 180 mm. From Fig. 4, the contour map of $f(a_1, a_2)$, we can find $f(a_1, a_2)$ is a totally a convex function. So BFGS method is used to optimize $f(a_1, a_2)$. Fig. 4 also demonstrates that it is not the optimum result when $a_1 = a_2 = 0$. Since it is obvious that the optimum result should be near the $a_1 = a_2 = 0$ point, we can choose $a_1 = a_2 = 0$ point to be the initial point in optimization process.

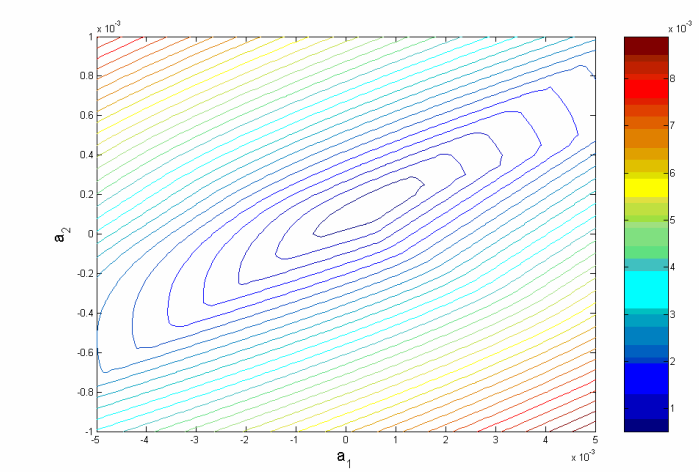


Figure 4: Contour map of the objective function $f(a_1, a_2)$.

The last step is we need to know the value of z_1 and z_4 which will produce the optimum value $a_{1,op}$ and $a_{2,op}$. Because $\boldsymbol{A}(\Delta z_1)$ and $\boldsymbol{B}(\Delta z_4)$ are monotonic functions, which means the equivalent potential $\delta\Phi(r)$ will change in the same or opposite direction with the shift of apex, Newton's Method can be used to solve this nonlinear problem and can be expressed as

$$\begin{bmatrix} z_1^{i+1} \\ z_4^{i+1} \end{bmatrix} = \left(\left(\begin{bmatrix} E_{1,1} & E_{1,2} \\ E_{2,1} & E_{2,2} \end{bmatrix}^{-1} \begin{bmatrix} a_{1,op} - a_1^i \\ a_{2,op} - a_2^i \end{bmatrix}\right)^T \begin{bmatrix} \Delta z_1 & 0 \\ 0 & \Delta z_4 \end{bmatrix}\right)^T + \begin{bmatrix} z_1^i \\ z_4^i \end{bmatrix} \tag{16}$$

In this iterative scheme, Δz_1 and Δz_4 are very small given value to calculate $E_{m,n}$ when $z_{1,0} = z_1^i$ $z_{4,0} = z_4^i$. And the value of a_1 and a_2 need to calculate each time. When this procedure converges, it produces good z_1 and z_4 to produce $a_{1,op}$ and $a_{2,op}$ we need.

2. RESULTS AND DISCUSSION

Figure 5 shows the algorithm for the example in Fig. 3, in which $\Delta z_1 = \Delta z_4 = 0.01$ mm and δ is a constant to determine whether to change the initial value and carry out the algorithm again or

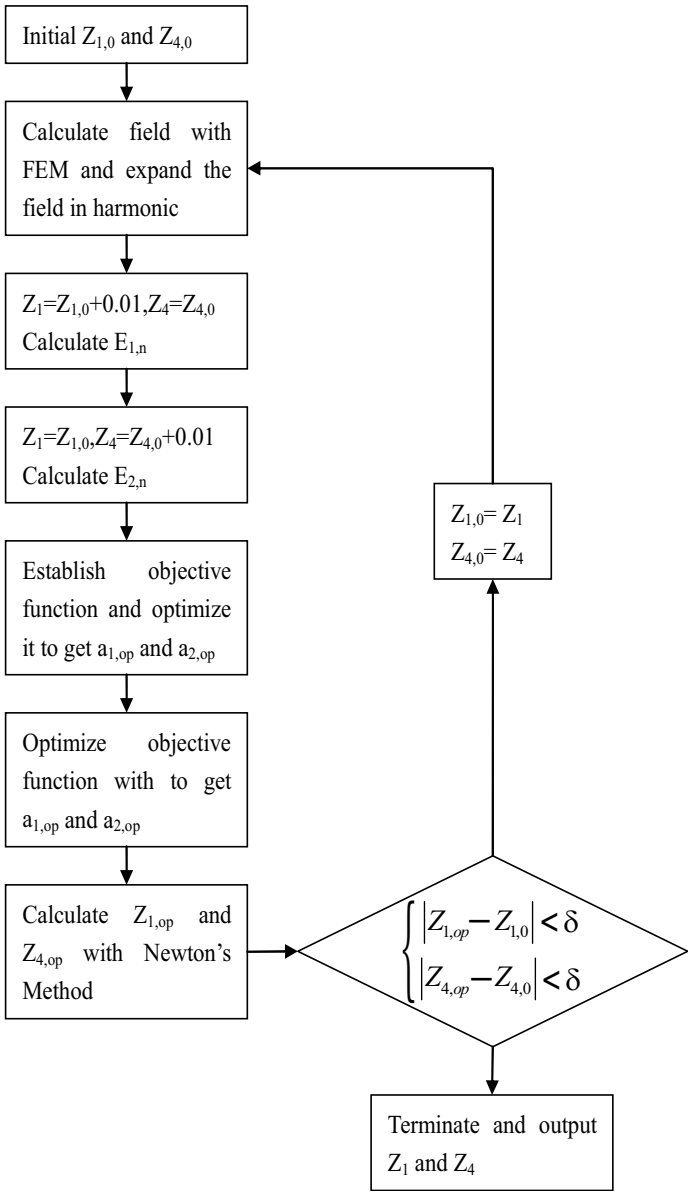


Figure 5: Flow chart of this algorithsm.

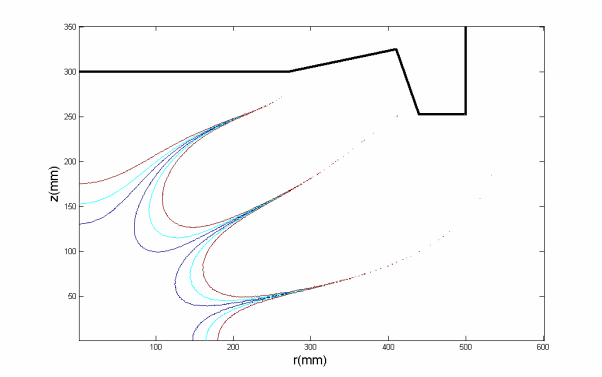


Figure 6: Equi-magnetic induction lines calculated with a 2-D FEM. (Three lines are 20 ppm, 50 ppm and 100 ppm contour lines).

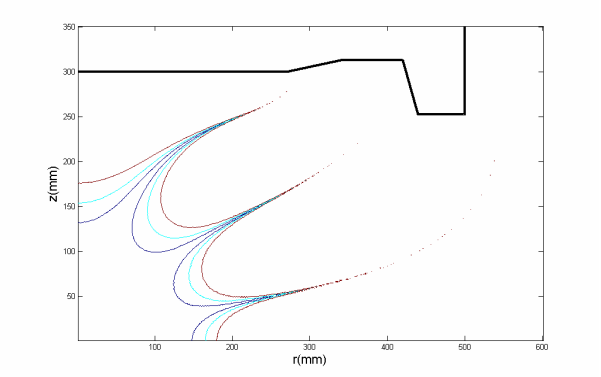


Figure 7: Equi-magnetic induction lines calculated with a 2-D FEM for double shimmed pole pieces with trapezoidal concave annulus. (Three lines are 20 ppm, 50 ppm and 100 ppm contour lines).

not. In this example δ may be 20 mm or more than 20 mm. Considering the lowest eight orders of harmonics are well describe the field, only the lowest eight orders of harmonics are utilized to establish the Field Alteration Function. From analysis in this process, $E_{m,n}$ just change a little at different z_1 and z_4, so $E_{m,n}$ needn't to be calculated at each time when using Newton's Method. Fig. 6 shows the result of this example with 212 ppm(peak to peak) uniformity in a sphere with radius 360 mm. This method can also be used to determine other kinds of pole shape. Fig. 7 shows a result of double shimmed pole pieces with trapezoidal concave annulus.

3. CONCLUSIONS

This new method can be applied to all kinds of passive pole pieces designs proposed in previous papers. Compared to the intelligent optimization methods, it is much faster to use this method, since FEM is avoided during optimization. Otherwise, the shape of pole pieces put forward by this method can produce the most uniform field in imaging area, which is more uniform than the result produced by some search methods [6].

REFERENCES

1. Abele, M. G. and J. H. Jensen, "Hybrid pole pieces for permanent magnets," *J. Appl. Phys.*, Vol. 79, No. 8, 5199–5201, 1996.
2. Abele, M. G., *Structures of Permanent Magnets*, Wiley, New York, 1993.
3. Abele, M. G., J. H. Jensen, and H. Rusinek, *13th International Workshop on Rare-earth Magnets and Their Applications*, 167–176, Birmingham, U.K., September 1994.
4. Molfino, P. and M. Repetto, "Design of an axisymmetric permanent magnet structure for magnetic resonance tomography," *IEEE Transactions on Magnetics*, Vol. 24, No. 2, 994–997, 1988.
5. Miyamoto, T., H. Sakurai, H. Takabayashi, et al., "A development of a permanent magnet assembly for MRI devices using Nd-Fe-B material," *IEEE Transactions on Magnetics*, Vol. 25, No. 5, 3907–3909, 1989.
6. Zhang, Y., D. Xie, and P. Xia, "Passive shimming and shape optimization of an unconventional permanent magnet for MRI, electrical machines and systems, 2003," *Sixth International Conference on ICEMS 2003*, Vol. 28, 99–902, 9–11 Nov. 2003.
7. Leonard, P. J. and A. M. Connor, "Pole shape optimization using a tabu search scheme," *IEEE Transactions on Magnetics*, Vol. 36, No. 4, July 2000.

Verification of BGA Package in RF Application

Zhihong Dong and Zhiping Yu
The Institute of Microelectronics of Tsinghua University, China

Abstract— The characteristics of BGA allow for higher pin counts, higher electrical and thermal performance and higher interconnect density; can meet the specifications of the next generation wireless communication systems, especially the use in RF module. In this paper I will verify the electrical performance of the BGA.

1. INTRODUCTION

An electronic package as defined in [1] is that portion of an electronic structure, which serves to protect an electronic/electrical element from its environment and the environment from its electronic/electrical element and should also allow for the complete testing of the packaged device. The basic function of a package is to provide signal distribution, heat dissipation, power distribution and circuit support and protection, but integrated circuit packaging is beginning to affect the performance of the integrated circuit with the rise in operating frequencies. Hence it is very stringent to have new package to meet this specifications.

BGA are commonly used for graphics IC, PLD, DSP, PC chipsets, communications, networking, microprocessors/controllers, ASIC, and other logic and mixed signal applications and are found in all types of computers and electronic equipment. In order to verify BGA in RF application, in this paper I will first run the simulation of BGA using HFSS, 3D FEM simulation tool, and then will model the package, that's to say to provide equivalent circuit models for the package, to provide better understanding of the effect from the BGA to the chip IC; At last, I will give some analysis of the BGA simulation result.

2. HFSS SIMULATION

HFSS (high frequency structural simulator) is a high-performance full-wave electromagnetic field simulator for arbitrary 3D volumetric passive device modeling. It employs the Finite Element Method (FEM), adaptive meshing to provide unparalleled performance and insight to all 3D EM problems.

The basic operation is to draw the structure, specify material properties for each layer, identify conductors, assign boundaries and specify port excitations. Then you need to specify the simulation options before running the simulation. At last you can get the S-parameters for the package model.

The structure building is very complex, as what we have are a series of 2D drawing of the die, top signal, bottom signal, via, ball etc. So we have to import each layer and assign the corresponding width and move to the corresponding position and with the 3D operation of unite, subtract to get the final 3D structure. The under test BGA package is a full module, its RF module includes LNA, PA, PLL, LO and Transceiver. As what we mind is the RF module, so the other module will be deleted from the structure, in order to accelerate the simulation process.

The air size should between $\lambda/10$ and $\lambda/4$, with the boundary of radiation.

Although the BGA under test is a two layer package (Figure 1), without power layer and ground layer, in order to assign lumped port to corresponding port we have to add ground layer to the structure with the boundary of perfect E.

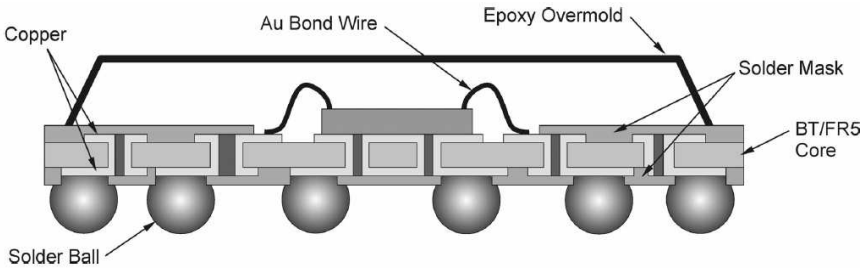


Figure 1: Two layer BGA structure.

I will concentrate on the performance of the transmitter. For the transmitter differential input (TXA and TXB) from the die, we should assign two ports from the ground to each die respectively. After assigning other ports and the simulation, we can get the simulation result of the S-parameters.

3. MICROWAVE OFFICE MODELING ACCORDING TO THE S-PARAMETERS

For accurate electrical characterization of an electronic package, it is important to have proper circuit models for the components that make up the package including the pad, bond wire, signal line, via, solder ball etc. What's more, we have to take all possible couple capacitance and couple inductance into account. According to the BGA structure in HFSS, we can get the equivalent model of the TX part, as in Figure 2, but the specific value of each RLC component is assigned randomly, we have to give each value a range and through the optimization to get the final value of each component.

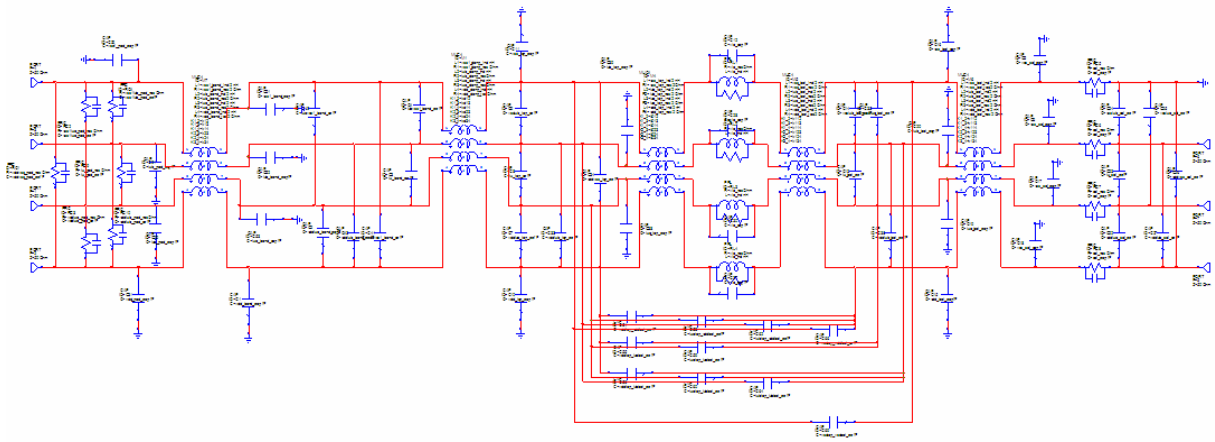


Figure 2: Lumped element representation of BGA package.

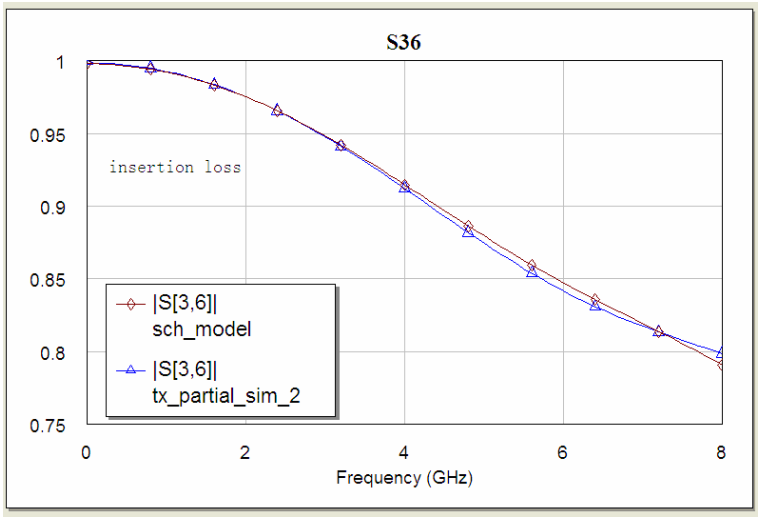


Figure 3: Comparison of equivalent model result and HFSS result.

The equivalent model simulation results & BGA simulation results are shown in Figure 3. Now I will give some analysis of the BGA simulation result with HFSS. There are some factors may have influence on the simulation result: first the BGA model in this paper is a two layer structure, with no power layer and ground layer, but during the simulation in HFSS, we have to assign lumped port, which is assigned from the ground to the specific plane of the port, so we have to add a ground layer at the bottom of the package model in HFSS in order to assign lumped port to the solder ball and the pad on the die, which may affect the simulation result; second, the size and height of solder ball (at the bottom of the package) will affect its parasitic capacitance to the ground, thus to affect the simulation result.

4. CONCLUSION

We have verified the performance of BGA package in RF front-end module, and give the equivalent circuit for the BGA package. The insertion loss is $0.8(-1.95\,\mathrm{dB})$ at 8 GHz, the reflection loss is $-33\,\mathrm{dB}$ at 8 GHz, which is very competent in RF application.

ACKNOWLEDGMENT

I want to thank Accelsemi. Inc. for providing the summer intern opportunity and support in HFSS using.

REFERENCES

1. Brown, W. D., "Advanced electronics packaging with emphasis on multi chip module," *IEEE Press Series on Microelectronic System.*
2. ANSOFT HFSS user guide.
3. Hu, H., "Investigation of the effects of ground vias, solder balls and wires in high-speed packages," ANSOFT workshop presentation.

Measurement of Radio Frequency Magnetic Field

K. Bartusek[1] **and E. Gescheidtova**[2]

[1]Institute of Scientific Instruments, Academy of Sciences of the Czech Republic, Czech Republic
[2]Faculty of Electrical Engineering and Communication, Brno University of Technology, Czech Republic

Abstract— A high quality of magnetic fields is one of the basic prerequisites for a perfect operation of the MR tomographic scanner. Magnetic susceptibility of materials that are used for body implants causes much distortion in MR images. The paper deals with mapping the radio frequency magnetic field induction in the vicinity of these materials with the aim of characterizing the magnetic field and comparing the results with fields simulated in the ANSYS program. Magnetic resonance method of radio frequency magnetic field mapping was verified and optimized for both the characterizing of MR imaging detecting coil and the effects of electrical conducting samples. The designed method is based on optimized spin-echo imaging sequence with small flip angle.

1. INTRODUCTION

The principle of reciprocity states a proportional relationship of the sensitivity during transmission of the RF (radio frequency) impulse relative to the amplitude of the received NMR signal using a receive transmit RF coil [1]. This relation has been used for absolute quantification in single voxel studies under the assumption that the B_1 field is homogenous throughout the voxel volume.

The local B_1 field strength in each examination can be obtained using repeated acquisitions with different RF impulse amplitudes ('transmit gain', 'flip angle') [2, 3]. The inhomogeneity can be corrected by the use of a map of the spatial distribution of the B_1-field acquired either through computer simulation, or through measurement of the field. There are two major advantages by actually measuring the B_1 distribution; it makes the registration of the map against MR-images easier and also a measured B_1 field compensates for imperfections in the coil design. The latter is not easily incorporated into a computer model.

High homogeneity of RF field in working space of MR tomograph depends on construction of a probe and on its adjustment. Mapping of RF magnetic field by the use of the described method is based on a image MR measurement of the homogenous testing specimen with optimally determined flip angle for obtaining the maximum contrast in the measured map of RF magnetic filed.

2. METHOD

Flip angle α is defined between magnetization vectors before and after excitation by RF impulse. An extension of the flip angle depends on the energy, which the excited nuclei obtain, it means on magnetic induction of B_1 field, on the duration of RF impulse (t_{RF}), and on gyromagnetic ratio of the nuclei γ. For a general, amplitude-modulated RF impulse is:

$$\alpha = \gamma \int_0^{t_{\mathrm{RF}}} B_1 \mathrm{d}t \,. \tag{1}$$

In the case of RF impulse with rectangular envelope we can the flip angle express by simple expression

$$\alpha = \gamma B_1 t_{\mathrm{RF}} \,. \tag{2}$$

In MR tomography the flip angle $\alpha = 90^\circ$ is interesting, for which the transversal component of magnetization M_{xy} is maximum (corresponds to amplitude of scanned signal). The size of xy part of flip angle is after end of excitation given by

$$M_{xy} = M_0 \sin \alpha \,. \tag{3}$$

Dependences of the maximum amplitude M in a image on the flip angle read form measured images and the theoretical course calculated according Equation (1) are predicted in Fig. 1.

For the image scanning in MR tomograph in Institute of Scientific Instruments, Academy of Sciences of the Czech Republic the RF amplitude-modulated impulse is used, which shape and

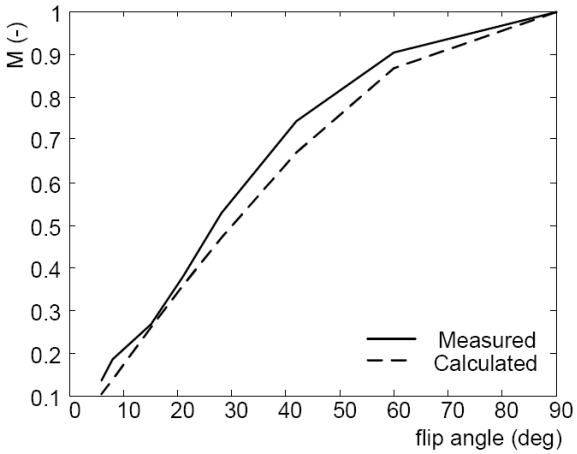


Figure 1: Dependences of maximum amplitude M in an image on the flip angle read from the measured images, and the theoretical course.

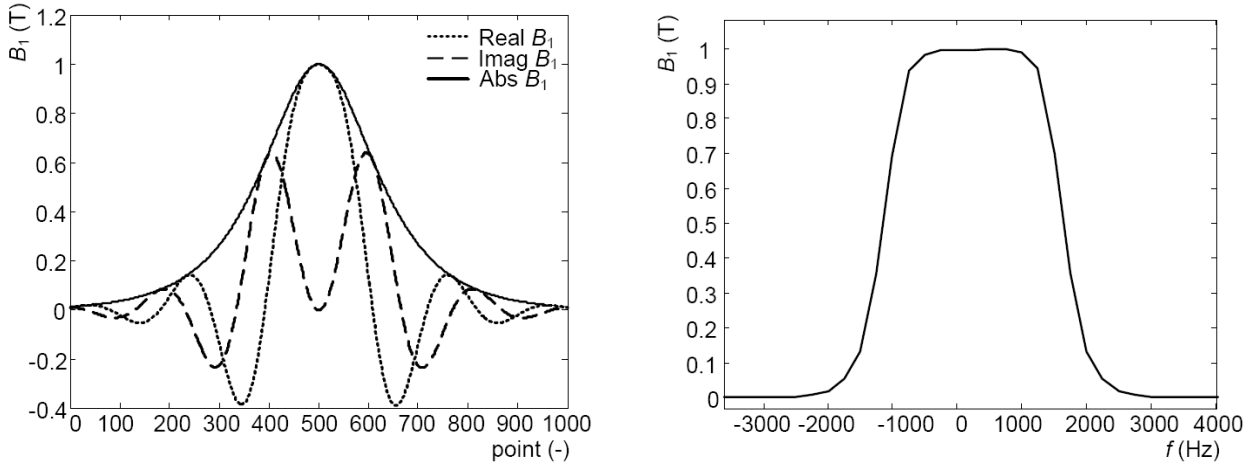


Figure 2: Amplitude-modulated RF impulse.

frequency spectrum are shown in Fig. 2. A flip angle of 90° obtained by integrating the RF impulse envelope (1) is

$$\alpha_{\mathrm{m90}} = \gamma \cdot B_{\mathrm{m90}} \cdot 4.3 \cdot 10^{-4}\mathrm{Ts}\,. \tag{4}$$

Magnitude of the magnetic field induction B_1, generated by this RF impulse, can be specified by comparing with induction B_1 generated by RF impulse of rectangular shape with known duration of 90° impulse. The used probe is of 90° impulse duration $t_{90} = 80\,\mu$s for rectangular shape of RF impulse with attenuation in RF channel 5 dB. The RF magnetic filed induction magnitude calculated according expression (2) is $B_{1,90} = 73.4\,\mu$T. This impulse flip angle is

$$\alpha_{90} = \gamma \cdot 5.88 \cdot 10^{-9}\mathrm{Ts}\,. \tag{5}$$

From Equation (4) and (5), the magnitude of magnetic induction of RF field $B_{\mathrm{m,90}}$ generated by RF impulse of 4 ms length in order to obtain 90° impulse can be evaluated

$$B_0 = \frac{5.88 \cdot 10^{-9}}{4.3 \cdot 10^{-4}} = 13.6\,\mu\mathrm{T}\,. \tag{6}$$

This lower level of the RF magnetic field is experimentally achieved by reduced magnitude of the RF power realized by an attenuation adjustment in a RF trace to value 21 dB.

The maximum amplitude $B_{\mathrm{m,90}} = 13.6\,\mu$T in image obtained by amplitude-modulated RF impulse of shape shown in Fig. 2 will be for the flip angle $\alpha = 90°$. For lower flip angles the maximum amplitude of induction B_1 will decline in correspondence with expression (3).

When mapping RF filed induction B_1 the lower flip angles are used in advance to sharp contrast for present inhomogeneity ΔB_1. To obtain the maximum contrast k in an image, the optimum size of the flip angle was determined.

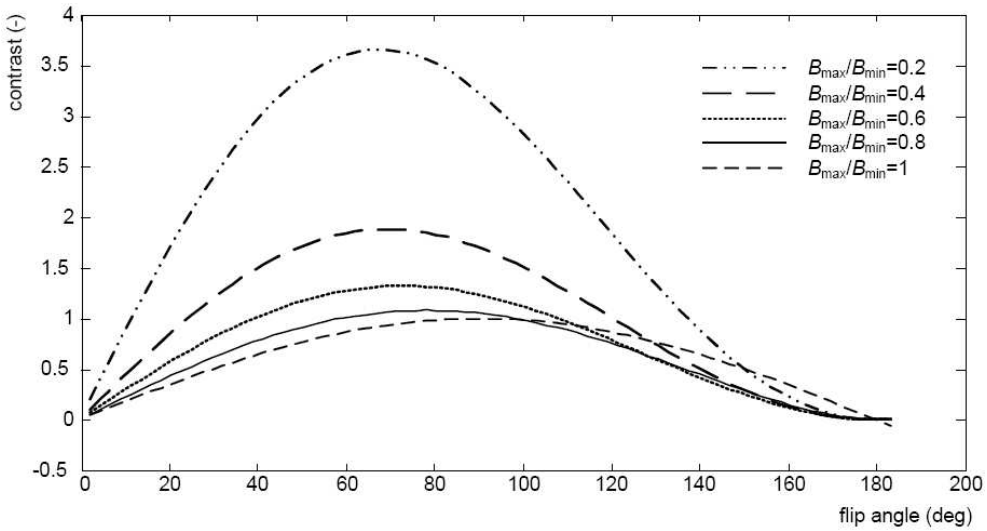


Figure 3: Dependence of the contrast k on the flip angle size.

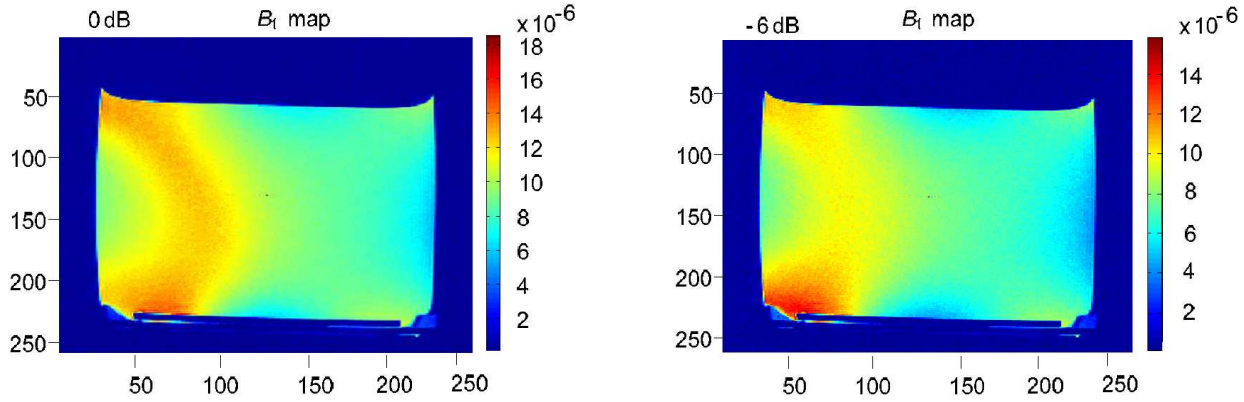


Figure 4: Amplitude-modulated RF impulse.

In a map of the RF field induction, scanned by MR technique, the dependence of the contrast k on the flip angle size is presented in Fig. 3. It can be seen that the optimum size of the flip angle is 66°.

The map the magnetic filed induction B_1 in the working space of MR tomograph was verified by measurement of MR image and by calculation. The used specimen was cubic vessel of $45 \times 45 \times 45$ mm dimension, filled by solution of H_2O, Ni_2SO_4, NaCl for reduction relaxation times T_1 (to 150 ms). This fact markedly shorts the time needed for the measurement.

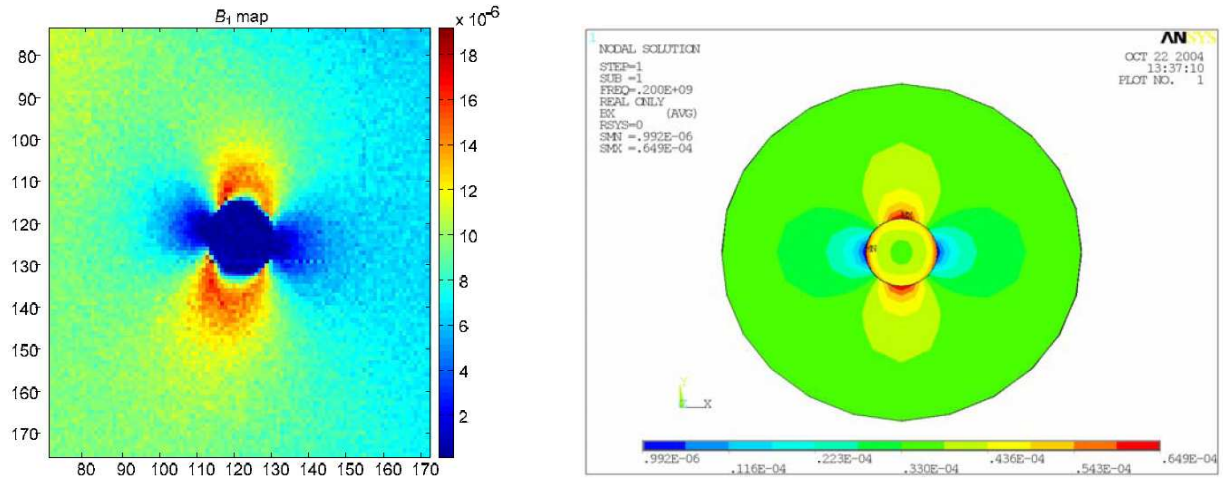
The MR images of dimension 60×60 mm (256×256 pixels) were measured by the spin echo (SE) method with the different flip angles. The variations of the flip angles form $\beta = 135°$ to $\beta = 30°$ were reached by the RF transmitter power changing with attenuation from 6 dB to -9 dB. The obtained images were normalized to the image maximum amplitude, corresponding to 90° exciting impulse. The amplitudes of the normalized images were converted into the RF magnetic field induction B_1 by use of

$$B_1 = B_{1,90} \arcsin\left(\frac{M}{M_0}\right) . \tag{7}$$

The obtained maps of field B_1 are depicted in Fig. 4. The optimum map is obtained for RF power attenuation -6 dB, in which the contrast k is maximum and the noise is minimum.

From the Fig. 4, it is obvious that for the plain orthogonal to axis and in the middle of the RF coil the field inductions are $B_{\max} = 13$ mT and $B_{\min} = 5$ mT. From this fact follows that the flip angle can for the optimum configuration rise up to 150°.

Using the method described above, the influence of the eddy current induced in the conducting specimens under measurement on the MR image deformation was examined. The MR images of 60×60 mm in dimension (256×256 pixels) were measured by the spin echo (SE) method with 66° flip angles. The specimen was a rod of Au-Pt-Ag alloy, used for dental implants. The alloy



(a) (b)

Figure 5: Map of RF magnetic field B_1 in the vicinity of the Safibond Bio rod a) measured b) simulated.

is produced by the Safina Company under the trade mark Safibond Bio. Electrical conductivity and susceptibility of this alloy were measured and established, the results are $\rho = 5.37\,\mathrm{Sm/m^2}$, $\chi = -24 \cdot 10^{-6}$. Fig. 5 shows the map of field B_1 in the vicinity of the specimen being measured.

To verify the measured map the simulation calculation using the 3D finite elements method was performed for the same configuration as in the measurement. In the homogenous RF magnetic field of 200 MHz frequency an area was created of 4 mm in diameter and 20 mm in length, whose conductivity and magnetic susceptibility were identical to those of the specimen measured. The results of the simulation are given in Fig. 5(b). It can be seen that both maps, the measured one and the simulated one, are identical.

3. CONCLUSION

The described method enables measurement of the RF magnetic field map inside the scanning coil of the MR tomograph. By the change of the measured slice location it is possible to measure 3 dimensional maps of the magnetic filed. The optimum flip angle for obtaining the maximum contrast of field B_1 and minimum disturbing noise in image $\alpha = 66°$. The method described enables measuring the RF magnetic field map inside the scanning coil of the MR tomographic scanner. By changing the location of the slice measured it is possible to measure 3D maps of the magnetic filed. The optimum flip angle for obtaining the maximum contrast of field B_1 and minimum disturbing noise in the image is $\alpha = 66°$.

Using the same method the map of B_1 magnetic field in the vicinity of a specimen of the material employed for dental implants was established. The method described is suitable for studying the deformation of RF magnetic field and MR images caused by eddy currents in specimens made of conducting materials.

ACKNOWLEDGMENT

The paper was prepared within the framework of B208130603 project of the Grant Agency of the Academy of Sciences of the Czech Republic and with the support of by the Grant Agency of the Ministry of Health of the Czech Republic IGA MZCR 8110-3/2004.

REFERENCES

1. Hoult, D. I., *Concepts Magn. Reson.*, No. 12, 173–187, 2000.
2. Kreis, R., et al., *J. Magn. Reson.*, No. 149, 245–250, 2001.
3. Helms, G., *NMR Biomed*, No. 13, 398–406, 2000.
4. Liang, Z. P. and P. C. Lauterbur, *Principles of Magnetic Resonance Imaging*, IEEE Press, New York, 2000.
5. Dahlqvist, O., L. Cohen, E. Lund, and P. Lundberg, "Absolute quantification of 31P muscle MRS using B1-field mapping," *Proc. Intl. Soc. Mag. Reson. Med.*, No. 13, 2768, 2005.

Temporal Analysis of Visual Search Task by Transcranial Magnetic Stimulation

Sheng Ge[1], **Shoogo Ueno**[2], **and Keiji Iramina**[1]
[1]Graduate School of Information Science and Electrical Engineering, Kyushu University, Japan
[2]Faculty of Engineering, Engineering, Kyushu University, Japan

Abstract— In this study, we examined the temporal aspects of the right posterior parietal cortex in easy feature "pop-out" visual search using transcranial magnetic stimulation (TMS). The transcranial magnetic stimulations were applied over the right posterior parietal cortex of subjects. Subjects received 4 tests which the TMS onset times were set as 100, 150, 200 and 250 msec after visual stimulus presentation. We found that, when SOA = 150 msec, compared to no-TMS condition, there was a significant elevation in response time when the TMS pluses were applied. However for the other SOA cases, there was no significant difference between TMS and no-TMS conditions. Therefore, we considered that "pop-out" visual search was processed in the right posterior parietal cortex at about 150 msec after stimulus present.

1. INTRODUCTION

The use of transcranial magnetic stimulation (TMS) in the investigation of neurological deficits provides an important method for human cognitive processes. Visual search, as a traditionally visual neglect sensitive measure, was studied by many researches. Much is already known about the involvement of right posterior parietal cortex [1, 2]. However, the study of temporal aspect of the posterior parietal cortex in visual search was not sufficient. In this study, to examine the temporal aspect of the posterior parietal cortex in "pop-out" visual search using TMS, we used different TMS stimulus onset asynchrony (SOA) and measured the visual search reaction times. The relationship between the SOA and reaction time was investigated.

2. MATERIALS AND METHODS

TMS Equipment:

The stimulator was a MagStim Super Rapid Stimulator (Magstim comp., Whitland, UK). Stimulus strength was set as the subjects' individual threshold for the motor evoked potential which under 55% of the maximum output. A figure-of-eight 70 mm coil was used.

TMS Stimuli:

2 pulses (20 ms interval) transcranial magnetic stimulations were applied over the right posterior parietal cortex of subjects.

Visual Search Task:

An easy feature "pop-out" search task was used in this study (Fig. 1). The target was a black backslash, and the distractor was black slash. The background was always gray. All the experiments were executed in a darkroom. The subject's head and sagittal midline was aligned with the centre

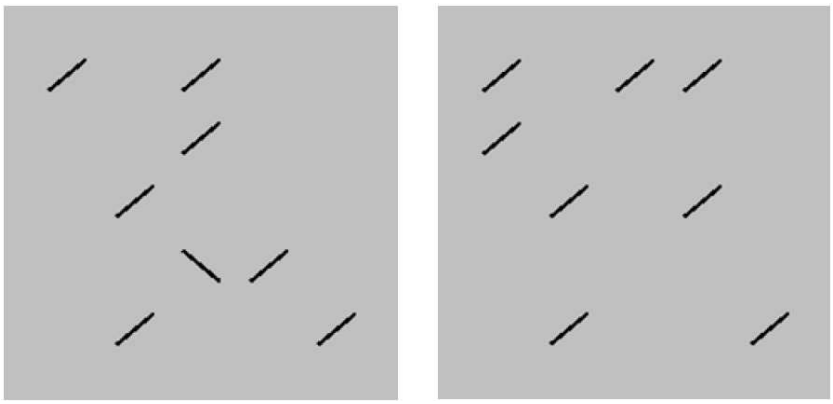


Figure 1: Examples of visual search stimuli.

of the monitor, and their head position was controlled by a chinrest. A visual search stimulus was

consisted of 8 items, which were presented on a 5.7×5.7 mm square range ($4.7° \times 4.7°$) on the center of PC monitor at a distance of 70 cm from the subject. The square range was divided into a 5 column$\times$ 5 row array of 25 virtual boxes. On any trial, each target or distractor could appear randomly in any one of 8 of these boxes. The target was present on 50% of trials and the target was unique among distractors.

Subjects:

5 subject, aged 21–31 years, all right hand, one female, four males.

Procedure:

The time sequence of experiment is shown in Fig. 2. Each trial was preceded by a central fixation cross ($0.9° \times 0.9°$) for 1500 msec, followed immediately by the visual search stimuli, which would be presented for 1500 msec. Subjects were asked to respond as quickly and as accurately as possible on a mouse button to indicate the presence or absence of the target (left button for target present and right button for target absent). The time from the visual search stimuli presentation till the button click was recorded as the response time. Transcranial magnetic stimulations were applied over the right PPC of subject at different time intervals after the visual search stimuli presentation. These time intervals were called as TMS stimulus onset asynchronies (SOA).

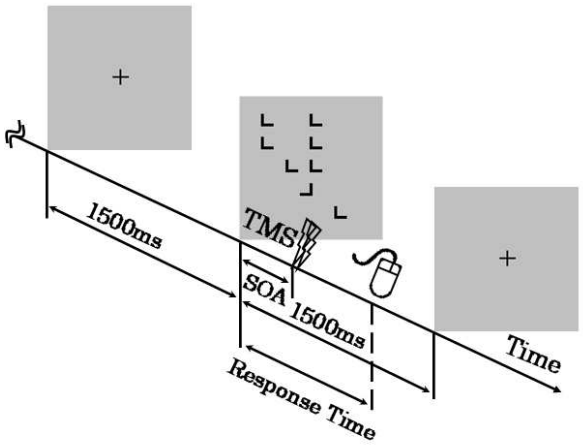


Figure 2: The time sequence of experiment.

Each subject received 60 trials for 2 times for TMS and no-TMS conditions respectively as one test. Subjects received 4 tests which the TMS stimulus onset asynchronies (SOA) were set as 100, 150, 200 and 250 msec after visual search stimuli presentation.

3. RESULTS AND DISSCUSSION

The average of target-present response times of each subject in the TMS condition was normalized to the no-TMS condition (set as the baseline = 1). The subtraction between the average of normalized response times of all the subjects and the baseline was taken to demonstrate the TMS effect.

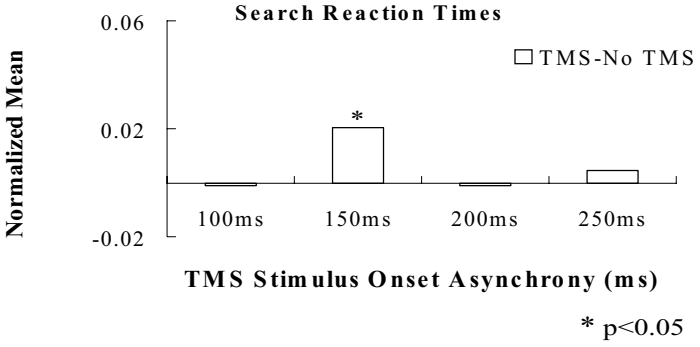


Figure 3: The target-present response times for each SOA (normalized).

The subtraction between the average of normalized response times of all the subjects and the baseline is shown in Fig. 3. In order to investigate the influence of TMS effect, one-way analysis

of variance (ANOVA) was used to analyze the difference between TMS and no-TMS conditions. We found that, when SOA = 150 msec, compared to no-TMS condition, there was a significant elevation ($p < 0.05$) in response times when the TMS pluses were applied. However for the other SOA cases, there was no significant difference between no-TMS and TMS conditions.

Corbetta et al.'s PET study reported activation of the PPC during feature conjunction search, but not during easy visual searches for single features [3]. Ashbridge et al.'s TMS study reported that TMS had detrimental effect on the performance of conjunction search but not on the feature search [2]. Whereas, both neuroimaging in humans [4] and single-unit recordings in monkeys [5] have revealed that the PPC subregions controlling spatial selection are also implicated in the selection of nonspatial features, suggesting that the involvement of the PPC in the visual search may not be binding-specific but rather reflect more general attention mechanisms. Furthermore Leonards et al. [6] and Donner et al. [7] also reported the PPC activations during the feature search task.

In the present study, when SOA = 150 msec, compared to the no-TMS condition, a significant elevation in response time was measured for easy feature search. Based on the past researches [2] and present study, it seems reasonable to support that the right PPC plays a dominant role not only in the conjunction search but also in the "pop-out" visual searches.

Furthermore, since a significant elevation in response time was only measured when SOA = 150 msec, we considered that "pop-out" visual search was processed in the right posterior parietal cortex at about 150 ms after stimulus presentation.

This TMS study opens up several new possibilities for understanding not only the role of the PPC in visual search but also the temporal aspect of the PPC involved in the "pop-out" visual search. The contribution to theories about the visual search dynamics is expected.

REFERENCES

1. Ellison, A., I. Schindler, L. L. Pattison, and A. D. Milner, "An exploration of the role of the superior temporal gyrus in visual search and spatial perception using TMS," *Brain*, Vol. 127, 2307–2315, 2004.
2. Ashbridge, E., V. Walsh, and A. Cowey, "Temporal aspects of visual search studied by transcranial magnetic stimulation," *Neuropsychologia*, Vol. 35, 1121–1131, 1997.
3. Corbetta, M., G. L. Shulman, F. M. Miezin, and S. E. Petersen, "Superior parietal cortex activation during spatial attention shifts and visual feature conjunction," *Science*, Vol. 270, 802–805, 1995.
4. Wojciulik, E. and N. Kanwisher, "The generality of parietal involvement in visual attention," *Neuron*, Vol. 23, 747–764, 1999.
5. Sereno, A. B. and J. H. R. Maunsell, "Shape selectivity in primate lateral intraparietal cortex," *Nature*, Vol. 395, 500–503, 1998.
6. Leonards, U., S. Sunaert, P. van Hecke, and G. A. Orban, "Attention mechanisms in visual search — an fMRI study," *Journal of Cognitive Neuroscience*, Vol. 12, supplement 2, 61–75, 2000.
7. Donner, T. H., A. Kettermann, E. Diesch, F. Ostendorf, A. Villringer, and S. A. Brandt, "Visual feature and conjunction searches of equal difficulty engage only partially overlapping frontoparietal networks," *Neuroimage*, Vol. 15, No. 1, 16–25, 2002.

Excitation of Hypersound Due to Coupling with Space Charge Waves in GaN Films

V. V. Grimalsky[1], **S. V. Koshevaya**[2], **L. M. Gaggero-S.**[1], **and F. Diaz-A.**[2]
[1]Faculty of Sciences, Autonomous University of Morelos, Cuernavaca, ZP 62210, Mor., Mexico
[2]Autonomous University of Morelos, CIICAp, Cuernavaca, ZP 62210, Mor., Mexico

Abstract— The excitation of hypersonic waves ($f = 30 - 200\,\mathrm{GHz}$) due to coupling with amplified space charge waves in GaN films is investigated theoretically. An amplification of space charge waves due to the negative differential conductivity in GaN films placed onto a semi-infinite substrate is considered and possible spatial increments are calculated. The amplified space charge waves can excite hypersonic waves at the same frequency due to piezoeffect and deformation potential mechanisms. It is demonstrated that the piezoeffect seems very effective for resonant excitation of hypersonic waves in the case of a full mechanic contact of GaN film and a non-piezoelectic substrate.

1. INTRODUCTION

There exists a problem of excitation of hypersound at the microwave frequencies $f > 10\,\mathrm{GHz}$ where the traditional methods, like using comb-like transducers, are not effective [1]. A possible solution of this problem is the resonant coupling of acoustic waves with the microwave electric field of space charge waves (SCW) in materials possessing negative differential conductivity (NDC) (like GaAs) [2, 3]. Namely, under propagation in the bias electric field higher than the critical value for observing negative differential conductivity, the SCW is subject to amplification, and its microwave electric field can achieve high values. In turn, due to piezoeffect, deformation potential, or electrostriction, this microwave electric field excites hypersonic acoustic waves (AW). In [2, 3] it was demonstrated that this excitation is of a resonant character with respect to the frequency and the thickness of the GaAs film. But the critical value of bias electric field in GaAs is $E_c = 3.5\,\mathrm{kV/cm}$ that limits the maximal values of the microwave electric field of SCW and output values of amplitudes of excited AW. Also, the frequency range of amplification of SCW in GaAs films is $f < 50\,\mathrm{GHz}$. To excite the powerful hypersonic AW at higher frequencies $f > 50\,\mathrm{GHz}$, it is rather better to use new materials possessing negative differential conductivity at higher frequencies $f = 100 - 300\,\mathrm{GHz}$, like gallium nitride GaN [4–7]. The attracting properties of GaN are: 1) a high critical bias field $E_c \sim 100\,\mathrm{kV/cm}$; 2) extended frequency range for observing NDC $f \leq 500\,\mathrm{GHz}$; 3) high values of piezoelectric constants.

The present paper is devoted to theoretical investigations of excitation of hypersonic AW in GaN films of a sub-micron thickness placed onto a non-piezoelectric substrate. The spatial increments of SCW due to NDC have been calculated. An amplification of SCW is possible up to the frequencies $f \sim 200 - 300\,\mathrm{GHz}$. The thin GaN film plays the role of a transducer of AW. The case of the full acoustic contact between GaN film and a substrate is considered. It has been demonstrated that the excitation of AW has a resonant character and the intensities of AW may reach the values of $1\,\mathrm{W/cm^2}$.

2. AMPLIFICATION OF SPACE CHARGE WAVES IN A THIN GAN FILM

Consider n-GaN film of a sub-micron thickness placed onto a semi-infinite dielectric substrate, see Fig. 1. The bias electric field is directed along OZ-axis, the SCW are excited by an input antenna I and propagate in Z-direction, too. The dynamics of SCW is described by the equations of motion of electrons jointly with the Poisson equation for the electric field. Because we consider the frequency range $f < 200\,\mathrm{GHz}$, it is possible to use the simplest hydrodynamic diffusion-drift equation for the electron fluid:

$$\begin{aligned} &\frac{\partial n}{\partial t} + div(\vec{v}(E)n - D_n\nabla n) = 0, \qquad \vec{v} = \mu(|E|)\vec{E}; \\ &div(\varepsilon_0\varepsilon(x)\nabla\varphi) = -e(n - n_0), \qquad \vec{E} = -\nabla\varphi + E_0 \end{aligned} \tag{1}$$

Here n is the electron concentration, φ is the electric potential of the alternative field, $\boldsymbol{v}$ is the electron velocity, n_0 is the equilibrium electron concentration (equal to the donor one), D_n is the

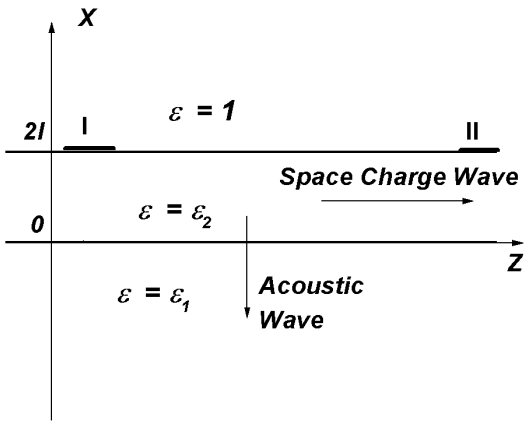


Figure 1: Geometry of the problem. The GaN film occupies the region $0 < x < 2l$, $x > 2l$ is vacuum, $x < 0$ is a semi-infinite substrate. I is an input antenna, II is output one.

diffusion coefficient, $\mu(E)$ is electron mobility, E_0 is the bias electric field. The data for GaN are taken from [8]. The coordinate frame is aligned along the crystalline axes. The lower indices 1, 2 relate to the substrate and the film, correspondingly.

The dependence of electron velocity on electric field for GaN is given in Fig. 2 [8].

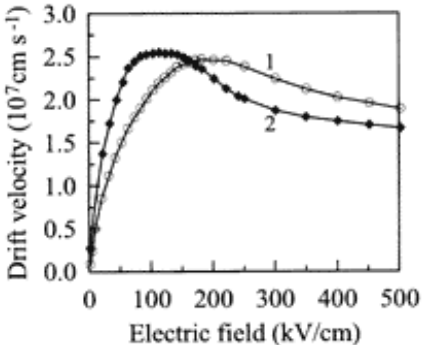


Figure 2: Dependencies of drift velocity on electric field for wurtzite (curve 1) and zinc blende (curve 2) GaN.

Amplification of SCW is investigated below. We use an approximation of a thin film where the electron concentration is assumed as two-dimensional. The linearized equations for the perturbation of electron concentration $\tilde{n}(z,t) = n - n_0$ ($|\tilde{n}| << n_0$) and for the electric potential φ are:

$$\begin{aligned} &\frac{\partial \tilde{n}}{\partial t} + v_0 \frac{\partial \tilde{n}}{\partial z} + n_0 \frac{dv}{dE} \frac{\partial \tilde{E}_z}{\partial z}|_{x=0} - D_n \frac{\partial^2 \tilde{n}}{\partial z^2} = 0; \qquad v_0 = v(E_0); \\ &\frac{\partial}{\partial x}\left(\varepsilon(x) \frac{\partial \varphi}{\partial x}\right) + \varepsilon(x) \frac{\partial^2 \varphi}{\partial z^2} = -\frac{e}{\varepsilon_0} 2\tilde{n} l \delta(x); \qquad \tilde{E}_z = -\frac{\partial \varphi}{\partial z} \end{aligned} \tag{2}$$

Consider the case of NDC: $dv/dE < 0$. For linear SCW we seek a solution of Eqs. (2) as: $\sim \exp(i(\omega t - kz))$. The dispersion equation for SCW is:

$$\omega - k v_0 - i \frac{2 n_0 l |e|}{\varepsilon_0 (1 + \varepsilon_1)} \frac{dv}{dE} k - i D_n k^2 = 0 \tag{3}$$

For spatial amplification of SCW (circular frequency ω is real, $k = k' + ik''$ is complex), in some frequency range one can obtain $k'' > 0$. We use the following parameters of GaN film: the electron concentration is $n_0 \approx 10^{17}\,\mathrm{cm}^{-3}$, the bias electric field is $E_0 \approx 150\,\mathrm{kV/cm}$, the diffusion coefficient is $D_n = 25\,\mathrm{cm}^2/\mathrm{s}$, the drift velocity is $v_0 = 2 \times 10^7\,\mathrm{cm/s}$, the thickness of the film is $2l = 0.1\,\mu\mathrm{m}$.

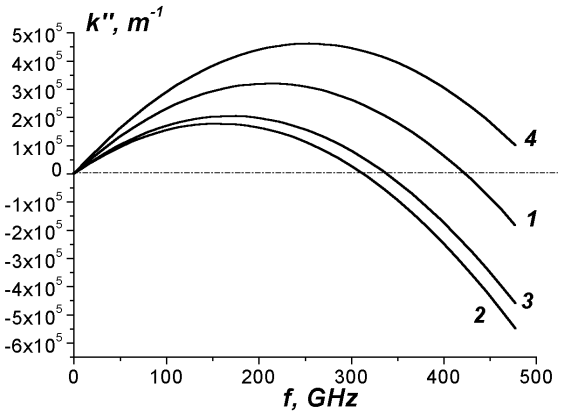


Figure 3: Spatial increment of instability $k''(\omega)$ of SCW in zinc blende GaN film. Curve 1 is for $E_0 = 150\,\mathrm{kV/cm}$, $n_0 = 10^{17}\,\mathrm{cm}^{-3}$, curve 2 is for $E_0 = 160\,\mathrm{kV/cm}$, $n_0 = 10^{17}\,\mathrm{cm}^{-3}$, curve 3 is for $E_0 = 150\,\mathrm{kV/cm}$, $n_0 = 0.8 \times 10^{17}\,\mathrm{cm}^{-3}$, curve 4 is for $E_0 = 150\,\mathrm{kV/cm}$, $n_0 = 1.2 \times 10^{17}\,\mathrm{cm}^{-3}$.

The dielectric permittivity of the substrate (like sapphire) is $\varepsilon_1 = 10$. The dependence of spatial increment of SCW k'' on frequency $f = \omega/2\pi$ is given in Fig. 3.

One can see that an amplification of SCW in GaN film occurs in a wide frequency range, and the maximal spatial increment is $k'' = 3 \times 10^3\,\mathrm{cm}^{-1}$ at the frequency $f = 150\,\mathrm{GHz}$. When compared with a case of the GaAs film, it is possible to observe an amplification of SCW in GaN films at essentially higher frequencies $f \geq 100\,\mathrm{GHz}$. To obtain an amplification of SCW $\approx 20\,\mathrm{dB}$, it is necessary to use the distance between input and output antennas of about $50\,\mu\mathrm{m}$. An amplification of SCW in GaN films should be realized in a pulse regime (of a duration $< 1\,\mu\mathrm{s}$), because of heating the GaN film. We do not consider here the frequency range $f > 200\,\mathrm{GHz}$, because an applicability of hydrodynamic equations is doubtful there [6].

3. EXCITATION OF HYPERSONIC ACOUSTIC WAVES

The microwave electric field of amplified SCW in the GaN film can excite the hypersonic AW at the same frequency due to piezoeffect. Because the sound velocity is two orders smaller than the velocity of SCW, the direction of propagation of excited AW is perpendicular to the surface of the film parallel to OX axis downwards. In the zinc-blende GaN (cubic symmetry), the nonzero piezoelectric module is $\beta = \beta_{3,12} = -0.4\,\mathrm{C/m}^2$. Therefore, the alternative electric field $\tilde{E}_z$ of SCW can excite the shear AW ($u = u_2$) that propagates along the OX axis. The equation of elasticity theory for mechanical displacement u takes the form:

$$\rho \frac{\partial^2 u}{\partial t^2} = \frac{\partial}{\partial x}\left(\rho s^2 \frac{\partial u}{\partial x}\right) + \frac{\partial}{\partial x}\left(\Gamma \frac{\partial^2 u}{\partial x \partial t}\right) \tag{4}$$

Here ρ is the density elastic medium, s is the shear acoustic velocity, Γ is viscosity. The expression for mechanical stress $\sigma = \sigma_{12}$ within the GaN film is:

$$\sigma = \rho_2 s_2^2 \frac{\partial u}{\partial x} + \beta \tilde{E}_z \tag{5}$$

Therefore, at the surfaces of the GaN film ($x = 0$ and $x = 2l$, see Fig. 1) there exists an exciting force due to piezoeffect, and the film serves as a transducer for AW. The excitation has a resonance character. To simplify the formulas, consider a case when the elastic properties of the film and substrate are the same. The expression of the amplitude of AW outgoing to the substrate is:

$$U = \frac{2i\beta\tilde{E}_z}{\rho s^2 K} sin^2(Kl) \exp(-2iKl); \qquad K = \frac{\omega}{s}\left(1 + \frac{i\omega\Gamma}{\rho s^2}\right)^{-1/2} \tag{6}$$

The effective excitation of AW occurs when the following resonant condition takes place:

$$Kl \approx \frac{\pi}{2}(2m+1), \qquad m = 0, 1, 2 \ldots \tag{7}$$

When the acoustic properties of the film and the substrate are different, the resonant condition (7) is valid, but the expression for the amplitude of the outgoing wave is more complicated. Note that at the frequency $f = 100\,\mathrm{GHz}$, the dissipation of shear AW is $|K''| \sim 500\,\mathrm{cm}^{-1}$ [1], so the condition $|K''| \sim 3 \times 10^{-3} << 1$ is satisfied. For the amplitude of alternative electric field of SCW $\tilde{E}_z = 10\,\mathrm{kV/cm}$, the output intensity of AW $P_a = (1/2)\rho s^3 |KU|^2$ can reach the values $P_a \sim 1\,\mathrm{W/cm}^2$ in the resonant case.

The excitation of AW has been considered here for the GaN film of the zinc blende structure. GaN films of the wurtzite structure are piezoactive too, and also can be used as transducers of AW.

An excitation of AW can also occur due to the deformation potential. The last interaction seems weaker than one due to piezoeffect. Our calculations have demonstrated that the output intensity is 10 times smaller, but the longitudinal AW (of a polarization u_1) can be excited there.

ACKNOWLEDGMENT

This work was supported by CONACyT, Mexico (Project 48955).

REFERENCES

1. Royer, D. and E. Dieulessaint, *Elastic Waves in Solids*, Vol. 2, Springer, New York, 2000.
2. Grimalsky, V., E. Gutierrez-D., A. Garcia-B., and S. Koshevaya, "Resonant excitation of microwave acoustic modes in n-GaAs film," in: *Proceedings of International Conference on Microelectronics ICM-2004*, No. 66, 4, Tunis, Tunisia, Dec. 2004.
3. Grimalsky, V., E. Gutierrez-D., A. Garcia-B., and S. Koshevaya, "Resonant excitation of microwave acoustic modes in n-GaAs," *Microelectr. Journ.*, Vol. 37, No. 3, 395–403, 2006.
4. Pearton, S. J., J. C. Zolper, R. J. Shul, and F. Ren, "GaN: processing, defects, and devices," *Journ. Appl. Phys.*, Vol. 86, No. 1, 1–79, 1999.
5. Jain, S. C., M. Willander, J. Narayan, and R. van Overstraeten, "III-nitrides: growth, characterization, and properties," *Journ. Appl. Phys.*, Vol. 87, No. 3, 965–1006, 2000.
6. Gruzhinskis, V., P. Shiktorov, E. Starikov, and J. H. Zhao, "Comparative study of 200–300 GHz microwave power generation in GaN TEDs by the Monte Carlo technique," *Semicond. Sci. Techn.*, Vol. 16, No. 8, 798–805, 2001.
7. Levinshtein, M. E., S. L. Rumyantsev, and M. S. Shur, *Properties of Advanced Semiconductor Materials: GaN, AlN, InN*, Wiley-Interscience, New York, 2001.
8. Internet site: http://www.ioffe.ru/SVA/NSM/Semicond/GaN/

Electron Spectrum of Single n-type δ-doped Quantum Wells in Si

I. Rodriguez-Vargas, L. M. Gaggero-Sager, V. V. Grimalsky
M. E. Mora-Ramos, and R. Pérez-Alvarez

Facultad de Ciencias, Universidad Autónoma del Estado de Morelos
Av. Universidad 1001, Col. Chamilpa 62210, Cuernavaca, Morelos, México

Abstract— We present the electron subband levels of single n-type δ-doped quantum wells in Si. The Thomas-Fermi and effective mass theories were used for the description of the conduction band bending and the subband level calculations, respectively. The energy levels, potential depth and subband occupancies are analyzed as a function of the impurity density. The possible trends of the transport properties based on the subband occupancy results are given. Our results are in good agreement with available experimental reports.

The advancement in growth and doping techniques allows the fabrication of semiconducting structures with very sharp doping profiles within a few atomic layers: the so-called δ-doping. Specifically, in the case of Si -with donor impurities- this advancement comes true in recent days [1–5], mainly due to the difficulties to incorporate impurities under accurately controlled conditions during the *MBE* growth [6, 7].

Due to the technological importance of this material and the particular properties of the δ-doped systems, they appear to be of interest for application in the electronic device industry as well as for basic investigation as [8–13].

In this work we present the results obtained for the electron spectrum and the subband occupancies in n-type δ-doped quantum wells in Si for different doping densities. Both the Thomas-Fermi (TF) and the effective mass approximations have been applied in order to obtain the conduction band potential and the subband level structure [14].

For a single δ-doped (SDD) quantum well, a direct relation between the density $n(z)$ and the Hartree potential $V_H(z)$ can be obtained in the one-dimensional TF approach within the Local Density approximation (LDA) [14]. It is written as:

$$n_{au}(z) = \frac{1}{3\pi^2}\left(\mu^* - V_H^*(z)\right)^{3/2}, \tag{1}$$

where $V_H^* = V_H/R_y^*$, $\mu^* = \mu/R_y^*$ are given in units of the effective Bohr radius and effective Rydberg, $a_0^* = \frac{\epsilon_r \hbar}{m^* e^2}$ and $R_y^* = \frac{e^2}{2\epsilon_r a_0^*}$. This particular scheme of effective units will be used throughout the work.

By including Eq. 1 in the corresponding Poisson equation, an explicit expression for the Hartree potential is derived. It has the form;

$$V_H^*(z) - \mu^* = -\frac{\beta^2}{\left(\beta|z| + z_0\right)^4}, \tag{2}$$

with $\beta = \frac{2}{15\pi}$ and $z_0 = \left(\frac{\beta^3}{\pi n_{2D}^{au}}\right)^{1/5}$.

In the framework of the LDA, the exchange-correlation potential for an electron gas can be written as [14]

$$V_x^*(z) = -\left[1 + \frac{0.7734 r_s}{21}\ln\left(1 + \frac{21}{r_s}\right)\right]\left(\frac{2}{\pi a r_s}\right), \tag{3}$$

where $r_s = \left(4\pi n_{au}(z)/3\right)^{-1/3}$ and $\alpha = (9\pi/4)^{-1/3}$.

Using the relation between $n_{au}(z)$ and $V_H^*(z)$ it is possible to write the total potential $V^* = V_H^* + V_x^*$ as

$$V^*(z) = -\frac{\beta^2}{\left(\beta|z| + z_0\right)^4} - c\left[\frac{\beta}{\left(\beta|z| + z_0\right)^2} + a\ln\left(1 + \frac{b\beta}{\left(\beta|z| + z_0\right)^2}\right)\right], \tag{4}$$

with $a = 0.7734/21\alpha$, $b = 21\alpha$ and $c = 2/\pi$. The latter equation summarized the model for the band bending profile. Instead of carrying out numerically troublesome self-consistent calculations, we simply solve two Schrödinger-like effective mass equations, thus obtaining the corresponding ladders of the longitudinal and transverse electron levels.

Due to the particular subband structure of the Si it is possible to write the subband occupancy as [8],

$$n_{jn} = \frac{m^*_{||j}}{\pi\hbar^2}\left[E_F - E_{jn}\right]\Theta\left(E_F - E_{jn}\right), \tag{5}$$

where $m^*_{||j} = m_t$ if $j = z, \bar{z}$ and $m^*_{||j} = (m_t m_l)^{1/2}$ if $j = x, \bar{x}, y, \bar{y}$. E_F represents the Fermi level and Θ the Heaviside step function.

The input parameters for the n-type δ-doped quantum wells are: $m^*_l = 0.9163m_0$, $m^*_t = 0.1905m_0$, $\epsilon_r = 12.5$ and $10^{12} \leq p_{2D} \leq 10^{14}\,\mathrm{cm}^{-2}$.

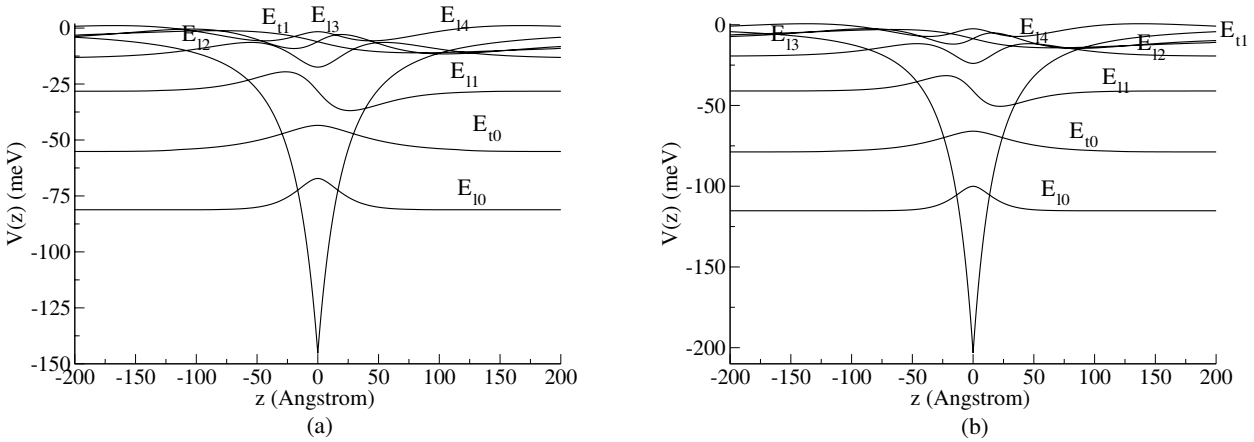


Figure 1: Potential profile and eigenfunctions of SDD QW's with (a) $n_{2D} = 8.0 \times 10^{12}\,\mathrm{cm}^{-2}$ and (b) $n_{2D} = 1.3 \times 10^{13}\,\mathrm{cm}^{-2}$.

In Fig. 1, the potential profile and wave functions of the single δ-doped quantum wells are depicted for two impurity densities, (a) $p_{2D} = 8 \times 10^{12}\,\mathrm{cm}^{-2}$ and (b) $p_{2D} = 13 \times 10^{12}\,\mathrm{cm}^{-2}$. It is seen that the main difference between these figures is the depth of the potential profile. This fact reflects in the number of confined states, which is a characteristic of the δ-doped systems.

Table 1: Subband occupancy $n_{jn}(10^{12}\,\mathrm{cm}^{-2})$ of single n-type δ-doped QW's for three different $n_{2D}(10^{12}\,\mathrm{cm}^{-2})$. $n_{jn}(TF)$ and $n_{jn}(S)$ represent our results (TF) and the selfconsistent ones obtained in Ref. [8] (S).

n_{2D}	n_{l0} (TF)	n_{l0} (S)	n_{t0} (TF)	n_{t0} (S)
8	4.85	3.33	3.14	4.22
13	7.13	5.08	5.86	6.83
65	29.96	20.04	38.03	31.04

In Table 1, the subband occupancies of the basic longitudinal and transverse levels for the different impurity densities are presented for both our results (TF) and the theoretical ones reported by Scolfaro et al. [8]. It is evident that there are discrepancies between both calculations. They can be explained by seeking the energy levels and the Fermi level reported by these authors, overall the lack of information about the determination of the latter. Taking into account the energies reported by those authors we are able to calculate the Fermi energy according to basic charge neutrality requirements. The recomputed values are: 60.3 meV, 186.5 meV and 210.2 meV for impurity densities of 8, 13 and 65 in units of $10^{12}\,\mathrm{cm}^{-2}$, respectively. These values clearly differ from those reported in that work, 85.7 meV, 100 meV and 327.8 meV, for the same impurity concentrations.

In [10] the quantum levels were experimentally determined through tunneling spectroscopy for $n_{2D} = 1.3 \times 10^{13}\,\mathrm{cm}^{-2}$. The dI/dV_g tunneling characteristic shows evidence -at positive V_g- for the occupied longitudinal ground level to be at $\approx 55\,\mathrm{meV}$. The same situation, with the same experimental techniques, [11] shows that the longitudinal basic level has an energy of 35 meV. So, taking the arithmetic average between these two findings, the experimental error is 10 meV. Therefore, the energy of the longitudinal basic level would be $45 \pm 10\,\mathrm{meV}$. For this very level our calculations give a value of 44.8 meV. It is important to mention that the inclusion of the exchange-correlation effects improves our latter result in 20% and that the error is about 1%, with respect to the experimental data.

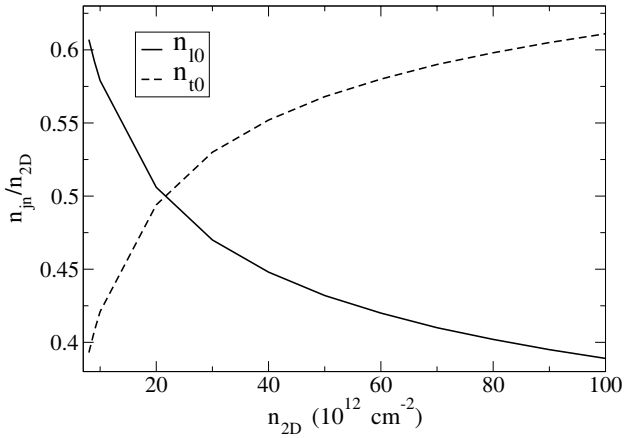


Figure 2: Relative subband density ($n_{jn} = n_{2D}$) versus the impurity density.

In Fig. 2, the relative subband density is shown as a function of the impurity concentration. As we can see from this figure, the basic longitudinal level has the 60% of the charge for the lowest impurity density depicted, while for the highest one occupy the 40%. The crossover between the longitudinal and the transverse ground occurs at $n_{2D} = 3 \times 10^{12}\,\mathrm{cm}^{-2}$. According to this behavior of the relative subband densities it is expected that the transport properties, specifically the mobility, should present an increase as functions of the impurity density. The reason for this is that the transverse ground level penetrates deeper in the potential barrier. Consequently, the scattering rate between impurities and carriers will be diminished thus leading to a higher mobility.

REFERENCES

1. Oberbeck, L., N. J. Curson, M. Y. Simmons, R. Brenner, A. R. Hamilton, S. R. Schofield, and R. G. Clark, "Encapsulation of phosphorus dopants in silicon for the fabrication of a quantum computer," *Appl. Phys. Lett.*, Vol. 81, No. 17, 3197–3199, 2002.
2. Oberbeck, L., N. J. Curson, T. Hallam, M. Y. Simmons, G. Bilger, and R. G. Clark, "Measurement of phosphorus segregation in silicon the atomic scale using scanning tunneling microscopy," *Appl. Phys. Lett.*, Vol. 85, No. 8, 1359–1361, 2004.
3. Leveque, P., J. S. Christensen, A. Y. Kuznetsov, B. G. Svensson, and A. N. Larsen, "In fluence of boron on radiation enhanced diffusion of antimony in delta-doped silicon," *J. Appl. Phys.*, Vol. 91, No. 7, 4073–4077, 2002.
4. Paliwal, V. K., A. G. Vedeshwar, and S. M. Shivaprasad, "Observation of novel superstructures at low Sb coverages on Si(111)-7 × 7 surface," *Surf. Sci. Lett.*, Vol. 513, No. 2, L397–L401, 2002.
5. Jeong, Y., M. Sakuraba, and J. Murota, "Epitaxial growth of N delta doped Si films on Si(100) by alternately supplied NH_3 and SiH_4," *Appl. Surf. Sci.*, Vol. 224, No. 1–4, 197–201, 2004.
6. Kiunke, W., E. Hammerl, I. Eisele, D. Schulze, and G. Gobsch, "Electrical transport between delta layers in silicon," *J. Appl. Phys.*, Vol. 72, No. 8, 3602–3605, 1992.
7. Jorke, H., H. Kibbel, F. Schäffer, A. Casel, H.-J. Herzog, and E. Kasper, "Properties of Si layers grown by molecular beam epitaxy at very low temperatures," *Appl. Phys. Lett.*, Vol. 54, No. 9, 819–821, 1989.
8. Scolfaro, L. M. R., D. Beliaev, R. Enderlein, and J. R. Leite, "Electronic structure of n-type δ-doping multiple layers and superlattices in silicon," *Phys. Rev. B*, Vol. 50, No. 12, 8699–8705, 1994.

9. Radamson, H. H., M. R. Sardela, Jr., O. Nur, M. Willander, B. E. Sernelius, W.-X. Ni, and G. V. Hansson, "Electron mobility enhancement in Si using doubly δ-doped layers," *Appl. Phys. Lett.*, Vol. 64, No. 14, 1842–1844, 1994.
10. Zeindl, H. P., T. Wegehaupt, I. Eisele, H. Oppolzer, H. Reisinger, G. Tempel, and F. Koch, "Growth and characterization of a delta-function doping layer in Si," *Appl. Phys. Lett.*, Vol. 50, No. 17, 1164–1166, 1987.
11. Li, H.-M., K.-F. Berggren, W.-X. Ni, B. E. Sernelius, M. Willander, and G. V. Hansson, "Tunneling current spectroscopy of electron subbands in n-type δ-doped silicon structures grown by molecular beam epitaxy," *J. Appl. Phys.*, Vol. 67, No. 4, 1962–1968, 1990.
12. Gaggero-Sager, L. M. and R. Perez-Alvarez, "A simple model for delta-doped filed-effect transistor electronic states," *J. Appl. Phys.*, Vol. 78, No. 7, 4566–4569, 1995.
13. Martinez-Orozco, J. C., L. M. Gaggero-Sager, and S. J. Vlaev, "A simple model for the differential capacitance profile in atomic layer doped field effect transistor (ALD-FET) in GaAs," *Mater. Sci. Eng. B*, Vol. 84, No. 3, 155–158, 2001.
14. Gaggero-Sager, L. M., "Exchange and correlation via functional of Thomas-Fermi in delta-doped quantum wells," *Modelling Simul. Mater. Sci. Eng.*, Vol. 9, No. 1, 1–5, 2001.

Shielding Design for Power Frequency Magnetic Field Produced by Substations

Haiyu Yu, **Li Hao**, **and Jiansheng Yuan**
The State Key Lab of Power Systems, Department of Electrical Engineering
Tsinghua University, Beijing 100084, China

Abstract— Some sensitive devices located upon substations may be interfered by the magnetic field caused by the transformers, cables and buses. Shielding is necessary in some cases to overcome this electromagnetic interference. The shielding design for a substation is presented in this paper, which is based on the calculation of magnetic field and shielding effectiveness. As a result, the best shielding effectiveness can reach 98.3%, and the value of the magnetic field in the location of the sensitive devices can be reduced to meet the demand of the electromagnetic environment.

1. INTRODUCTION

According to IEC 61000-4-8(1993) [1], the power frequency magnetic field environment is divided into five levels, which are 1, 3, 10, 30, 100 A/m. Different device is able to endure different level of magnetic field environment. CRT (Cathode-ray tube) displaying devices may be the most sensitive devices to the power frequency magnetic field. They can't normally display, e.g., the dithering of the screen, when the background magnetic field intensity reaches 1.5 A/m. Generally, if such sensitive devices just locate above underground substation with large capacity, they will be interfered by the high magnetic field.

To solve this electromagnetic interference problem, the magnetic shielding technique can be employed. Two kinds of shielding approaches to shielding the field of underground substation are analyzed in this paper. The first is laying a plate with high permeability or high conductivity on the floorboard of the up floor or as the ceil of the substation. The other is installing shells with high permeability or high conductivity around bus bars and cables. By calculating the shielding effect of these two approaches with different size, different material, and different position, the best shielding design is found. Considering the engineering practice, the second approach is further modified in structure.

2. MODEL OF THE SUBSTATION AND THE CARED REGION

A typical underground substation (as shown in Fig. 1) includes transformers, middle-voltage (MV) supply cables and bus bars, low-voltage (LV) distribution cables and bus bars, etc. Transformers are usually equipped by ferrite shell, so that they are not considered as the interference sources. Additionally, since the current of MV cables and bus bars is lower than that of LV cables and bus bars, MV cables and bus bars are not considered as the interference sources either. Consequently, only the LV distribution bus bars and cables are regarded as the critical sources in indoor MV/LV substation, with respect to the magnetic interference at power frequency.

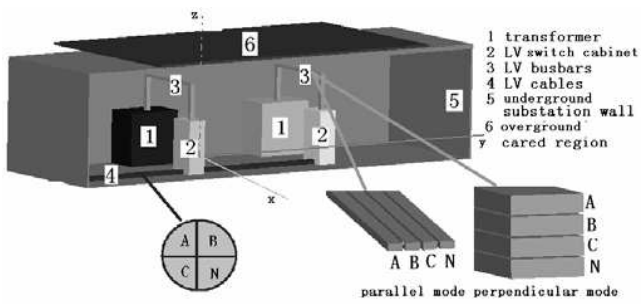


Figure 1: Structure of the underground substation and the cared region.

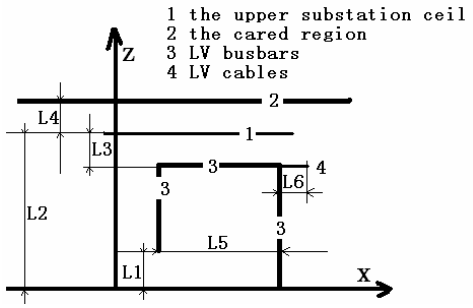


Figure 2: Location of bus bars, cables and the cared region.

In the substation considered in the paper there are 24 transformers distributing on the floor. Since they are located by making a couple as a group, only the bus bars and cables relative to a

couple of transformers are considered in the calculation of the magnetic field and the design of the shielding. A group of bus bars and cables as well as the cared region where the sensitive devices are located are shown in Fig. 2, in which the cables are in the direction of Y axis and located on the floor. The rated current in the bus bar is 3000 A.

The LV bus bars and the cables are of four conductors, which are phase ABC and the natural line. The four bus bars from the transformer to the low voltage switch cabinet can be arranged compactly in two modes, parallel or perpendicular to the floorboard of the substation. The four core-lines of cable from low voltage switch cabinet to the load are equally arranged in cables, i.e., they are located on the vertices of a square.

In the calculation model shown in Fig. 2, L1 to L5 are respectively 2300, 4899, 1000, 200, 2000, and 200 mm. The distance between two transformers in one couple is 108.9 mm, and that between low voltage switch cabinet and the load is 6800 mm. The center of bus bars is 6 mm to the floor, both for parallel and perpendicular modes. The center of cables is 20.67 mm to the floor, and the diameter of the square, on whose vertices the cables are located, is 35.91 mm.

The cared region is the up floor of the substation, which is the 2008 Olympic Games television program broadcast center. In the calculation, the plane on which the magnetic field is considered is of 5.1 m to the substation floor.

3. CALCULATION OF THE INTERFERENCE MAGNETIC FIELD

The interfering sources are the LV bus bars and cables. The magnetic field in the cared region can be calculated by the Biot-Savart's Law, which is as follows:

$$\vec{B}(x_p, y_p, z_p) = \sum_{i=1}^{N} \frac{\mu_0 I_i}{4\pi} \int_{li} \frac{d\vec{l} \times \overrightarrow{r_{i,p}}}{r_{i,p}^3} \tag{1}$$

where I_i is the current in line segment i of the bus bars and cables, and N is the total number of the line segments of bus bars and cables.

It is not difficult to calculate the field by (1). However, since the current in a phase is varying in sine model, and the phases are different in phase lines, the magnetic field at a position are varying and the location of the maximum value of the magnetic field is also varying. Therefore, the field at different time must be calculated to find the maximum value.

Usually, when the power system operates in symmetrical model, i.e., the magnitude of the currents in three phases is the same and the current in the natural line is zero, the interference magnetic field is less than that in the asymmetrical operations. Considering the symmetrical and asymmetrical operations, in the calculation the currents in three kinds of operations are taken into account, which are listed in Table 1.

Table 1: Magnitude and phase of current in three kinds of operations.

Index	Symmetrical Operation	20% [1]asymmetrical	100% [2]asymmetrical
Phase_A, Mag(kA), Angle(degree)	3.039, 0	2.533, 0	3.039, 0
Phase_B, Mag, Angle	3.039, −120	3.039, −120	0, 0
Phase_C, Mag, Angle	3.039, 120	3.039, 120	0, 0
Phase_N, Mag, Angle	0, 0	0.506, 0	3.039, 180

[1]The load of phase A is 20% larger than the rated load. [2]Phase B and C are broken or do not operate.

On the other hand, to get the maximum value of the interference magnetic field in the cared region, the position order of the cable phase lines and the bus bar phase line must be set properly, since the total field is the vector sun of those produced by every phase line. Otherwise, the maximum value may not be obtained.

Based on the structure data and the calculation approaches described above, the maximum interference magnetic field in the cared region is calculated, and the results are listed in Table 2.

From the results in Table 2 we can see that the interference magnetic field caused only by the cables in symmetrical operation is even larger than that in 20% and 100% asymmetrical operation. The reason is that the purpose is to get maximum value of the total field produced by cables and bus bars, but not that produced only by the cables. At the time when the field produced only

Table 2: Value of maximum flux density in three kinds of operations(μT).

Interference Sources	Symmetrical Operation	20% asymmetrical	100% asymmetrical
LV busbars[1]	2.65	2.85	4.95
LV busbars[2]	3.80	4.11	6.55
LV cables	1.54	1.42	1.26
LV busbars[1] and cables	3.57	3.70	5.32
LV busbars[2] and cables	3.97	4.25	6.60

[1]Bus bars are arranged in parallel mode. [2]Bus bars are in perpendicular mode.

by the cables arrives maximum value, the total field does not arrives its maximum value. The magnitude distribution of the total field in the cared region in symmetrical operation is shown in Fig. 3.

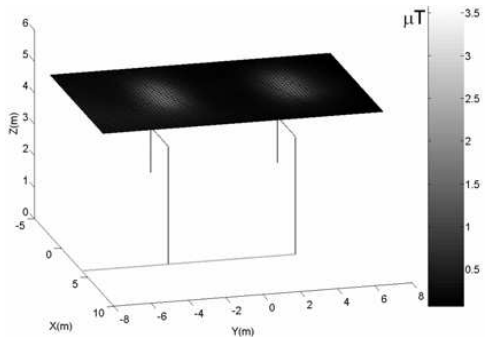


Figure 3: Magnitude distribution of total field in symmetrical operation.

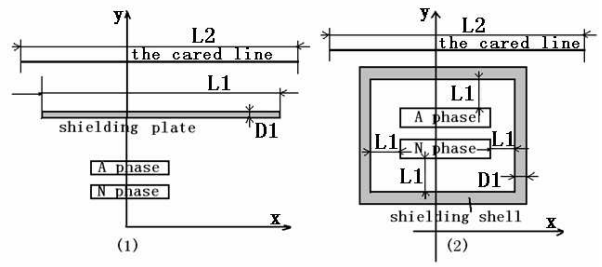


Figure 4: Two shielding schemes, (1) lay a shielding plate, (2) install a shielding shell.

The magnetic field caused by the cables are smaller than that by the bus bars, since the cables are further away from the cared region than the bus bars. The magnetic field caused by bus bars in parallel mode are smaller than that in perpendicular mode. Whatever in any case, the interference magnetic field is much higher than the value of environment level one and two. Therefore, the shielding must be employed. To design the shielding, the field caused by only bus bars in perpendicular mode in 100% asymmetrical mode is taken into account as a sample.

4. DESIGN OF SHIELDING SCHEMES

There are two schemes employed in the practice engineering to shield the power frequency magnetic field [2, 3]. The first is laying a plate with high permeability or high conductivity on the floorboard of the up floor or as the ceil of the substation. The other is installing shells with high permeability or high conductivity around bus bars and cables, as shown in Fig. 4. A square shell is adopted in the analysis below. The material of the shielding plate or shielding shell is as follows: aluminum with conductivity of 3.8E7 S/m and relative permeability of 1, steel with 1.026E7 S/m and 2000, and a fictitious material with conductivity of null and relative permeability of 10000, which is denoted to material 3 in the following.

To analyze the shielding effect, the magnetic field with and without the shielding should be calculated and compared. The shielding effect is defined as follows:

$$\delta = \frac{B_{\max\_N} - B_{\max\_Y}}{B_{\max\_N}} \times 100\%, \tag{2}$$

where B_{max_N} and B_{max_Y} are the maximum flux density in the cared region without and with the shielding.

With the dimensions shown in Fig. 4, we take two-dimensional model for the calculation of the field. The cared region is represented by a cared line as shown in Fig. 4. The calculation of the magnetic field with the shielding is conducted by Ansoft (a commercial electromagnetic analysis

software). The maximum values of the interference magnetic field and the shielding effects on the cared line are listed in Table 3 and Table 4 with different size of shielding frames respectively. The magnetic flux density distribution along the cared line without and with plate and shell shielding is shown in Fig. 5.

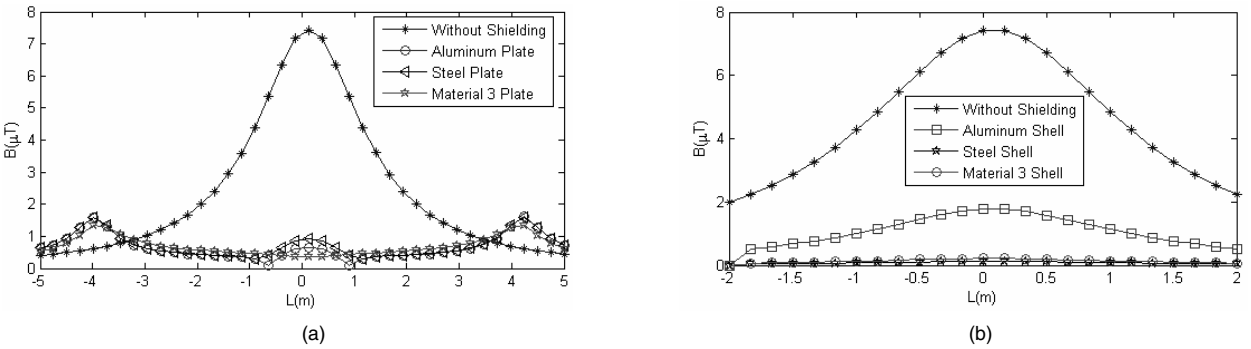


Figure 5: Distribution of the magnetic flux density along the cared line, (a) with the shielding plate of $L1 = 8\,\text{m}$ and $D1 = 2\,\text{mm}$ (b) with the shielding shell of $L1 = 0.02\,\text{m}$ and $D1 = 2\,\text{mm}$.

Table 3: Value of maximum flux density and the shielding effects on the cared line with shielding plate of different materials.

Material	Aluminum	Steel	Material3	Aluminum	Steel	Material3
L1(m)	8	8	8	2	2	2
D1(mm)	2	2	2	2	2	2
L2(m)	10	10	10	4	4	4
B_max(μT)	1.65	1.45	0.72	5.69	3.57	3.39
Effect(%)	77.9	80.6	90.4	23.8	52.2	54.6

Table 4: Value of maximum flux density and shielding effects on the cared line with shielding shell of different materials.

Material	Aluminum	Steel	Material3	Aluminum	Steel	Material3
L1(m)	1	1	1	0.02	0.02	0.02
D1(mm)	2	2	2	2	2	2
L2(m)	4	4	4	4	4	4
Bmax:μT	0.86	0.14	0.72	1.77	0.13	0.22
Effect(%)	88.5	98.1	90.4	76.3	98.3	97.1
L1(m)	0.02	0.02	0.02	0.01	0.01	0.01
D1(mm)	4	4	4	1	1	1
L2(m)	4	4	4	4	4	4
Bmax:μT	0.88	0.0023	0.11	5.13	0.66	0.23
Effect(%)	88.2	99.97	98.5	31.3	91.1	96.9

According to results listed in Table 3 and Table 4, we can see that the shielding effect is enormous. Considering the weight of the material and the shielding effect, the optimal design is setting a steel shielding shell with L1=0.02 m and D1=2 mm, by which the shielding effect is 98.3%, and the maximum magnetic flux density in the cared line is reduced to 0.13 μT. Obviously, the environment of the magnetic field can be managed to reach level 1, and the CRT devices can normally display.

5. FURTHER CONSIDERATION TO THE SHIELDING SHELL

The optimal shielding scheme is the closed square shell shaped by welding. However, considering the engineering practice, it is not easy to weld the shell and make the shell close properly. Suppose the shell is not welded and an aperture exists. To check the shielding effect of such kind shell, the model with an aperture on the bottom corner as shown in Fig. 6 is calculated.

With $L3 = 10\,\text{mm}$, and the other parameters are the same as those in the optimal shielding shell described above, the shielding effect is calculated in three different shielding materials. Maximum flux density and the shielding effect is 1.78 μT 76.2%, 0.37 μT 95.1%, 0.41 μT 94.5% in the cared line with aluminum, steel and Material 3 respectively. Fig. 7 gives the magnetic force lines when the steel shielding shell with an aperture is used. The distribution of the magnetic flux density along the cared line is shown in Fig. 8. With this structure, the shielding shell is easier to be put into practice, and the shielding effect is also enormous.

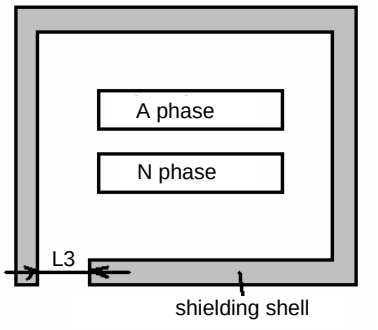


Figure 6: Shielding shell with an aperture.

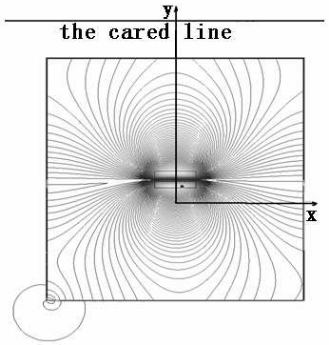


Figure 7: The magnetic force lines when the steel shielding shell with an aperture is used.

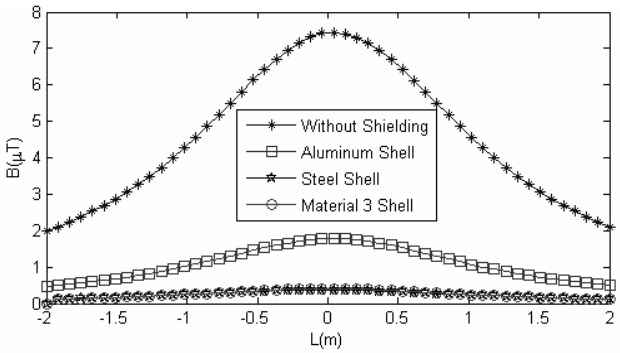


Figure 8: Distribution of the magnetic flux density along the cared line.

6. CONCLUSIONS

1. Generally, the interference of power frequency magnetic field with generic electrical devices can be neglected. However, the interference with the CRT displayers is remarkable.
2. When the sensitive devices such as CRT displayers are located on the over ground of the underground MV/LV substation, this interference effect is very serious, and the magnetic shielding measures must be adopted.
3. Comparing the effect of the shielding plate and the shielding shell with different materials, the ideal shielding scheme in engineering is steel shell, whose shielding effectiveness can react to 98%, and the maximum magnetic flux density in the cared region can be limited to 0.2 μT, which is much less than the value of magnetic environment level 1.
4. Considering the engineering practice, the shielding shell may not be welded, and on the bottom of which a small aperture can exist. With this structure the shielding effect can still reach to 95%.

ACKNOWLEDGMENT

This work was supported by the National Natural Science Foundation of China (No. 50677028), SRFDP (No. 20050003007) and Tsinghua Basic Research Foundation.

REFERENCES

1. IEC 61000-4-8:1993, *Electromagnetic Compatibility (EMC) — Part 4: Testing and Measurement Techniques — Section 8: Power Frequency Magnetic Field Test*, Basic EMC Publication.
2. Buccella, C., M. Feliziani, and A. Prudenzi, "Active shielding design for a MV/LV distribution transformer substation," *Electromagnetic Compatibility, 2002 3rd International Symposium*, 350–353, 21–24 May 2002.
3. Garzia, F. and A. Geri, "Active shielding design in full 3D space of indoor MV/LV substations using genetic algorithm optimization," *Electromagnetic Compatibility, 2003 IEEE International Symposium*, Vol. 1, 197–202, 18–22, Aug. 2003.

Analytical Design and Simulation Rescaling of Magnetically Insulated Transmission Lines

Xiangmin Jin and Jiansheng Yuan

The State Key Lab of Power Systems, Department of Electrical Engineering
Tsinghua University, Beijing 100084, China

Abstract— Analytical solution and numerical simulation are two effective approaches for analyzing and designing the vacuum magnetically insulated transmission lines (MITLs) in multi-terawatts applications. It is introduced in this paper a scheme for designing a tapered-electrode-radius MITL used in the linear induction accelerator utilizing the laminar flow model to match self-limited anode currents of each segment. This preliminary design is validated with particle-in-cell numerical simulations. In order to enhance the power transportation efficiency the MITL has to be rescaled with the help of particle-in-cell simulations, so that the MITL operates in the load-limited equilibria with smaller amount of space electron flow. The design based on the numerical simulation is presented in this paper. However, the modification will cause more simulations and complicated manufactures.

1. INTRODUCTION

In modem pulsed power systems due to the high power density, very high electric fields are produced at the metal surfaces in vacuum transmission lines. For example, a power density of 1 TW/cm^2 generates a 20MV/cm electric field. Generally, when the local electric field on a cathode surface exceeds 200 kV/cm, electrons explosively emitted from the cathode and form a plasma sheath near the surface [1]. As a result of the radial electric field, electrons emitted from the cathode plasma sheath are pulled toward the anode and shunt the high voltage gap. Consequently, ordinary transmission lines have an energy flux limit of 100 MW/cm^2. However, the magnetic fields generated by the currents flowing in the system generate forces perpendicular to the electric fields to confine the electrons. When the currents become sufficiently large, the self-magnetic field can keep these electrons flowing parallel to the cathode and greatly reduce the loss of electrons to the anode. Thus, magnetic insulation is established. The electron trajectories and currents of magnetic insulation in a coaxial geometry are shown in Fig. 1.

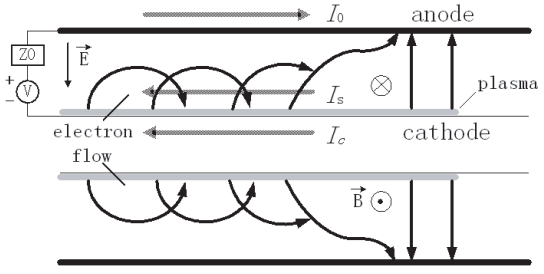


Figure 1: Schematics of currents and electron flow in an MITL.

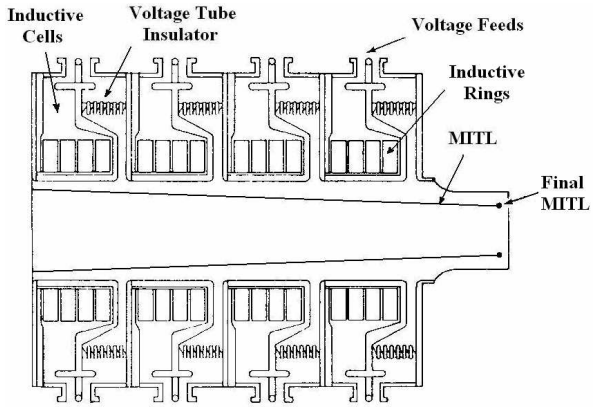


Figure 2: Schematic cross section of the voltage adder MITL.

2. NUMERICAL METHOD

During the past decades analytical models have been improved to give an understanding of the magnetically insulated transmission lines (MITLs) [2]. There are basically two types of analytical equilibrium models, laminar flow and quasi-laminar flow models. Both models assume the energy and canonical angular are conserved. The difference between the two models is the character of the electron trajectories in the sheath. In laminar flow model the gap voltage is increased on a

time scale long compared with the electron cyclotron period, thus the $\mathrm{d}E/\mathrm{d}t$ drift moves electrons onto laminar flow orbits parallel to the plane of the gap. In quasi-laminar flow model the voltage in the gap is applied instantaneously and the resulting orbits are cycloid-like between the cathode and the edge of the electron sheath.

Besides, electromagnetic particle-in-cell (PIC) as used for the coupled simulations of charged particles and their electromagnetic fields is also a useful method for analyzing MITLs [3]. PIC utilizes a set of computational macro-particles for representing the electron clouds in the MITLs, and numerically solves the coupled self-consistent solutions of the micro-particle motion equations and the Maxwell equations in the simulation space. Large-scale PIC simulations are sometimes used for designing a whole large system, but these generally are fairly expensive, and suffer from long turnaround times [4]. However, PIC simulations are quite valuable for developing or modifying the results original obtained by analytical models. Therefore, we describe here the design scheme combining the analytical models and PIC simulations.

3. DESIGN OF MITL USING EQUILIBRIUM MODELS

The analytical equilibrium models have been proven to be useful in the design of cylindrical MITLs with no radii changes [5]. Generally in practice, MITLs have radial changes in the axial direction to fit for the loads dimensions. As shown in simulations and analysis, in an MITL with geometrical discontinuity, although a large amount of electrons inject from the upstream of geometries discontinuity, most of them are absorbed by the cathode downstream and have been replaced by emitted electrons with an increase in electron density and electron cloud boundary [6]. Therefore, the injected electrons from upstream act ittle influence on the downstream. As a result, equilibrium models can still be used in design and analysis of MITLs with radii geometrically changes.

Here we use a voltage adder MITL installed in the linear induction accelerator (shown in Fig. 2) as an example to describe the design process [7]. The voltage adder MITL sums the voltages of the pulses delivered from the voltage feeds. The voltages are added to the right so that all arrive simultaneously at the load Z_L. The voltage adder has a structure with a constant outer anode radius and tapered inner cathode radii along the power-flow direction. Assume in order to drive a terminating load the adder MITL delivers a power with 750 kA and 20 MV. If the input voltages from the voltage feeds have the same peak value of 4 MV, so there will be 5 cavities on the voltage adder MITL. Besides, the outside radius of the anode cylinder is 38.1 cm which is defined by the system insulating stacks.

As we know that in an MITL that has no equilibrium electron loss, the anode current is constant along the length of the MITL. Consequently, operating in the efficient equilibrium the voltage adder MITL has a same anode current throughout every segment. Besides, assume that there is a final cylindrical MITL installed behind the voltage adder MITL which is sufficiently long to isolate the adder MITL from the terminating load, so when the MITL system operate in the equilibrium the anode current is determined by the self-limited current of the final MITL. Numerical simulations and measurements from many magnetically insulated devices have shown that the self-limited current tends to be near the minimum current on any voltage [8]. Accordingly, the anode current of the voltage adder MITL is the minimum current of the final MITL under its equilibrium voltage. The vacuum impedance Z_0 of each section of the voltage adder MITL can then be chosen so that the minimum anode current for each of the section voltage is the same. This choice allows matched self-limited flow for the local voltage and vacuum impedance all the way through the voltage adder MITL. Such an approach can only be used in the system which normally has a long final MITL; its current is determined by MITL itself but not the terminating load impedance. Determined by operation requests above the anode current all through each segment is 750 kA and the voltage adder MITL output voltage is 20 MV. Therefore, the local voltage of the five segments is respectively 4 MV, 8 MV, 12 MV, 16 MV and 20 MV. Because the voltage adder is negative-polarity (anode outside and cathode inside) and has a constant anode radius, the concept of matching self-limited current was applied to calculate the inner radii of each section of the voltage adder MITL.

With the initial parameters and laminar flow model equation for minimum anode current I_l [8]:

$$I_l = I_0 g \gamma_l^3 \ln\left[\gamma_l + \left(\gamma_l^2 - 1\right)^{1/2}\right] \tag{1}$$

where the $I_0 = 2\pi m_0 c/\mu_0 e \approx 8500A$, m_0 is the rest mass of the electron, c is the velocity of light, e is the charge of the electron and μ_0 is the vacuum permeability. $g = 1/\ln(r_a/r_c)$ is the cylindrical

geometrical factor where r_a and r_c are the anode radius and cathode radius. γ_l is the electron flow sheath boundary defined by

$$\gamma = 1 + eV/m_0c^2 \tag{2}$$

where the V is the voltage along the radial direction. Besides, γ_l has the relation with anode voltage γ_a as followed [8]:

$$\gamma_a = \gamma_l + \left(\gamma_l^2 - 1\right)^{3/2} \ln\left[\gamma_l + \left(\gamma_l^2 - 1\right)^{1/2}\right] \tag{3}$$

Hence for the given local voltage and required anode current for each segment, the corresponding cathode radii can be calculated using Eq. (1)∼(3), so that in self-limited equilibra each segment has the same minimum anode current. The anode current curves for all segment voltages of the laminar flow model are given in Fig. 3 in which the abscissa represents the electron sheath edge voltage. The triangles on curve are the locus of the minimum current of all the segments. The solution for all the segments cathode radii of is shown in Table 1.

Table 1: Design results of laminar flow model and PIC simulation.

MITL Segment	Segment Voltage (MV)	Model Cathode Radius (cm)	Rescaling Cathode Radius 1 (cm)	Rescaling Cathode Radius 2 (cm)	Rescaling Cathode Radius 3 (cm)	Actual Cathode Radius (cm)
D1	4	33.8	27	29	29	33.7
D2	8	30.6	26	28	27.5	30.5
D3	12	27.8	25	26	26	27.6
D4	16	25.3	24	25	24.5	25.1
D5	20	23	23	23	23	22.7

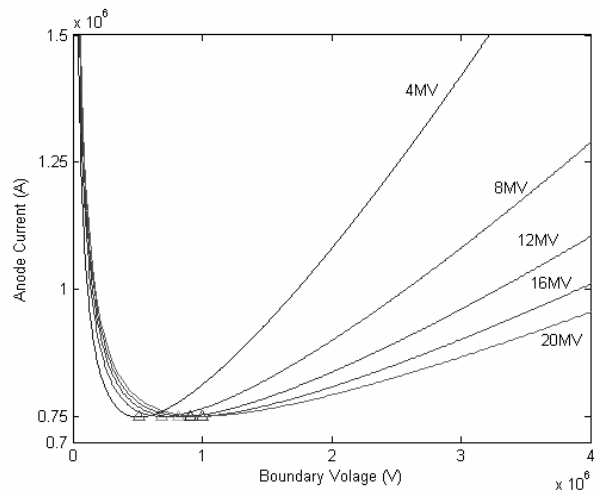


Figure 3: Anode currents curves and minimum equilibiua of MITL.

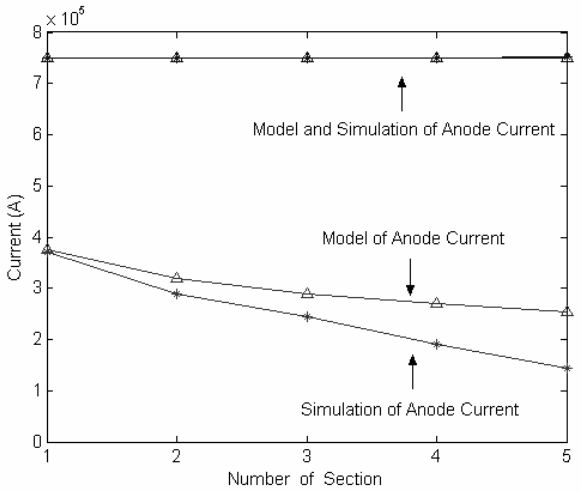


Figure 4: Anode and cathode currents of each segments of MITL.

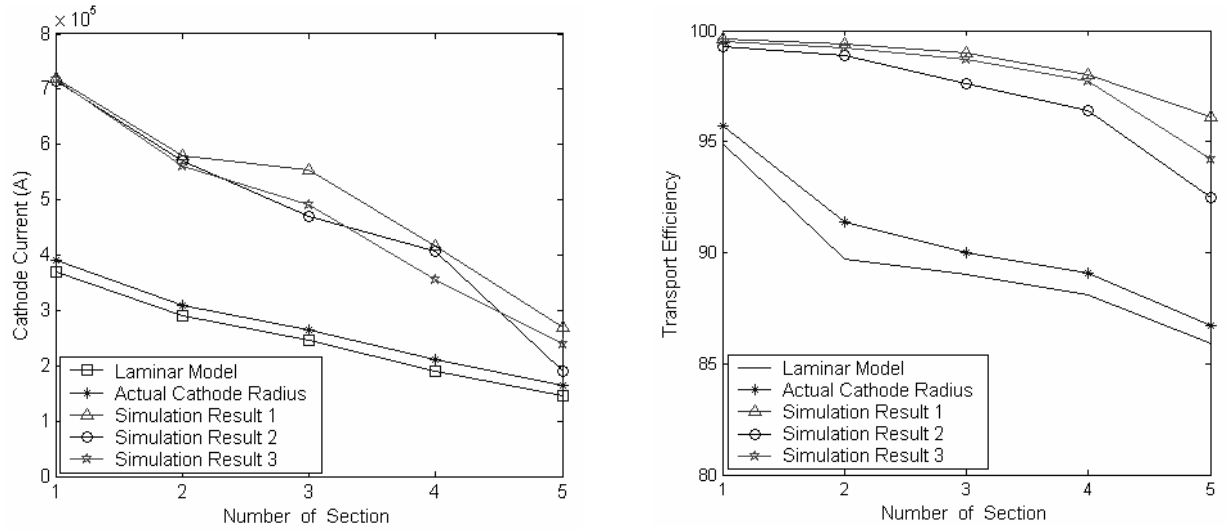
The solution is not sensitive to the particular equilibrium flow model (laminar flow or quasi-laminar flow model) used in the calculation. Matching self-limited currents section to section produces the result that each section of the voltage adder MITL behaves like an infinitely long MITL with the local voltage.

4. PARTICLE-IN-CELL SIMULATIONS VALIDATION AND RESCALING

The voltage adder MITL designed above is now simulated by PIC simulations to exam the working features. In simulations the input voltage pulses applied at the five feeds are all chock waves with 10-ns rise-time and afterwards maintaining the voltage amplitude. The MITL output boundary

condition is set to be a load-matched boundary which represents the isolation from the terminating load. Shown in Fig. 4 were the anode current and cathode current values at each segment, respectively from the laminar flow model and the simulation results at 35 ns into the pulse when a self-limited equilibrium has been established. As a result of matched self-limited design the simulation values of all of the segments are operating in the self-limited equilibrium with 750-kA anode currents. Whereas, from the simulations we can see that cathode currents at the 2nd to the 5th segment are lower than the analytical results. This is because at each geometrical discontinuity the disturbed electromagnetic field causes more electrons emit from the cathode and the electrons downstream tend to change their density to minimize the influence of the geometrical discontinuity [6]. While in the laminar flow model results are deduced by ignoring the effects of geometrical discontinuity of MITLs. Therefore, the cathode currents from the laminar model are higher than the simulations. Furthermore, in Fig. 4 the difference between the anode and cathode currents is the space electron flow which couples an amount of transportation power. From the results we can see that in the matched self-limited equilibrium more than half of the current is carried by the electron flow. If the load does not under-match the MITL, all these electrons and the currents associated with them will be lost before the load. Therefore, operating in the self-limited equilibrium the MITL has a limited transportation efficiency.

In order to enhance the transportation efficiency and lower the electron flow, dimensions of the MITL will have to be rescaled. As we know that the rate of the cathode to the anode current is higher when an MITL works in the load-limited equilibrium with the anode current locating on the super-insulated branches (left branches) of the curves in Fig. 3 and the electron flow sheath holding close to the cathode. Consequently, the MITL can be rescaled to make each segment operate under the load-limited equilibrium in order for high power transportation efficiencies. Since the last segment is connected with the final MITL, its cathode radius is still 23 cm as above. However the other segments can not be designed using the laminar flow method, because the free parameter in the model, the electron sheath boundary voltage, can only be determined beforehand in the self-limited equilibrium. Therefore, the cathode radii have to be calculated with the help of PIC simulations.



(a) (b)

Figure 5: Cathode currents and transport efficiencies at measurements in Table 1.

Based on analysis above we use PIC simulations to modify the cathode radii for the 1st to 4th segment of the adder MITL so as to adjust the anode current to be 750-kA at the local voltages. Three subsets of the modified dimensions are shown in Table 1. In order to decrease the manufacturing complexity each of the segment cathode radius is larger than the precedes by integral centimeters. It is shown in Figs. 5(a) and (b) that cathode currents and transportation efficiencies at each segment of these voltage adder MITLs compared with the analytically designed MITL. We can see that the cathode currents and transportation efficiencies of the rescaling MITLs are much higher than those of the analytically designed MITL. Therefore, one rule of rescaling MITLs for enhancing the transportation efficiency is that operating in the load-limited equilibra. Besides,

comparison of the three rescaling MITLs shows that the MITL with smaller step-downs in cathode radii has higher cathode currents and transportation efficiencies than the other two. And although the other two have the same starting and ending segments radii, the one with average cathode radius changes operates in a better condition than the other one. As known that at the geometrical discontinuity there are more electrons emitting from the cathode downstream to maintain the efficient impedance. Therefore the more abrupt the discontinuity is, the more electrons emit out so that the less the transportation efficiency will be. Accordingly, we get another designing rule to average and reduce the geometrical changes of MITLs.

However, with the changes in cathode radii the dimensions of the voltage feeds will have to be adjusted in order to match with the segments impedance and maintain the local voltages. Thereby, the simulation scales and manufacturing difficulties will be fairly large. Consequently, in actual project the radii of MITLs (also shown in Table 1) are commonly calculated based on analytical model results. The cathode current and transportation efficiency of the actual MITL are also shown in Fig. 6. Furthermore, it is used in actual projects the inductive rings (shown in Fig. 2) placed at the end of each feed which can help to reduce the electron flow by nearly 50%.

5. CONCLUSIONS

In practical applications, due to different geometrical features and operation requirements it is more practical combining analytical models and numerical simulations to optimize the design of MITLs. Based on the analysis and calculations conducted in this paper the design rules for MITLs can be summarized as follows.

1. The analytical models can be used in MITL designs even if the MITL has axial geometrical discontinuity, when a MITL is long enough compared with the input pulse duration. For a MITL with tapered radii segments, analytical model design methods of matched self-limited anode current for each of the MITL segment produces the result that each section operates in its self-limited equilibrium with the local voltage.
2. The power transportation efficiency of a MITL operating in the self-limited equilibrium is limited, because part of the power is coupled by the space electron flow that can not be delivered to the load. Consequently, the MITL will have to be rescaled with the help of PIC simulations to operate in the load-limited equilibrium in order to reduce the space electron flow and enhance the transportation efficiency.
3. In rescaling MITLs a good rule to follow is to average and reduce the geometrical changes in the radial direction.

ACKNOWLEDGMENT

This work was supported by the National Natural Science Foundation of China (No. 50677028), SRFDP (No. 20050003007) and Tsinghua Basic Research Foundation.

REFERENCES

1. Humphries Jr., S., *Charged Particle Beams*, Wiley-Interscience, New York, 1990.
2. DiCapua, M. S., "Magnetic insulation," *IEEE Trans. Plasma Sci.*, Vol. 11, No. 3, 205–215, 1983.
3. Goplen, B., L. Ludeking, D. Smithe, and G. Warren, "User-configurable MAGIC for electromagnetic PIC calculations," *Comp. Phys. Comu.*, Vol. 86, 54–86, 1995.
4. Mazarkis, M. G., J. W. Poukey, D. L. Smith, et al., "SABRE modification to a higher voltage high impedance inductive voltage adder (IVA)," *Proc. 11th IEEE Plasma Science Conf.*, 192, 1996.
5. Lockner, T. R. and J. W. Poukey, "Comparison of one-dimensional simulations of magnetically insulated gaps with analytic models," *J. Appl. Phys.*, Vol. 57, No. 7, 2652–2653, 1985.
6. Hiraoka, K., M. Nakajima, M. Shiho, et al., "Electron flow through geometrical discontinuity in coaxial magnetically insulated transmission lines," *Jpn. J. Appl. Phys.*, Vol. 39, No. 9, 5280–5286, 2000.
7. Mazarkis, M. G., J. W. Poukey, D. L. Smith, et al., "SABRE modification to a higher voltage high impedance inductive voltage adder (IVA)," *Proc. 11th IEEE Plasma Science Conf.*, 192, 1996.
8. Wang, M. Y. and M. S. DiCapua, "Operating point of long magnetically insulated vacuum transmission lines," *J. Appl. Phys.*, Vol. 51, No. 11, 5610–5614. 1980.

Analysis on the Shielding Effect of the Power Transformer Tank

Duo Chen, **Haiyu Yu**, **and Jiansheng Yuan**

State Key Lab of Power Systems, Department of Electrical Engineering
Tsinghua University, Beijing 100084, China

Abstract— The MV/LV transformer without a tank in the underground substation can produce remarkable interference with the computer Cathode-Ray Tube (CRT) displayers above the ground, because of the generated power frequency magnetic field. To resolve such problem, a tank with high conductivity or/and high permeability is needed to enclose the transformer. To evaluate the shielding effectiveness, the finite element method is adopted to calculate the leakage magnetic field outside the transformer tank with different materials. The calculated results show that the steel tank is able to achieve an enormous shielding effectiveness.

1. INTRODUCTION

The shielding effectiveness of the power transformer is a very important target in transformer design. When the transformer is located near to the computer Cathode-Ray Tube (CRT), even very small power frequency magnetic field intensity can generate remarkable interference. According to the IEC 61000-4-8 (1993) [1], the environment of power frequency magnetic field is defined as five levels based on its magnitude. Different electronic devices can work normally at different level. According to our test, the CRT can not normally display or the screen will dither obviously, when the background magnetic field intensity is greater than 1A/m (or the magnetic flux density is greater than $1.256\,\mu$T)—worse than the first level environment [1]. If such sensitive electronic devices locate just above the underground substation, they will be interfered and can not work normally.

Usually, the dry transformers are employed in such MV/LV (Middle voltage/Lower voltage) substations. Therefore, from the engineering point of view, it is not necessary to set any tank to the transformers. However, from the electromagnetic inference point of view, a tank of metal or ferromagnetic material must be set as the electromagnetic shielding. Due to the high conductivity of metal materials, the induced eddy current in the metal tank is large, which can reduce the leakage field outside the tank. Due to the high permeability of ferromagnetic material, the magnetic field is bounded inside the tank, so that the field outside the tank is reduced. By considering such two kinds of behavior together for ordinary materials (e. g., steel), the analysis and the design of the tank become complex. It is almost impossible to calculate the leakage field outside the tank by analytical solutions. In contrast, the numerical computation method must be used to obtain the result of such complex problem.

In this paper, aluminum, steel and an assumptive material with zero-conductivity and a high relative permeability of 10000 are chosen to analyze and compare their shielding effect. The three-dimensional finite element method is adopted to calculate the leakage field outside the transformer tank. The paper is organized as follows: Section 2 introduces the parameter of a MV/LV transformer and its three-dimensional finite element model. Section 3 analyzes the simulation results and compares the shielding effectiveness of tanks with different materials.

2. THE TRANSFORMER AND ITS FINITE ELEMENT MODEL

The data of the concerned three-phase MV/LV transformer is in Table 1. To simplify the finite element simulation model, only the A phase is analyzed. The schematic diagram of the transformer in the underground substation and the cared region is shown in Fig. 1. The thickness of the tank is 2 mm, and the observed line of the cared region is 2 m above the top of the transformer tank.

The finite element method is adopted to calculate the leakage magnetic field outside the transformer tank. Due to the high conductivity of aluminum and steel, the induced eddy current in the tank is large, which can generate magnetic field and ultimately contributes to the shielding effectiveness. Therefore, to correctly evaluate this eddy current effect, the transformer tank region is finely meshed. The meshed model of the transformer and the metal tank are shown in Fig. 2.

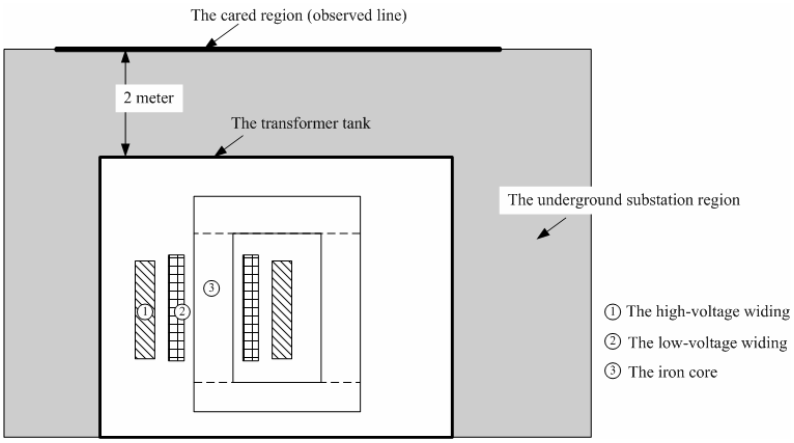


Figure 1: The schematic diagram of the transformer in the underground substation and the cared region.

Table 1: Characteristics of the three-phase power transformer.

RATED POWER	2000 KVA
Rated High Voltage	10 kV
Rated Low Voltage	400 V
Rated Frequency	50 Hz
No-Load Current	0.5%

3. CALCULATION RESULTS AND ANALYSIS

The leakage magnetic flux density along the observed line is calculated. In order to obtain the shielding effectiveness, the leakage magnetic flux density outside the metal tank is compared with the no-tank situation. In the simulation model, the no-tank situation is calculated by setting the tank material parameter as air. The shielding effectiveness is evaluated by using the reduction ratio of the leakage field, which is defined as follows:

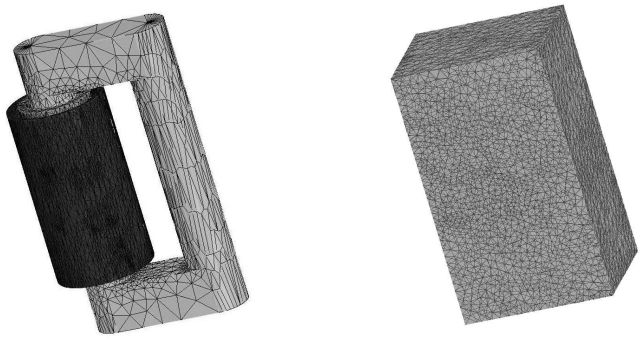
$$R = \frac{|B_{\max}^{\text{material}} - B_{\max}^{\text{air}}|}{|B_{\max}^{\text{air}}|} \times 100\% \tag{1}$$

The comparison results of aluminum, steel and material 3 (the assumptive one defined above) are shown in Table 2.

Table 2: Shielding effectiveness of different material tank.

Material	Conductivity (S/M)	Relative Permeability	Maximum B (μT)	Reduction Ratio
Air	0	1	4.20	
Steel	1.03E+07	2.00E+03	0.134	96.81%
Aluminum	3.80E+07	1	0.219	94.79%
Material 3	0	1.00E+04	0.647	84.60%

From the results in Table 2, it can be seen that when there is no tank to enclose the transformer, the leakage magnetic flux density in the cared region reaches 4.20 μT, which exceeds 1.256 μT, the first environment level of the IEC 61000-4-8 (1993), so that the display of the computer CRT will be influenced. However, when the transformer is enclosed by the high conductivity metal or the high permeability material tank, the leakage magnetic flux density in the cared region is reduced greatly to be under the value of the first environment level.



(a) (b)

Figure 2: 3D finite element model of the transformer: (a) the meshed model of the core and windings, (b) the meshed model of the transformer tank.

In addition, Table 2 shows that both the steel and the aluminum can obtain better shielding effectiveness than Material 3, because the induced eddy current on the high conductivity material tank is large, which can generate inverse direction magnetic flux density to counteract the original leakage magnetic field. Besides, the steel achieves the best among the three materials, which means that the shielding effect of the material with both high conductivity and high permeability is better than the one only with high conductivity, and the one only with high permeability. This is an important conclusion for the power frequency magnetic shielding. The transformer enclosed by such material tanks can maintain the leakage magnetic field in a very small magnitude. Therefore, the computer CRT can be in its normal working status. The leakage magnetic flux density distribution along the observed line is plotted in Fig. 3, which also shows the enormous shielding effectiveness of the materials.

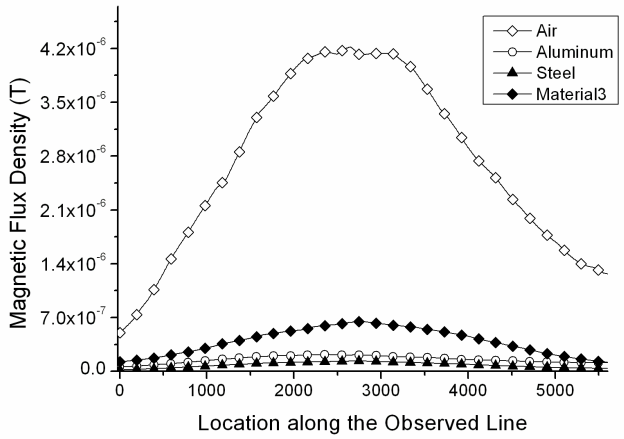


Figure 3: Leakage magnetic flux density along the observed line.

4. CONCLUSIONS

The MV/LV transformer without a tank in the underground substation can generate great leakage magnetic field, which exceeds the value of the first environment level defined by the IEC 61000-4-8 (1993). In this case, sensitive devices such as computer CRT displayers can not display normally near the underground substation. In order to resolve such interference, a tank of high conductivity or/and high permeability is needed to enclose the transformer to shield the leakage magnetic field. Among the three analyzed materials, aluminum, steel and Material 3, the steel tank with both high conductivity and high permeability achieves the best shielding effectiveness. With such tank, the maximum leakage magnetic field in the cared region can be reduced twenty times, so that the field is much smaller than the value of the first environment level.

ACKNOWLEDGMENT

This work was supported by the National Natural Science Foundation of China (No. 50677028), SRFDP (No. 20050003007) and Tsinghua Basic Research Foundation.

REFERENCES

1. IEC 61000-4-8, "Testing and measurement techniques-Section 8: Power frequency magnetic field test," *Electromagnetic Compatibility (EMC)-Part 4*, Basic EMC Publication, 1993.

A Miniaturized 2.45 GHz RFID Tag Antenna Using Planar Impedance Transformer

Zhiguang Fan[1], **Shan Qiao**[1,2], **Jiangtao Huangfu**[1], **and Lixin Ran**[1]

[1]Department of Information and Electronic Engineering, Zhejiang University
Hangzhou 310027, China
[2]Zhejiang University City College, Zhejiang University
Hangzhou 310027, China

Abstract— We proposed a miniaturized 2.45 GHz RFID tag antenna consisting of a dipole radiator and a planar impedance transformer formed on a copper foil layer cladding on 0.5 mm thick FR4 substrate. The impedance transformer comprises two inductively coupled spiral inductor to realize both resistance boosting and inductive compensation for the short dipole radiator to be conjugate-matched to the microchip. The length of the dipole radiator and the side length of the spiral inductor can be easily adjusted to tune the input resistance and inductance of the tag antenna. Finite Difference Time Domain (FDTD) 3D electromagnetic simulation method is utilized to simulate and optimize the tag antenna design. The final design occupies only an area of 33.55 mm×8.54 mm. The omnidirectional radiation pattern, the peak gain value of −0.2 dBi and the radiation efficiency of 0.56 are also achieved, which demonstrate acceptable electrical performance for miniaturized tag design.

1. INTRODUCTION

Passive Ultra High Frequency (UHF) Radio Frequency Identification (RFID) technology has become increasingly popular in various areas of automatic identification of objects such as access control, animal tracking, inventory management, asset identification and manufacturing industry [1, 2]. A basic passive UHF RFID system consists of a passive tag with stored information that is attached to an object and a reader that transmits commands and energy to active tag by electromagnetic wave. Communication from tag to reader is based on electromagnetic wave backscattering modulation by the data stored on tag [3]. Generally, a RFID tag consists of an Application Specific Integrated Circuit (ASIC) microchip connected to an antenna. There is no internal source of energy in a passive tag just like the meaning of "passive". Compared with active one, passive tag is characteristic of simpler circuits, lower cost and smaller size, which enhance its usability in commercial applications of large scale. Particularly, small tag size is often strongly desired for being easily attached to an article with limited volume. A tag antenna typically has a size comparable to a wavelength of working frequency, which is generally much larger than the tag microchip's. Therefore, an important consideration in miniaturized tag design is to adopt those antenna designs with both low profile and compact size. Some reported RFID tag antennas such as half-wavelength dipole antenna, planar inverted F-antenna [4], folded dipole antenna [5], etc., apparently can not meet the above requirements adequately. Meander Line Antenna (MLA), as an attractive choice for electrically small antennas, has been proposed and designed for RFID tag applications by means of genetic algorithms (GA) in [6]. However, since GA is kind of optimisation method based on conditional random searching, its method is inherently computationally intensive and hardly provides intuitive design information.

Another important consideration for tag design is the impedance matching between tag microchip and antenna. Under the condition of conjugate matching, the power transfer from a tag antenna to its terminated microchip is maximized and consequently the reading distance is optimized. A dipole-like small antenna has low input resistance and high input capacitance, which makes it severely mismatch with tag microchip which is generally of high input resistance and high input capacitance. However, due to high fabrication cost, it is usually prohibitive to add an external matching network with discrete elements including inductors and capacitors in tag design. Therefore, an ideal tag antenna must be capable of self-tuning to specified impedance in order to directly match with tag microchip.

In this paper, we proposed a miniaturized 2.45 GHz RFID tag antenna that consists of a dipole radiator and a planar impedance transformer formed on a copper foil layer cladding on 0.5 mm thick FR4 ($\varepsilon_r = 4.4$, $\tan\delta = 0.02$) substrate. The impedance transformer comprises two inductively coupled spiral inductor to realize both resistance boosting and inductive compensation for the

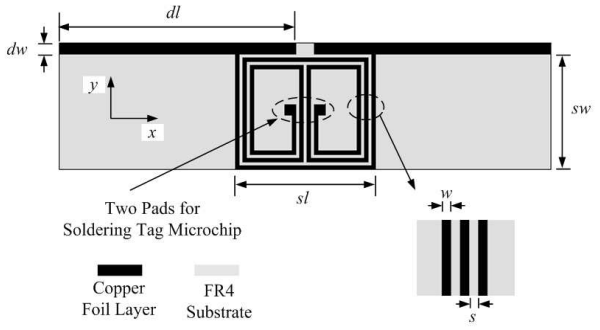


Figure 1: Miniaturized 2.45 GHz RFID tag antenna using planar impedance transformer.

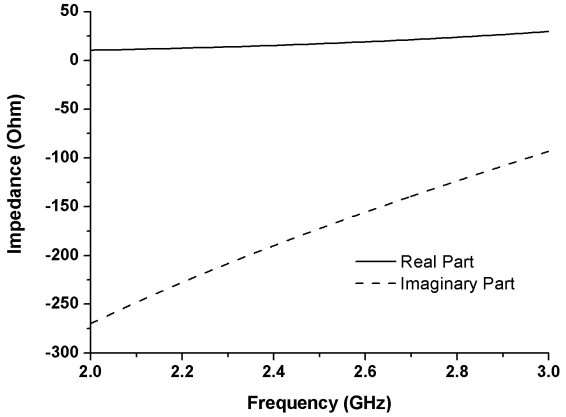


Figure 2: Input impedance of the dipole radiator.

short dipole radiator to be conjugate-matched to the microchip. Two dimension parameters, i.e., the length of the dipole radiator and the side length of the spiral inductor can be easily adjusted to tune the input resistance and inductance of the tag antenna. Finite Difference Time Domain (FDTD) 3D electromagnetic simulation method is utilized to simulate and optimize the tag antenna design. The final design occupies only an area of 33.55 mm × 8.54 m. The omnidirectional radiation pattern, the peak gain value of −0.2 dBi and the radiation efficiency of 0.56 are also achieved, which demonstrate acceptable electrical performance for miniaturized tag design.

2. TAG ANTENNA STRUCTURE

The proposed miniaturized 2.45 GHz RFID tag antenna is shown in Fig. 1. It consists of a dipole radiator and a planar spiral impedance transformer which both are formed by etching the copper foil cladding on FR4 dielectric substrate ($\varepsilon_r = 4.4$, $\tan\delta = 0.02$). The copper foil is 0.035 mm thick and the substrate is 0.5 mm thick. A tag microchip can be bonded to two square pads of 0.75 mm ∗ 0.75 mm at the centre of the spirals. The dipole radiator has a length much less than half-wavelength of 2.45 GHz in free space. Thus, it has both a low input resistance and also a high input capacitive reactance as illustrated in Fig. 2, for example, 16.1–j181.0 Ω at 2.45 GHz. In our application, the tag microchip has an input impedance of 80.0–j232.0 Ω. In order to make the dipole radiator to be conjugate-matched to the microchip, two inductively coupled spiral inductor are designed to compose an impedance transformer to realize both resistance boosting and inductive compensation. Both the line-width w and line-space s of the two spiral inductors are 0.25 mm. The single-arm length dl and width dw of the dipole radiator are 15.4 mm and 0.75 mm, respectively. The side length sl of the outer inductor is 9.29 mm, which is always 1.5 mm bigger than its side width sw. The input impedance $Z_T = R_T + jX_T$ of the tag antenna varies while sw and dl being tuned, as is illustrated in Fig. 3. It can be observed that the resistance R_T increases while dl increases and the reactance X_T decreases while sw decreases. Thus, we can make use of these two dimension parameters to tune the input resistance and reactance of the tag antenna.

3. RESULTS

Based on the design analyses in Section 2, we used Finite Difference Time Domain (FDTD) 3D electromagnetic simulation method to simulate and optimize the tag design. When dl is tuned to 15.4 mm and sw to 7.79 mm, the tag antenna impedance takes on 80.3–j228.9 Ω, which is approximately perfectly conjugate-matched to the tag microchip. The final tag design occupies a rectangular area of 33.55 mm × 8.54 mm. The diagonal dimension of the rectangular is 34.62 mm, which is far less than half-wavelength 61.22 mm at 2.45 GHz in free space.

Figures 4(a), (b) and (c) show the simulated tag antenna radiation patterns in xy-plane, yz-plane and xz-plane, respectively. As expected, its radiation patterns are similar to those of a classic dipole antenna. In yz-plane, it demonstrates the omnidirectional properties. In both xy-plane and xz-plane, the radiation nulls can be observed in the x-direction. The peak gain value of −0.2 dBi and the radiation efficiency of 0.56 are achieved, which are good enough to be acceptable for a miniaturized UHF RFID tag design. Particularly, since FR4 material has very high dielectric loss at 2.45 GHz ($\tan\delta = 0.02$), it can be expected that the radiation efficiency and the gain would be

improved much if other substrate of lower loss were used.

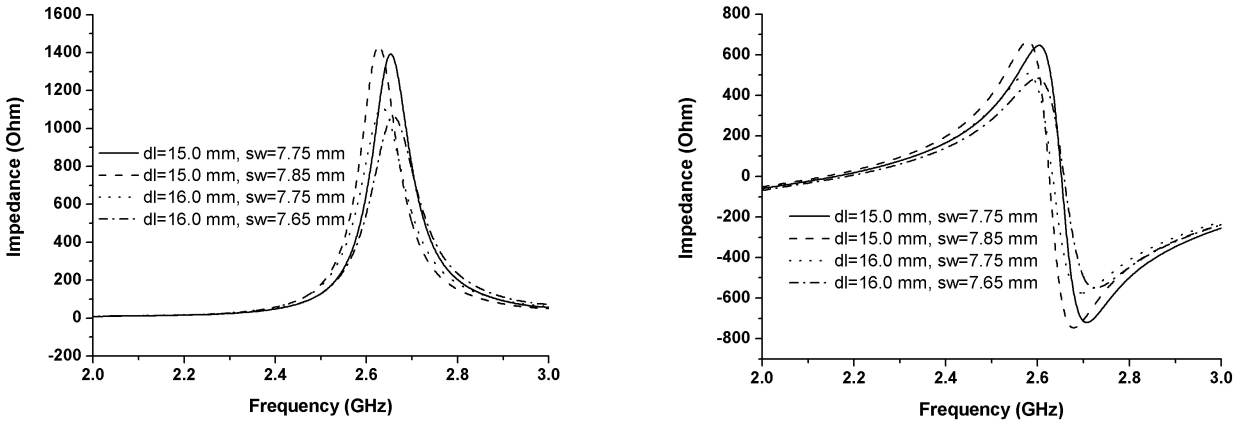


Figure 3: Input impedance of the miniaturized 2.45 GHz RFID tag antenna. (a) Real part, (b) Imaginary part.

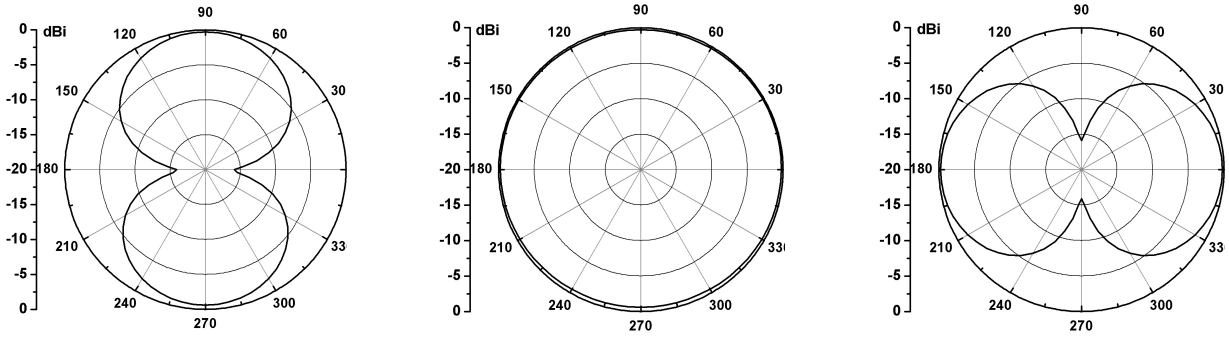


Figure 4: Radiation patterns of the miniaturized 2.45 GHz RFID tag antenna. (a) xy-plane, (b) yz-plane, (c) xz-plane.

4. CONCLUSIONS

In this paper, we proposed a miniaturized 2.45 GHz RFID tag antenna that consists of a dipole radiator and a planar impedance transformer formed on copper foil cladding FR4 substrate. We have used Finite Difference Time Domain (FDTD) 3D electromagnetic simulation method to simulate and optimize the tag design. The tag antenna was tuned to be approximately perfectly conjugate-matched to the tag microchip to maximize power transfer and consequently the reading range. This tag antenna achieves the omnidirectional radiation pattern, the peak gain value of -0.2 dBi and the radiation efficiency of 0.56. The final design occupies a small area of 33.55 mm $\times$ 8.54 mm. All these properties make this tag design a good choice in compact UHF RFID applications.

ACKNOWLEDGMENT

This work is supported by ZJNSF R105253, and in part by NSFC 60671003 and 60531020.

REFERENCES

1. Tuttle, J. R., "Traditional and emerging technologies and applications in the radio frequency identification (RFID) industry," *Proc. IEEE RFIC Symp.*, 5–8, Denver, CO, Jun. 1997.
2. Bansal, R., "Coming soon to a Wal-Mart near you," *IEEE Antennas Propag. Mag.*, Vol. 45, 105–106, Dec. 2003.

3. Stockman, H., "Communication by means of reflected power," *Proc. IRE*, 1196–1204, Oct. 1948.
4. Hirvonen, M., P. Pursula, K. Jaakkola, and K. Laukkanen, "Planar inverted-F antenna for radio frequency identication," *Electron. Lett.*, Vol. 40, 848–850, Jul. 2004.
5. Qing, X. and Y. Ning, "A folded dipole antenna for RFID," *Proc. IEEE Antennas and Propagation Soc. Int. Symp.*, Vol. 1, 97–100, Jun. 2004.
6. Marrocco, G., "Gain-optimized self-resonant meander line antennas for RFID applications," *Antennas Wireless Propag. Lett.*, Vol. 2, No. 21, 302–305, 2003.

Simulation Study for the Decoding of UHF RFID Signals

Shengli Wang[1], **Shan Qiao**[1,2], **Shaoyuan Zheng**[1], **Zhiguang Fan**[1]
Jiangtao Huangfu[1], **and Lixin Ran**[1]

[1]Department of Information and Electronic Engineering, Zhejiang University
Hangzhou 310027, China
[2]Zhejiang University City College, Zhejiang University
Hangzhou 310027, China

Abstract— In this paper, we present a simulation model for the decoding of the UHF (Ultra High Frequency) RFID (Radio Frequency IDentification) signals utilizing MATLAB/Simulink. A digital phase-locked loop (DPLL) was designed for recovering the external clock which is used for the bit synchronization and sampling. The IDs are well recovered under the existence of an additive signal noise ratio (SNR).

1. INTRODUCTION

Radio Frequency Identification (RFID) is a technology using radiated and reflected RF power to identify a variety of objects. A typical RFID system consists of a transceiver, or RFID reader, and a transponder, or tag. Generally, an RFID reader includes an RF transmitter, one or more antennas and an RF receiver. A generic RFID tag consists of a microchip attached to an antenna. In a UHF RFID system, communication between the reader and the transponder is via backscatter reflection [1, 2].

According to the Class 1, Gen 2 UHF Air Interface Protocol Standard [3], the return link data from the transponder to the reader is encoded by a scheme of Bi-phase space (FM0). Figure 1 illustrates the basic functions of the FM0 data coding. FM0 inverts the baseband phase at each symbol boundary, and a bit 0 has an additional mid-symbol phase inversion.

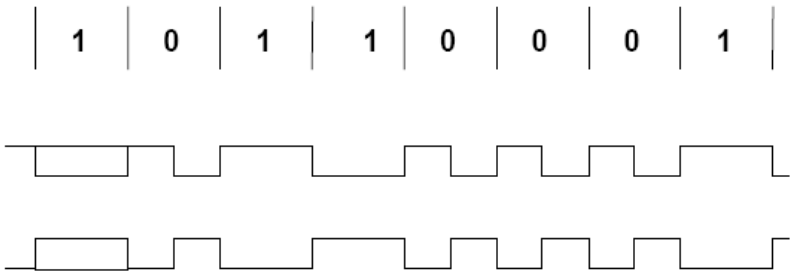


Figure 1: Fm0 data coding.

In this paper, a decoding model of the UHF RFID signals is described by use of MATLAB/ Simulink [4]. In the simulation model, a DPLL is designed for recovering the external clock which is used for bit synchronization and sampling. Sub-blocks of the decoding model are described in detail, and the results of the simulation model are presented under the existence of an additive signal noise ratio (SNR).

2. DECODING MODEL DESCRIPTION

Figure 2 is the simulation environment of the decoding model realized in Simulink. In the FM0 Code Generator block, the state machine produces a digital stream carrying a random noise based on the communication protocol implementation. The stream enters into the FIR (Finite Impulse Response) block to filter out the noise. The digital phase-locked loop (DPLL) was designed for recovering an external clock from an input serial data stream directly. Then the signal is sampled by the synchronization clock coming from the DPLL and is converted to a bit stream described by binary digit '0' or '1'. The Decoding block is used to finally recover the source IDs. The outputs from the FM0 Code Generator block and the Decoding block are simultaneously sent to the Error Rate Calculation block, to compute the Bit Error Rate (BER). The Scope and Display blocks are used for demonstraing the results of the simulation.

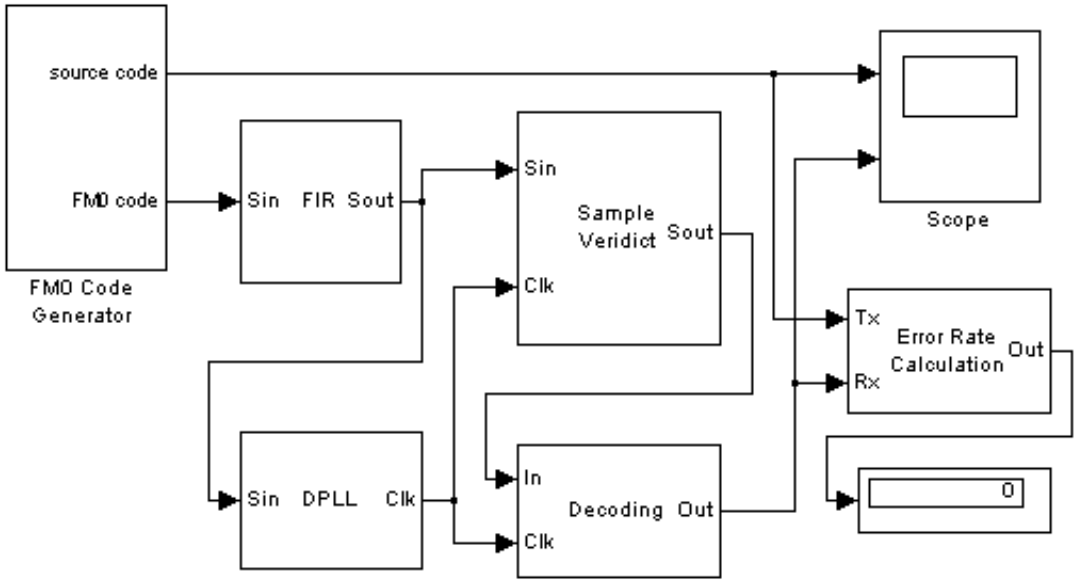


Figure 2: Top-level block diagram of the decoding model in this paper.

3. FM0 CODE GENERATOR MODULE

Figure 3 shows the basic blocks of the FM0 Code Generator Module in Simulink. The Bernoulli Binary Generator block generates a 40 kbps random binary stream satisfying a Bernoulli distribution. The FM0 Encoding block is an S-functions (system-functions) block, providing a powerful mechanism for extending the capabilities of Simulink. In this paper it was written in M-language of MATLAB. The Band-Limited White Noise block produces normally distributed random numbers that are suitable for the use in continuous or hybrid systems. Finally, the output of this module is a sumation of the above two with a constant offset of -0.5. Figure 4 shows the waveform produced by this module.

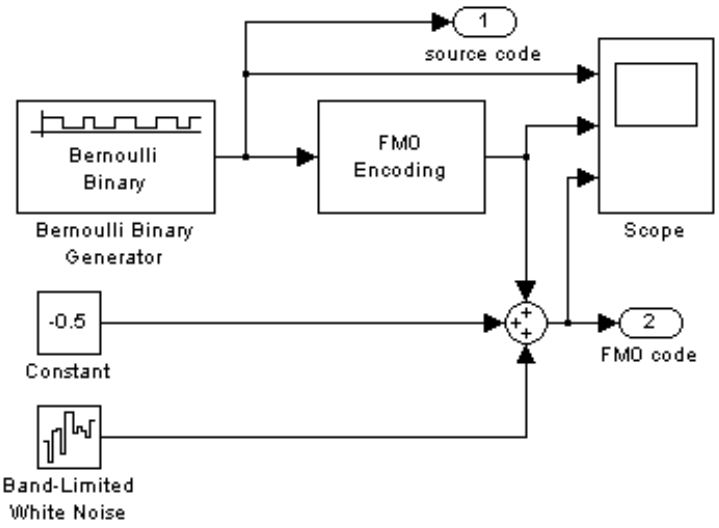


Figure 3: Implementation of the FM0 Code Generator Module.

4. DPLL MODULE

The block diagram of the digital phase-locked loop (DPLL) used to directly recover the external clock from the input stream is shown in Figure 5. The Local Clock block produces two clocks with a phase difference of π, and both with the frequency f0 equals mRb, where Rb is the data rate of the input stream and m is the dividing factor of the divider. The input data stream passes through

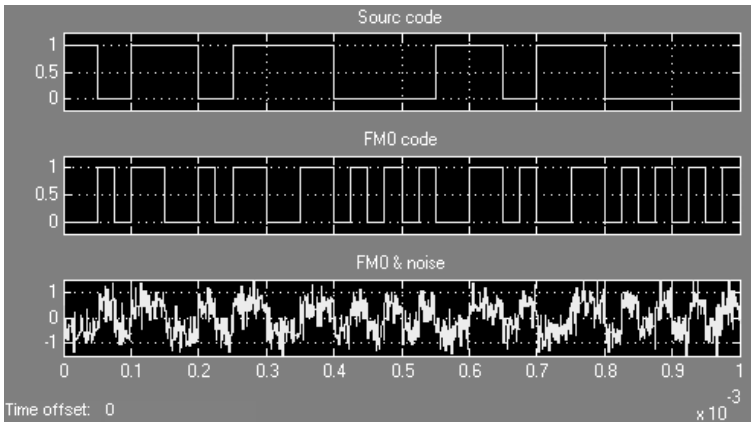


Figure 4: Waveform of the FM0.

the Edge Detector circuit and was converted to a narrow pulses sequence. The gate 'AND1' will be a pulse output when the phase of the said narrow pulses sequence lags the divider's clock pulse output, then it will be added to the divider through the Controller circuit to add a clock pulse which will be divided by the divider, in order to adjust the phase of the clock output. If the phase of the said narrow pulses leads the said clock pulses, the gate 'AND2' will be a pulse output which will employ the divider to minus a clock pulse to make the clock output a little delay. Until the Phase Comparator has no pulse output, the loop is locked. Figure 6 is the simulation results of the DPLL, in which the lower waveform is the clock recovered from the input steam.

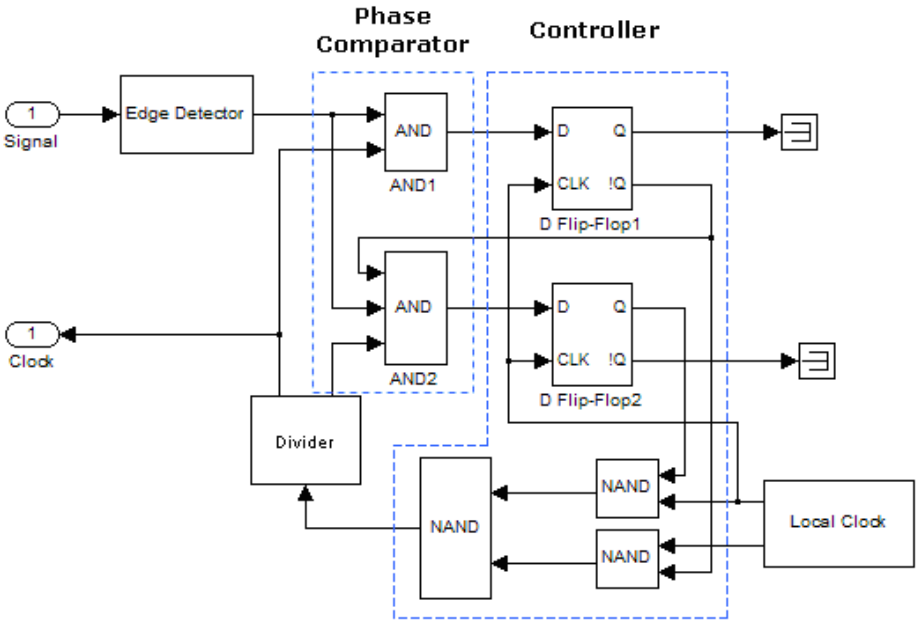


Figure 5: Simulation blocks of the DPLL.

5. SIMULATION RESULTS

As shown in Figure 7, the IDs are well recovered from the noisy FM0 code stream. Comparing with the source IDs, the recovered just have a time delay. In the simulation, we adjusted the noise power in the model to obtain the Bit Error Rate (BER) by the Error Rate Calculation block. The results (not shown in this paper) indicate that when the SNR $> -1.0\,\mathrm{dB}$, the BER is still zero, which implies a good performance of the simulation proposed in this paper.

To verify the simulation aforementioned, we collected the baseband data from a real UHF RFID reader, as shown in Figure 8, where Figure 8(a) is the waveform of the digital signal that was sampled from the ADC, and 8(b) shows the waveform of the data that was filtered by the FIR

filter. Then the data steam was imported into the simulation environment as a block instead of the FM0 Code Generator module. It was finally decoded successfully with the simulation model.

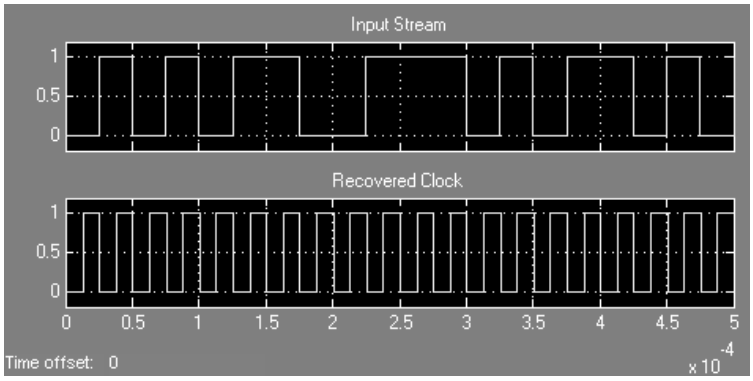


Figure 6: The input stream and the recovered clock of the DPLL.

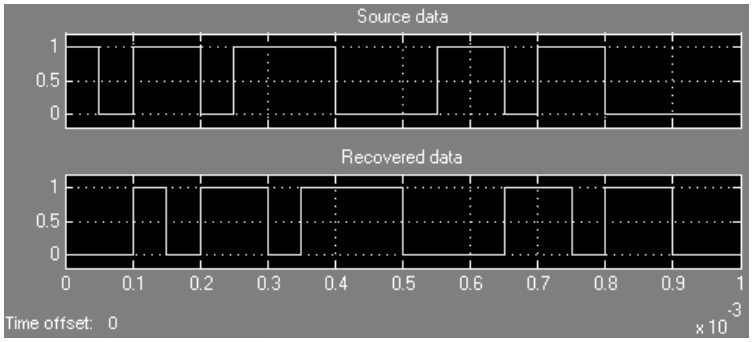
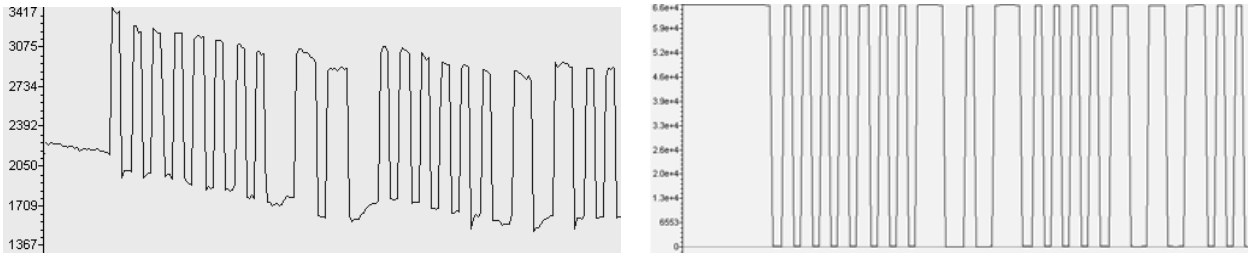


Figure 7: Waveforms of the simulation results.



(a) (b)

Figure 8: The baseband signal from the DSP, (a) Signal data form the ADC, (b) The filtered waveform.

6. CONCLUSION

A simulation environment for the decoding of the UHF RFID signals has been discussed in this paper. A DPLL module was designed in the simulation model for recovering the synchronous clock, and the results of the simulation show that the ID can be well decoded whether the FM0 code stream is provided by the Simulink module or it was sampled from an actual RFID reader. Such simulation model has been realized in a TI C54X series DSP of a UHF RFID reader.

ACKNOWLEDGMENT

This work is supported by ZJNSF R105253, and in part by NSFC 60671003 and 60531020.

REFERENCES

1. Landt, J., "The History of RFID," *IEEE Potentials*, Vol. 24, No. 4, 8–11, Oct.–Nov. 2005.

2. Finkernzeller, K., *RFID Handbook: Fundamentals and Applications in Contactless Smart Cards and Identification*, second edition, Wiley and Sons, NJ, 2003.
3. Class 1 Generation 2 UHF Air Interface Protocol Standard. http://www.epcglobalinc.org/standards/.
4. Simulink manual from http://www.mathworks.com/products/simulink/.

A System Design for the Reader of Microwave Radio Frequency Identification

Changzhan Gu[1], **Zhiguang Fan**[1], **Shaoyuan Zheng**[1], **Jiangtao Huangfu**[1,2], **and Lixin Ran**[1,2]

[1]Department of Information and Electronic Engineering, Zhejiang University
Hangzhou 310027, China
[2]The Electromagnetics Academy at Zhejiang University, Zhejiang University
Hangzhou 310027, China

Abstract— RFID (Radio Frequency Identification) is now attracting many interests for its bright future in commercial and daily use. For the low-frequency (LF, typically 125 KHz) and the high-frequency (HF, typically 13.6 MHz) RFIDs, the Readers have already come into our daily life out of their matured technologies, while the ultra-high frequency (UHF) RFID Readers are still in the process for the reliability and functionality. This paper represents a systematic design of a UHF RFID reader. Received backscatter modulation waves (carrier wave frequency 915 MHz) are directly demodulated to baseband signals using the Zero IF technology. The reader is designed as a quadrature (I/Q) one, which can avoid null points of the received signals [4] and enhance the demodulation sensitivity.

1. INTRODUCTION

RFID is the short for "Radio Frequency Identification", which is a generic term for non-contact technologies that supply means to automatically identify people and objects on the basis of radio waves. RFID is, to a certain extent, similar to the bar coding, however, it greatly enhances the data process and acts as a complement to the present existing technologies. RFID covers a wide variety of applications such as the building access control, toll collection, vehicle parking access control, anti-theft system, animal tracking, inventory management and so on [3]. It is widely believed that RFID will supersede bar coding in the following few years.

An operational RFID system typically falls into three parts: an RFID tag (a combination of a microchip and an antenna), an RFID reader, and a host computational system that communicates with the reader [2, 3]. According to the frequency used in an RFID system, there are usually four broad RFID categories: low frequency (LF, usually around 125 and 134.2 KHz), high frequency (HF, usually around 13.56 MHz), ultra-high frequency (UHF, usually around 868 and 928 MHz) and microwave (usually around 2.45 and 5.8 GHz). For tags, there are two major types: the passive tags, which draw their power from the transmission of readers through electromagnetic coupling; and the active tags, which have their own power supplies. This paper introduces an RFID reader that communicates with the passive tags which are much more widespread used. The reader transmits a carrier (unmodulated wave) to the tag and powers up the microchip inside the tag, and then the tag reacts to the reader by transmitting a series of data, the ID. Backscatter modulation is applied in this process, which can be described as follows: the reader transmits an unmodulated carrier and receives a modulated backscattering from the tag [1].

2. RECEIVER DESIGN

The reader is constructed as a quadrature structure whose receiver chain includes I (in-phase) and Q (quadrature) subchains. The block diagram of the reader's architecture is shown in Figure 1.

The design is based on a direct conversion topology, which is known as "homodyne", in which received signals are down converted (demodulated) directly from RF to baseband. This requires that the Local Oscillator (LO) must be tuned to and synchronized in-phase with the carrier frequency [2], which is 915 MHz in this paper. Input signals, which has been backscatter modulated by the tag, are usually very weak signals with a power level only measured several decades mV. Thus it is probable that the incoming signals are fully submerged in the relatively more powered carrier. I/Q demodulator is employed in the receiver chain to act as an advantageous component to help improve overall signal to noise ratio (SNR) as well as the LO carrier leakage suppression [2]. The carrier is generated by an ADI's ADF4360-7, which integrates an integer-N synthesizer and a VCO. By choosing appropriate external inductors, we get our required RF carrier. LO carrier for the receiver chain goes through a 2-way power splitter that outputs two quadrature carriers — the in-phase (I) channel and the (Q) quadrature channel. Dual matched MMIC amplifier MERA-556 amplifies

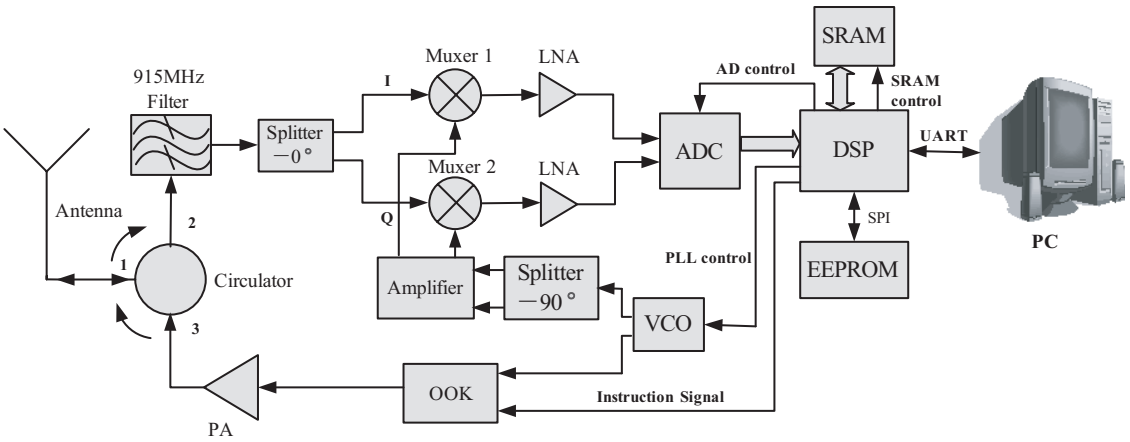


Figure 1: Block diagram of UHF RFID reader system.

the two quadrature carriers to a proper power level to meet the requirements for the inputs of the mixers. Consequently, quadrature carrier signals used for I/Q demodulation are acquired.

Consider an RF carrier from VCO

$$X(t) = A\cos(w_0 t + \theta_0). \tag{1}$$

Passing through the splitter leads to a -90° phase shift in one output channel,

$$X_1(t) = A\cos\left(w_0 t + \theta_0 - \pi/2\right) = A\sin(w_0 t + \theta_0), \tag{2}$$

while other channel is in-phase with the initial RF carrier, i.e.,

$$X_2(t) = X(t) = A\cos\left(w_0 t + \theta_0\right), \tag{3}$$

where $X_1(t)$ and $X_2(t)$ are the two quadrature carriers waves.

Coupling through the circulator, incoming signals are filtered via a SAW filter whose centre frequency is 915 MHz. Another 2-way power splitter is employed to divide the incoming signals into two identical and equal parts without any phase shift. Two mixers with differential inputs demodulate the above signals directly to baseband, as shown in Figure 2. It is seen from the upper waveform that the V_{P-P} of the baseband signals is only 32.5 mV, which does not satisfy the ADC's input requirement. Therefore, additional low-noise amplifiers (LNA) are used and the amplified waveform is shown in the lower part of Figure 2.

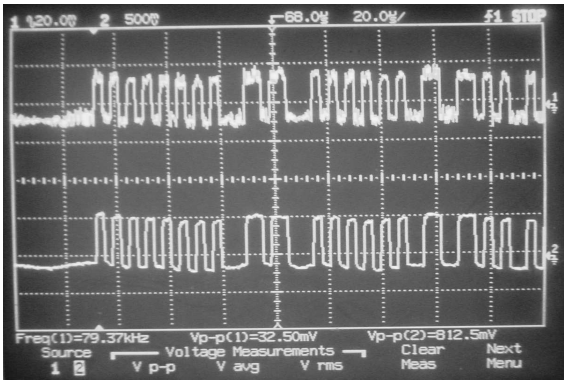


Figure 2: Waveforms of the demodulation output (Channel 1, $V_{P-P} = 32.5$ mV) and the LNA output (Channel 2, $V_{P-P} = 812.5$ mV).

3. TRANSMITTER DESIGN

Most of the passive tags can be read by a reader directly, but some tags must be read according to specific communication protocol. When a reader is reading a tag, it is also providing tags with

power, which enables the chips on the tags to work normally. The reader sends out ASK modulated signals, then the chip on the tag derives its operating power from the RF beam transmitted by the reader. The RF beam is received and rectified by the chip. When the tag gets enough energy and the correct commands, it will return its own unique User ID. The reader must send out a continuous carrier wave to provide energy to the circuit on the tag, and the tag transmits its factory-programmed code back to the reader by varying the amount of energy that is reflected from the chip antenna circuit.

The commands signal produced by DSP is modulated in OOK (On-Of-Keying) method, then magnified by power amplifier, and finally emitted by an antenna. All commands are transmitted from the reader to the tags by means of pulse interval encoding, and the average bit rate is 33 kbps.

The output power of Power Amplifier is limited, and generally no more than 1 Watt in USA. A Power Amplifier with a typical output power of 1.2 Watts is used in the system, and a tag in a distance of about 4 meters away from the antenna can access enough energy. A circulator, whose typical isolation is 22 dB and typical insertion loss is 0.25 dB, is used to insulate the transmitter signal and the receiver signal.

4. BASEBAND DESIGN

Baseband chain is composed of ADC, DSP, EEPROM and SRAM. An ADC with two 12 bits channel is used to sample and hold differential input analog signal. Maximum SNR performance will be achieved with the ADC set to the largest input span of 2V p-p, and common-mode voltage of the input signal is easily set to the half of power voltage. The lowest typical conversion rate of the ADC is 1 MSPS. The separate clock inputs for each channel should have a nominal 50% duty cycle and commonly a 5% tolerance.

A DSP is a system controller and baseband processor used for configuring PLL, controlling ADC, encoding, decoding and communicating with PC. The TMS320VC5402 fixed-point, digital signal processor (DSP) is based on an advanced modified Harvard architecture that has one program memory bus and three data memory buses, with 10-ns execution time for a single-cycle, fixed-point Instruction (100 MIPS). It has 4K × 16-Bit On-Chip ROM and 16K × 16-Bit Dual-Access On-Chip RAM, and we use a 64K × 16-Bit SRAM to extend data and program space.

An EEPROM is included for storing system programs, and communicating with the DSP through Serial Peripheral Interface (SPI). The DSP bootloader is used to transfer codes from the external source EEPROM into internal program memory following power-up. This allows codes to reside in slow, non-volatile memory externally, and be transferred to high-speed memory to be executed. The EEPROM clock rate is set to 400 kHz for a 100 MHz device.

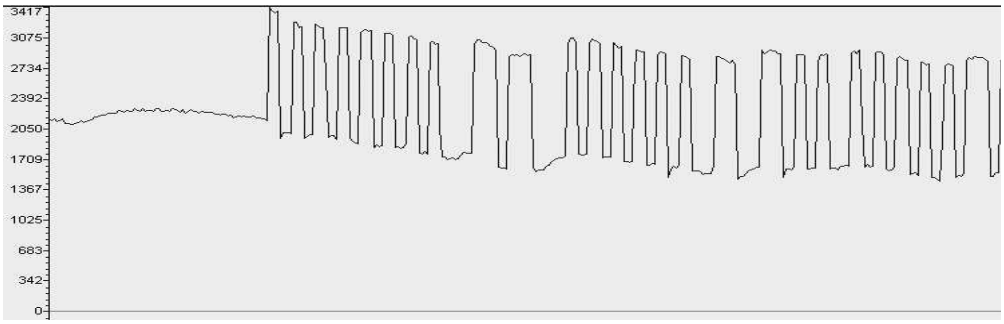


Figure 3: Figure for ADC sampling data.

5. CONCLUSION

We have introduced the detailed structure of a UHF RFID reader that functions well experimentally. For the receiver chain, Zero-IF demodulation is incorporated, as well as quadrature I/Q channels. The designed system proves a good operational structure as data signals have been received and exposed as the above schematics.

ACKNOWLEDGMENT

This work is supported by ZJNSF R105253, and in part by NSFC 60671003 and 60531020.

REFERENCES

1. Kossel, M., H. R. Benedickter, R. Peter, and W. Bachtold, "Microwave backscatter modulation systems," *Microwave Symposium Digest, 2000 IEEE MTT-S International*, Vol. 3, 11–16 June, 2000.
2. Boles, T., R. Houlihan, and C. Weigand, "An HMIC I/Q modulator/demodulator for RFID applications," *Microwave Conference Proceedings, 2005. APMC 2005. Asia-Pacific Conference Proceedings*, Vol. 2, 4–7 Dec., 2005.
3. Roberts, C. M., "Radio frequency identification (RFID)," *Computers & Security*, 25, 18–26, 2006.
4. Droitcour, A. D., O. Boric-Lubecke, V. M. Lubecke, J. Lin, and G. T. A. Kovacs, "Range correlation and I=Q performance benefits in single-chip silicon doppler radars for noncontact cardiopulmonary monitoring," *IEEE Trans. on Microwave Theory and Techniques*, Vol. 52, No. 3, March 2004.

Effects of Giant Optical Anisotropy in R-plane GaN/AlGaN Quantum Wells by Valence Band Mixing

Chun-Nan Chen, **Kao-Feng Yarn**
Department of Electronic Engineering, Far East University
Hsin-Shih, Tainan 744, Taiwan, China

Win-Jet Luo
Department of Refrigeration and Air Conditioning
Chin-Yi University of Technology, Taichung, Taiwan, China

Jih-Chen Chiang, **Ikai Lo**, **Wan-Tsang Wang**, **Ming-Hong Gau**, **and Hsiu-Fen Kao**
Department of Physics and Center for Nanoscience and Nanotechnology
Sun Yat-Sen University, Kaohsiung, Taiwan, China

Meng-En Lee
Department of Physics, Kaohsiung Normal University
Kaohsiung Country, Taiwan, China

Wei-Ching Chuang
Department of Electro-Optics Engineering, Formosa University, Yulin, Taiwan, China

Wen-Chung Chang and Tsung-Chan Cheng
Department of Electronic Engineering, Southern Taiwan University of Technology, Tainan, Taiwan, China

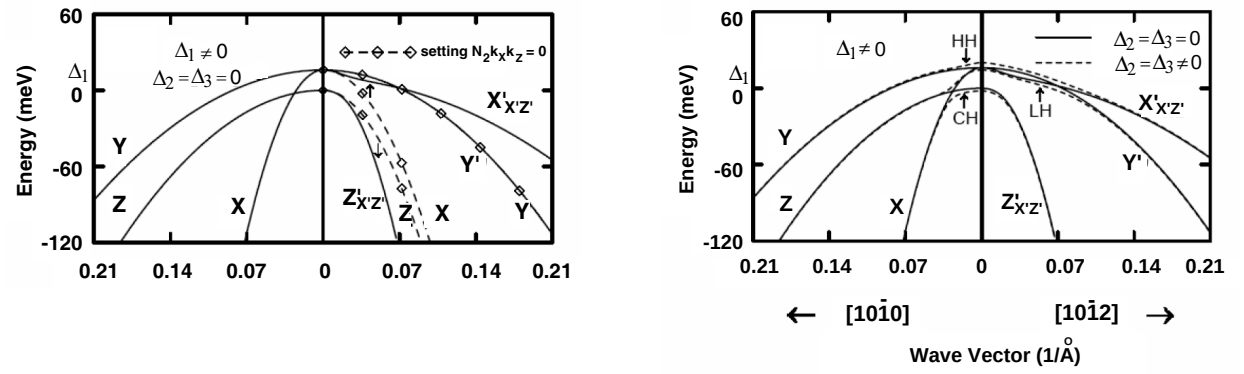
Abstract— Investigation of optical anisotropy spectra in the R-plane (i. e., the $[10\bar{1}2]$-oriented layer plane) of $GaN/Al_{0.2}Ga_{0.8}N$ quantum wells with different widths is studied. The optical matrix elements in the wurtzite quantum wells are calculated using the **k·p** finite difference scheme. The calculations show that the valence band mixing effect produces giant in-plane optical anisotropy in $[10\bar{1}2]$-oriented $GaN/Al_{0.2}Ga_{0.8}N$ quantum wells with a narrow width. The nature of the in-plane optical anisotropy is found to be dependent on the well width. Specifically, it is found that the anisotropy changes from x'-polarization to y'-polarization as the well width increases.

1. INTRODUCTION

GaN-based semiconductors have received increasing attention in the past decade as a result of their unique properties and potential applications in the electronics and optoelectronics fields. [1–5] Recent advances in crystal growth techniques now enable the fabrication of high-quality |||-nitride based semiconductor heterostructures on substrates with orientations other than the conventional [0001] direction. [6, 7] The optical and electronic properties of such semiconductor heterostructures are quite different from those of semiconductor crystals grown on conventionally orientated substrates. For example, Rau, et al. [8] and Sun, et al. [9] identified the existence of giant in-plane optical anisotropy (over 90%) in non-polar |||-nitride based wurtzite quantum wells, while Sharma, et al. [10] observed weaker optical anisotropy (∼32%) in semi-polar |||-nitride based wurtzite quantum wells. This study employs an arbitrarily-oriented [hkil] Hamiltonian potential matrix to conduct a comprehensive investigation into in-plane optical anisotropy in semi-polar wurtzite quantum wells. The optical anisotropy spectrum is calculated for semi-polar $[10\bar{1}2]$-oriented $GaN/Al_{0.2}Ga_{0.8}N$ quantum wells of different widths. It is shown that the valence band mixing (VBM) effect results in giant optical anisotropy in thin $[10\bar{1}2]$-oriented $GaN/Al_{0.2}Ga_{0.8}N$ quantum wells. The physical origin of the giant in-plane optical anisotropy (over 90%) observed in non-polar quantum wells is discussed elsewhere, and is therefore not considered here. The current optical transition calculations are based on the formalism introduced by Lew Yan Voon and Ram-Mohan [11] and are solved using the **k·p** finite difference scheme. [12, 13] In the calculations, the directions of the x-, y- and z-axes are assumed to lie along the $[10\bar{1}0]$, $[12\bar{1}0]$ and [0001] directions, respectively. Furthermore, the three unit vectors in the prime coordinates (x', y', z') are given by $\hat{y}' = \hat{y}$, $\hat{z}' = \hat{z}\cos\theta + \hat{x}\sin\theta$ and $\hat{x}' = \hat{x}\cos\theta - \hat{z}\sin\theta$, respectively, where θ is the polar angle and $\hat{z}'$ is the growth direction. Note that the polar angles $\theta = 0°$, $43.2°$ and $90°$ correspond to the $z' = [0001]$, $[10\bar{1}2]$ and $[10\bar{1}0]$ growth directions, respectively.

2. RESULTS AND DISCUSSION

Figure 1(a) shows the GaN valence band structures along the $[10\bar{1}0]$- and $[10\bar{1}2]$-directions without taking the spin-orbit interaction into consideration (i. e., $\Delta_2 = \Delta_3 = 0$). In this figure, the growth direction is chosen to be $[10\bar{1}2]$, and hence in the prime coordinate system (x', y', z'), $z' = [10\bar{1}2]$ and $k_{z'}//[10\bar{1}2]$. It is observed that the valence bands along the $[10\bar{1}0]$ direction are pure $|Y\rangle$- (labeled Y), pure $|Z\rangle$- (labeled Z) and pure $|X\rangle$- (labeled X) states, respectively. These pure states arises because the off-diagonal terms (i. e., $N_1 k_x k_y$, $N_2 k_x k_z$ and $N_2 k_y k_z$) in the 3×3 valence band Hamiltonian given in Eq. (26) of Ref. [14] are all equal to zero in the $[10\bar{1}0]$ direction. [14] However, in the $[10\bar{1}2]$ direction, the off-diagonal term, i. e., $N_2 k_x k_z$, is not equal to zero, and hence mixing occurs between the $|X\rangle$- and $|Z\rangle$-bands. Prior to mixing, as shown by the dashed lines with open squared symbol, the $|X\rangle$-band (labeled X) and the $|Z\rangle$-band (labeled Z) have a lower energy than the $|Y\rangle$-band (labeled Y). However, after mixing (please refer to the solid lines), the upper X-Z mixing band (labeled $X'_{X'Z'}$) has a higher energy than the $|Y\rangle$-band (labeled Y') when the wave vector $k_{z'}$ is greater than $\sim 0.07(1/\text{Å})$, because the coupling term, i. e., $N_2 k_x k_z$, increases rapidly with increasing $k_{z'}$. In this figure, the upper (lower) X-Z mixing band is labeled as $X'_{X'Z'}(Z'_{X'Z'})$ because it is not only an $|X\rangle$-like ($|Z\rangle$-like) X-Z mixing band, but also an $|X'\rangle$-like ($|Z'\rangle$-like) X'-$'$mixing band. Additionally, the $|Y\rangle$-band is labeled as Y' because the $|Y\rangle$- and $|Y'\rangle$-bands are identical. Clearly, the X-Z valence band mixing leads to a crossover of the Y' and X'_{XZ} bands in the $[10\bar{1}2]$ direction. The dashed lines in Fig. 1(b) show the GaN valence band structures along the $[10\bar{1}0]$- and $[10\bar{1}2]$-directions when the spin-orbital interaction is taken into consideration (i. e., $\Delta_2 = \Delta_3 \neq 0$). Here for convenience, the valence bands are labeled (from top to bottom) as the heavy-hole (HH) band, the light-hole (LH) band, and the crystal-field (CH) band, respectively. It is seen that strong mixing between the Y' and $X'_{X'Z'}$ bands near their crossing point causes the HH (LH) band along the $[10\bar{1}2]$-direction to change from a $|Y'\rangle$-like ($|X'\rangle$-like) state to an $|X'\rangle$-like ($|Y'\rangle$-like) state as the wave vector k increases. Additionally, it is also seen that the $X'\rangle$-like HH and $|Y'\rangle$-like LH bands are widely separated when $k_{z'}$ is larger than $\sim 0.1(1/\text{Å})$. In this study, these phenomena are referred to as the valence band mixing (VBM) effect. As discussed later, the VBM effect implies that strong in-plane x'-polarization anisotropy exists in a $[10\bar{1}2]$-oriented unstrained quantum well with a narrow well width (L) (e. g., $L < 30\text{Å}$).



(a) (b)

Figure 1: (a) Valence band structures of $[10\bar{1}2]$-oriented GaN crystal for $\Delta_2 = \Delta_3 = 0$ (solid lines). In comparison with the solid lines, the dashed lines with open squared symbols show the valance band structures when the coupling between the $|X'\rangle$- and $|Z'\rangle$-states is not included (i. e., assuming $N_2 k_X k_Z = 0$), (b) Valence band structures of $[10\bar{1}2]$-oriented GaN crystal for $\Delta_2 = \Delta_3 \neq 0$ (dashed lines) and $\Delta_2 = \Delta_3 = 0$ (solid lines).

Figure 2 shows the in-plane (i. e., $x' - y'$ plane) valence subband structures of a $[10\bar{1}2]$-oriented $GaN/Al_{0.2}Ga_{0.8}N$ quantum well with a well width of $L = 15\text{Å}$. Reading from top to bottom, the valence subbands correspond to the first heavy-hole (HH1) band, the first light-hole (LH1) band, the second heavy-hole (HH2) band, and the first crystal-field (CH1) band, respectively. It is observed that near the Γ-point, the HH1 band is a strongly $|X'\rangle$-like symmetric $|X'^S\rangle$-like) state, the LH1 band is a strongly $|Y'\rangle$-like symmetric ($|Y'^S$-like) state, the HH2 band is a strongly $|X'\rangle$-

like asymmetric ($|X'^A\rangle$-like) state, and the CH1 band is a strongly $|Z'\rangle$-like symmetric ($|Z'^S\rangle$-like) state. Meanwhile, in the $k_{x'}$ direction, the Γ-point effective mass of the HH2 band is negative and the Γ-point effective mass of the HH1 band is lighter than that of the LH1 band. As a result, strong 2-band mixings of (X'^S-Y'^S) and (X'^S-X'^A) occur near $k_x = 0.03(1/\text{Å})$ and $0.07(1/\text{Å})$, respectively. Consequently, as $k_{x'}$ increases, the HH1 band changes from an $|X'^s\rangle$-like state to a $|Z'^S\rangle$-like state, while the LH1 band transits from a $|Z'^S\rangle$-like state, through an X'^s-state, to an $|X'^A\rangle$-like state. In the present study, this phenomenon is referred to as the quantum-well valence-band mixing (QWVBM) effect. Similarly, the QWVBM effect leads to 2-band mixing of X'^A-Y'^S near $k_{y'} = 0.07(1/\text{Å})$. As $k_{y'}$ increases, the LH1 band therefore changes from a $|Y'^s\rangle$-like state to an $|X'^A\rangle$-like state. However, it is observed that the HH1 band remains in an $|X'^s\rangle$-like state.

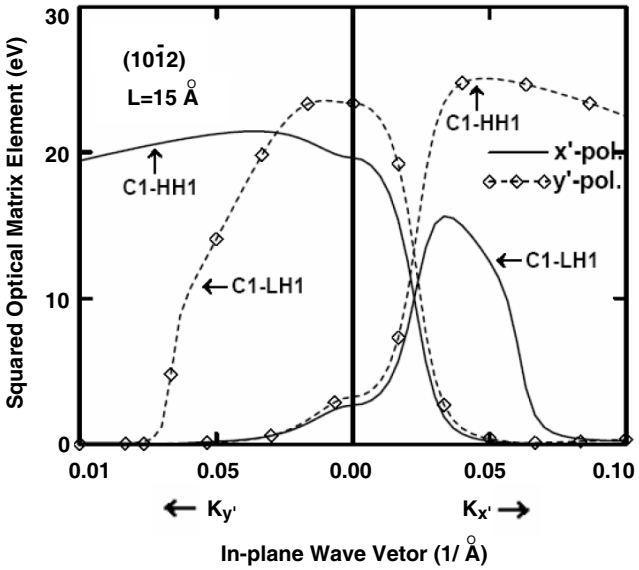


Figure 2: In-plane (i. e., x'-y' plane valence subband structure of [10$\bar{1}$2]-oriented GaN/$Al_{0.2}Ga_{0.8}$N quantum well with 15Å well width.

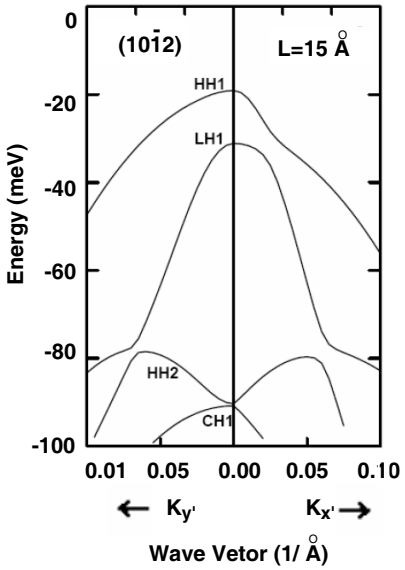


Figure 3: Squared optical matrix elements of in-plane (i. e., x'-y' plane) optical transitions for [10$\bar{1}$2]-oriented GaN/$Al_{0.2}Ga_{0.8}$N quantum well with 15Å well width.

Figure 3 shows $|M_\alpha|^2_{c1-v1}$ as a function of the in-plane (i. e., the x'-y' plane) wave vector for a [10$\bar{1}$2]-oriented GaN/$Al_{0.2}Ga_{0.8}$N quantum well with a well width of $L = 15$Å, where $|M_\alpha|^2_{c1-v1}$ denotes the squared optical matrix elements for an α-polarization ($\alpha = x'$ and y')transition from the first conduction subband ($C1$) to $v1$ valence subbands (where $v1$ =HH1, LH1 and CH1, respectively). Note that $|M_\alpha|^2_{C1-CH1}$ ($\alpha = x'$, y') is not shown in Fig. 3, mainly because the energy of the CH1 band is far less than that of either the HH1 band or the LH1 band and therefore it does not contribute to the optical properties of the GaN/$Al_{0.2}Ga_{0.8}$N quantum well near the zone center, the strongly $|X'^s\rangle$-like HH1 ($|Y'^s\rangle$-like LH1) subband is a weak X'^S-Y'^S mixing band, and hence $|M_{x'}|^2_{C1-HH1}$ ($|M_{y'}|^2_{C1-LH1}$)is much larger than $|M_{y'}|^2_{C1-HH1}$ ($|M_{x'}|^2_{C1-LH1}$). This indicates that the weak X'^S-Y'^S mixing causes the photoluminescence generated by the C1-HH1 (C1-LH1) transition to be strongly polarized along the x'-axis (y'-axis). Certainly, the X'^S-Y'^S mixing becomes weaker as the energy separation between the HH1 and LH1 bands increases. In a thin [10$\bar{1}$2]-oriented wurtzite GaN unstrained quantum well ($L < 30$Å), the VBM_{X^s} effect results in a large energy difference between the $|X'^s\rangle$-like HH1 band and the $|Y'^s\rangle$-like LH1 band near the zone center, and this implies the presence of strong in-plane x'-polarization anisotropy. In Fig. 3, as $k_{x'}$ increases toward the 2-band (X'^s_2-Y'^s) mixing region ($k_{x'} \approx 0.03(1/\text{Å})_2$), the QWVBM ($X'^S$-$Y'^S$ mixing) effect causes $|M_{x'}|^2_{C1-HH1}$ and $|M_{y'}|^2_{C1-LH1}$ to fall rapidly to zero and $|M_{y'}|^2_{C1-LH1}$ to increase to its maximum value. Similarly, the QWVBM effect causes $|M_{y'}|^2_{C1-LH1}$ to vanish for $k > 0.07(1/\text{Å})$ and to assume a single-peak structure with a peak near $k_{x'} = 0.03(1/\text{Å})$.

Figure 4 shows the in-plane optical anisotropies at the zone center (i. e., $\rho(k_{x'} = k_{y'} = 0)$) of the z'-oriented GaN/ $Al_{0.2}Ga_{0.8}N$ quantum wells. Note that the in-plane optical anisotropy, ρ, is defined as

$$\rho = \frac{|M_{x'}|^2_{C1-HH1} - |M_{y'}|^2_{C1-HH1}}{|M_{x'}|^2_{C1-HH1} + |M_{y'}|^2_{C1-HH1}}$$

It can be seen that the largest optical anisotropy (y'-polarization) occurs in the $[10\bar{1}0]$-oriented quantum wells, while subsidiary anisotropy (x'-polarization) appears in the thin $[10\bar{1}2]$-oriented quantum wells ($L < 30$Å). The strong anisotropy in the thin $[10\bar{1}2]$-oriented quantum wells is x'-polarized and is the result of the VBM effect. However, the strong anisotropy in the $[10\bar{1}0]$-quantum wells is y'-polarized and is caused by the crystal field effect (which will be discussed elsewhere). Most interestingly, in the $[10\bar{1}2]$-oriented quantum wells, it is observed that as the well width increases and reaches a critical value of L_C ($L_C \approx 30$Å), the in-plane anisotropy changes from x'-polarization to y'-polarization. Again, this phenomenon is caused by the VBM effect (i. e., the HH band shown in Fig. 1 changes from an $|X'\rangle$-like state to a $|Y'\rangle$-like state as the wave vector $k_{z'}$ decreases). The critical value, L_C, decreases (increases) with increasing compressive (tensile) strain since compressive (tensile) strain enlarges (reduces) the crystal field split energy and therefore weakens (enhances) the coupling between the $|X\rangle$-and $|Z\rangle$-bands (refer to Fig. 1). Clearly, when the compressive (tensile) strain is sufficiently high, only y'-polarization (x'-polarization) anisotropy exists. Finally, in-plane optical anisotropy is not found in the [0001]-oriented quantum wells because this particular quantum well belongs to the D_{6h} high-symmetry point group. The results presented in this study are in good agreement with the experimental findings reported in the literature. For example, Rau, et al. [8] and Sun, et al. [9] identified the presence of giant in-plane optical anisotropy (over 90%) in non-polar |||-nitride based wurtzite quantum wells, while Sharma et al. [10] observed weaker optical anisotropy (∼32%) in semi-polar |||-nitride based wurtzite quantum wells.

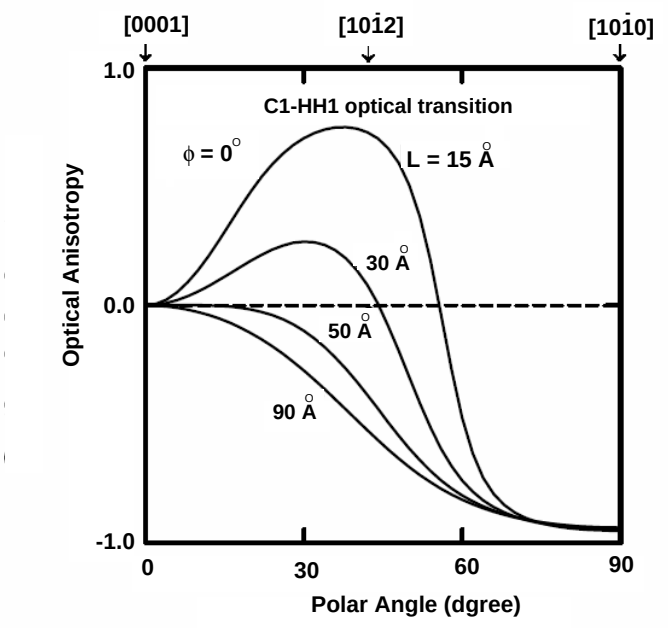


Figure 4: Well-width dependence of optical anisotropy for C1-HH1 transition in $GaN/Al_{0.2}Ga_{0.8}N$ quantum wells as function of substrate orientation (1011)with $1 = 0 \sim \infty$ (polar angle $\theta = 90^\circ \sim 0^\circ$ and azimuthal angle $\varphi = 0^\circ$).

3. CONCLUSION

In conclusion, this study has calculated the optical transition of $GaN/Al_{0.2}Ga_{0.8}N$ wurtzite quantum wells of different well widths using the $\mathbf{k}\cdot\mathbf{p}$ finite difference scheme. The results have shown that giant in-plane optical anisotropy exists in the R-plane (i. e., $[10\bar{1}2]$-oriented layer plane) of $GaN/Al_{0.2}Ga_{0.8}N$ quantum wells with a narrow width as a result of the valence band mixing effect. The anisotropy in $[10\bar{1}2]$-oriented quantum wells is width dependent, i. e., the in-plane anisotropy

changes from x'-polarization to y'-polarization as the well width increases. The results presented in this study are in good agreement with the experimental findings reported in the literature and provide valuable guidelines for the design of polarization stabilization devices based on polarization control or selectivity.

REFERENCES

1. Morkoç, H. and S. N. Mohammad, *Science*, Vol. 267, 51, 1995.
2. Akasaki, I. and H. Amano, *J. Cryst. Growth*, Vol. 175–176, 29, 1997.
3. Nakamura, S., *Science*, Vol. 281, 956, 1998.
4. Lo, I., W. T. Wang, M. H. Gau, S. F. Tsay, and J. C. Chiang, *Phys. Rev. B*, Vol. 72, 245329, 2005.
5. Lo, I., W. T. Wang, M. H. Gau, S. F. Tsay, and J. C. Chiang, *Appl. Phys. Lett.*, Vol. 88, 82108, 2006.
6. Hwang, S. J., W. Shan, R. J. Hauenstein, and J. J. Song, *Appl. Phys. Lett.*, Vol. 64, 2928, 1994.
7. Khan, M. A., R. A. Skogman, and J. M. van Hove, *Appl. Phys. Lett.*, Vol. 56, 1257, 1990.
8. Rau, B., P. Waltereit, O. Brandt, M. Ramsteiner, K. H. Ploog, J. Puls, and F. Henneberger, *Appl. Phys. Lett.*, Vol. 77, 3343, 2000.
9. Sun, Y. J., O. Brandt, M. Ramsteiner, H. T. Grahn, and K. H. Ploog, *Appl. Phys. Lett.*, Vol. 82, 3850, 2003.
10. Sharma, R., P. M. Pattison, H. Masui, R. M. Farrell, T. J. Baker, B. A. Haskell, F. Wu, S. P. DenBaars, J. S. Speck, and S. Nakamura, *Appl. Phys. Lett.*, Vol. 87, 231110-1, 2005.
11. Lew Yan Voon, L. C. and L. R. Ram-Mohan, *Phys. Rev. B*, Vol. 47, 15500, 1993.
12. Chen, C. N., *Phys. Rev. B*, Vol. 72, 085305, 2005.
13. Chen, C. N., *J. Appl. Phys.*, Vol. 96, 7374, 2004.
14. Chuang, S. L. and C. S. Chang, *Phys. Rev. B*, Vol. 54, 2491, 1996.

Monitoring the Earth, Ocean, and Atmosphere with Hyperspectral Remote Sensors

Daniel K. Zhou
NASA Langley Research Center, Hampton, VA 23681, USA

Abstract— The success of the NAST-I program enables us to understand the great benefits achieved with current and future hyperspectral remote sensors. Validated retrieval algorithms and results demonstrate the advancement of hyperspectral remote sensing capabilities in monitoring the Earth, ocean, and atmosphere.

1. INTRODUCTION

The NPOESS (National Polar-orbiting Operational Environmental Satellite System) Airborne Sounder Testbed-Interferometer (NAST-I) is designed to support the development of future satellite temperature and moisture sounders. The NAST was developed in 1998 and it has been collecting data through out numerous field campaigns. The methodology for thermodynamic parameters retrieved under cloud-free conditions was developed and validated; the algorithm was used to demonstrate that an improved moisture profile is retrieved from hyperspectral radiances [4]. Accurate retrieval of the surface properties (i.e., surface emissivity and temperature) is simultaneously achieved in order to obtained accurate terrestrial boundary layer (TBL) moisture. In addition to thermodynamics parameters retrieved, atmospheric trace gas (e.g., CO) retrieval is critical in air quality observation, modeling, and forecasting. NAST-I CO retrieval algorithm was developed to demonstrate the retrieval sensitivity and accuracy of this type of hyperspectral sounders under cloud-free conditions [5–7]. The globe hosts many vast cloudy regions and a great deal of effort has gone into cloud detection and cloud-clearing processes [2]. Nevertheless, the schemes dealing with cloud detection and cloud-clearing [1] remain a major source of error in the final retrieval products. Recently, fast molecular and cloud transmittance models [3] have been used to enable the infrared radiances to be used under cloudy conditions with the accuracy required for sounding retrieval processing [5, 6, 8]. With this recently developed algorithm, cloud parameters as well as atmospheric profiles are retrieved simultaneously from infrared spectral radiance observations. A large amount of samples, retrieved and validated, are used to demonstrate a state-of-art retrieval algorithm based on the NAST-I hyperspectral data and how well this algorithm applies to the AIRS data. Experiences with the NAST-I and AIRS retrieval put us in a position of confidence that accurate retrievals will be obtained from future satellite hyperspectral instruments such as the IASI, the CrIS, the GIFTS, and the HES.

2. RETRIEVAL DEMONSTRATION

High-resolution infrared radiance spectra obtained from near nadir observations provide atmospheric, surface, and cloud property information. A fast radiative transfer model, including cloud effects, is used for atmospheric profile and cloud parameter retrieval. The retrieval algorithm is presented along with its application to field experiment data from the NPOESS Airborne Sounder Testbed-Interferometer (NAST-I). The retrieval accuracy dependence on cloud properties is discussed. The initial EOF regression has laid a first step in dealing with infrared sounding data under cloudy conditions which is now significantly improved by the physical iteration inversion described in this study. Results achieved with airborne NAST-I observations show that accuracies close to those achieved in totally cloud-free conditions can be achieved down to cloud top levels. The accuracy of the profile retrieved below cloud top level is dependent upon the optical thickness and fractional coverage of the clouds. Retrieval accuracy of temperature and moisture profiles is greatly improved by the physical inversion. Details can be found elsewhere [5, 6, 8]. The physical retrieval results of cloud and thermodynamic parameters are shown in Figure 1 (after [8]). Results achieved with airborne NAST-I observations show that accuracies close to those achieved in totally cloud-free conditions can be achieved down to cloud top levels. The accuracy of the profile retrieved below cloud top level is dependent upon the optical thickness and fractional coverage of the clouds. Retrieval accuracy of temperature and moisture profiles is greatly improved by the physical inversion as shown by dropsonde validation. Therefore, the radiances can be accurately simulated by using physically inverted results which is important for direct assimilation into a forecast model.

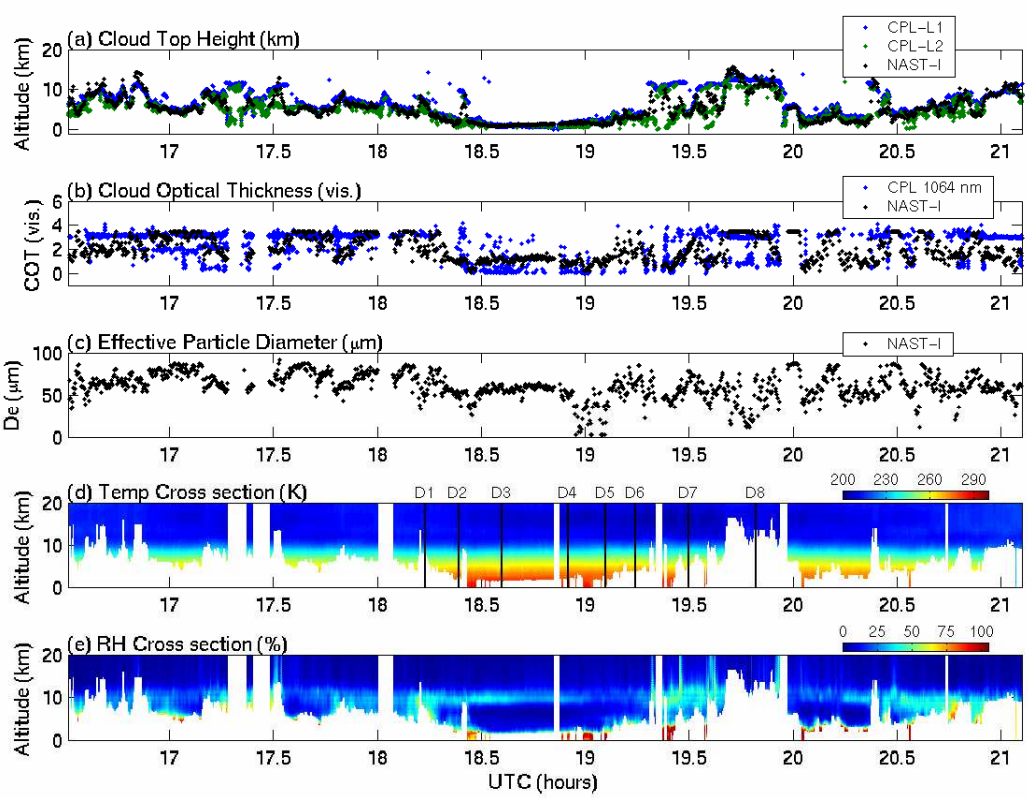


Figure 1: Panel (a) plots NAST-I physically retrieved cloud top height compared with the CPL measured cloud top heights of the top 2 layers (L1 and L2). Panel (b) plots NAST-I retrieved visible cloud optical thickness (COT) compared with the CPL measurement. Panel (c) plots NAST-I retrieved cloud particle size. Panels (d) and (e) plot NAST-I physically retrieved temperature and relative humidity vertical cross sections, respectively. The areas wiped off are under the top layer of clouds where the cloud visible optical thickness is larger than one and under the lower "opaque" cloud. The black vertical bars in panel (d) indicate dropsonde locations.

The Atmospheric InfraRed Sounder (AIRS) on the Earth Observing System (EOS) Aqua satellite was launched on 4 May 2002. Experience with NAST-I greatly helps the development of AIRS satellite data retrieval. The NAST retrieval algorithm has been used for the AIRS team algorithm validation [9]. As the satellite data covers a large variety of surface and atmospheric conditions, the retrieval scheme is tested with a large variety of conditions. Initial study with AIRS data has been performed and more detailed validation over the land will be conducted. One granule of AIRS data (10 September 2004; ∼01:00 UTC; local nighttime), shown in Figure 2, covers water and land including the vicinity of the Sahara Desert. The AIRS single field of view (∼13.5 km at nadir) data are used; few cloudy spots were seen as the effective skin temperature contains cloud features (i.e., cooler "skin temperature"). However, this granule, for the most part, was collected under cloud-free conditions. The distribution of the surface emissivity images near 11 μm and 8.15 μm are plotted, capturing the feature of the surface emissivity variation. The spectacular features over the land, especially in the vicinity of the Sahara Desert, are clearly evident. Typical emissivity spectra retrieved from water, land, and desert are plotted in comparison with laboratory measurements. These reasonable retrieved surface properties greatly support the accurate atmospheric retrievals which are found elsewhere [9].

The NAST-I field campaign data are also used to provide radiometric measurements including tropospheric trace species such as carbon monoxide (CO). The NAST-I CO inversion scheme is combined with a three-step procedure: (1) EOF (i.e., empirical orthogonal function) regression retrieval, (2) simultaneous matrix inversion, and (3) CO profile enhancement inversion. The NAST-I tropospheric CO retrieval algorithm has been developed to retrieve CO from NAST-I radiances and to investigate tropospheric CO vertical profile retrieval accuracy from a satellite ultraspectral sounder. The CO evaluation study having co-incident radiance and in-situ measurement data sets

enables an understanding of the accuracy of our current CO retrieval algorithm and validates the results based on theoretical simulations [5, 6]. Examples of CO retrievals from the EAQUATE [7] indicate that the CO first guess profile plays a major role in the CO profile accuracy. The TBL CO accuracy, mainly determined by the first guess, affects the free tropospheric CO retrieval. However, CO variations in the free troposphere can still be captured while CO amount variations in the TBL are not retrieved very well. Thus, it is a challenge to obtain accurate CO profiles (especially in the TBL) from remote sounders such as NAST-I.

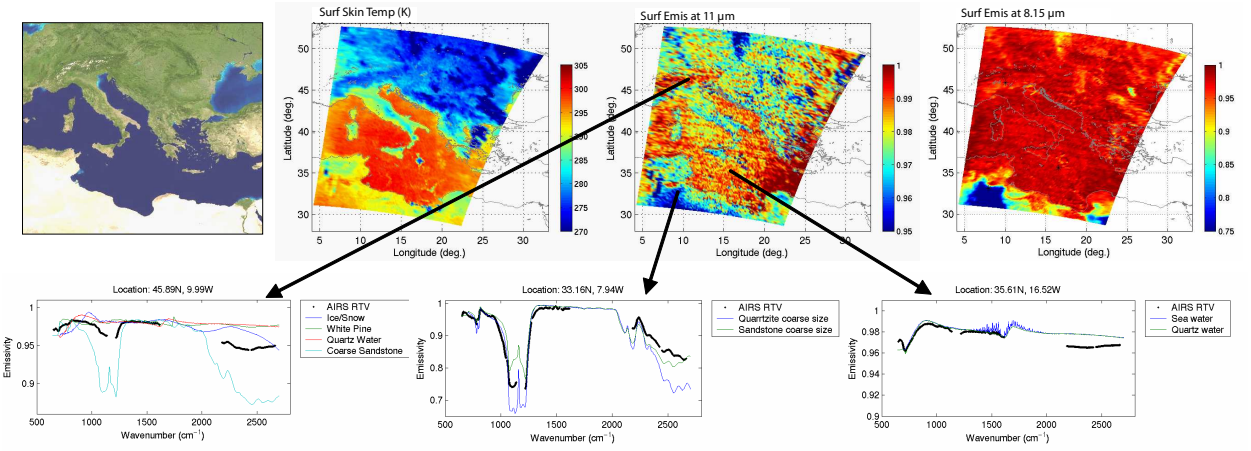


Figure 2: Surface skin temperature and emissivity retrieved from the AIRS data, indicating a large variety of surface types is captured by the hyperspectral data and this inversion scheme.

3. CONCLUSION

The advantage of using NAST-I is that continuous spectrum of atmospheric radiation is obtained with a high spatial resolution as needed to develop retrieval algorithms. Lessons learned from the NAST program are beneficial to current and future satellite hyperspectral instruments such as AIRS, IASI, CrIS, GIFTS, and HES.

ACKNOWLEDGMENT

The NAST-I program is supported by the NPOESS Integrated Program Office, NASA Headquarters, and NASA Langley Research Center. The authors express their sincere thanks to the NAST-I team members from various organizations.

REFERENCES

1. Smith, W. L., "An improved method for calculating tropospheric temperature and moisture from satellite radiometer measurements," *Mon. Wea. Rev.*, Vol. 96, 387–396, 1968.
2. Smith, W. L., D. K. Zhou, H.-L. Huang, J. Li, X. Liu, and A. M. Larar, "Extraction of profile information from cloud contaminated radiances," paper presented at the *ECMWF Workshop on the Assimilation of High Spectral Resolution Sounders in NWP*, June 28–July 1, 2004.
3. Yang, P., B. C. Gao, B. A. Baum, Y. Hu, W. Wiscombe, S.-C. Tsay, D. M. Winker, and S. L. Nasiri, "Radiative properties of cirrus clouds in the infrared (8–13 μm) spectral region," *J. Quant. Spectros. Radiat. Transfer*, Vol. 70, 473-504, 2001.
4. Zhou, D. K., W. L. Smith, J. Li, H. B. Howell, G. W. Cantwell, A. M. Larar, R. O. Knuteson, D. C. Tobin, H. E. Revercomb, and S. A. Mango, "Thermodynamic product retrieval methodology for NAST-I and validation," *Appl. Opt.*, Vol. 41, 6957–6967, 2002.
5. Zhou, D. K., W. L. Smith, X. Liu, J. Li, A. M. Larar, and S. A. Mango, "Tropospheric CO observed with the NAST-I: retrieval methodology, analyses, and first results," *Appl. Opt.*, Vol. 44, 3032–3044, 2005.
6. Zhou, D. K., W. L. Smith, X. Liu, A. M. Larar, H.-L. A. Huang, J. Li, M. J. McGill, and S. A. Mango, "Thermodynamic and cloud parameters retrieval using infrared spectral data," *Geophys. Res. Lett.*, Vol. 32, L15805, doi:10.1029/2005GL023211, 2005.

7. Zhou, D. K., A. M. Larar, X. Liu, G. W. Sachse, W. L. Smith, J. P. Taylor, S. M. Newman, and S. A. Mango, "NAST-I tropospheric CO retrieval validation during INTEX-NA and EAQUATE," submitted to the *Q. J. R. Meteorol. Soc.* (available upon request), 2007.
8. Zhou, D. K., W. L. Smith, Sr., X. Liu, A. M. Larar, S. A. Mango, and H.-L. Huang, "Physically-retrieving cloud and thermodynamic parameters from ultraspectral IR measurements," *J. Atmos. Sci.*, in press, 2007.
9. Zhou, D. K., W. L. Smith, V. Cuomo, J. P. Taylor, C. D. Barnet, P. Di Girolamo, G. Pappalardo, A. M. Larar, X. Liu, S. M. Newman, C. Lee, and S. A. Mango, "Retrieval validation during the European AQUA Thermodynamic Experiment," submitted to the *Q. J. R. Meteorol. Soc.* (available upon request), 2007.

Two-dimensional Cross Embedded Metamaterials

J. Zhang[1,2], **H. Chen**[1,2], **L. Ran**[1,2], **Y. Luo**[1,2], **and J. A. Kong**[1,3]

[1]The Electromagentics 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, Massachusetts 02139, USA

Abstract— Traditional two dimensional (2D) left-handed metamaterials were composed of honeycomb structures. In this paper, we experimentally realized a 2D S-shaped metamaterial structure with cross embedded form. We show that metamaterials with cross embedded arrangements exhibit wider left-handed band and lower loss than honeycomb structures.

1. INTRODUCTION

In 2001, left-handed metamaterial was experimentally realized [1] by using metamaterial fabricated with periodical patterns composed of a unit split-ring resonator (SRR) and conducting wire. Since then, a variety of designs have been proposed and studied, such as the Ω-like structure, S-shaped resonators, while transmission line structures and photonic crystals have also been shown to exhibit left-handed properties. The general problem with the one-dimensional (1D) LHM (left handed metamaterials) is the inherent asymmetry of the resulting lattice of structures. R. A. Shelby has achieved 2D isotropy by placing the SRRs along two orthogonal axes in a lattice [2]. And so far the two dimensional LHM are usually realized by arranging the rings into a honeycomb structure as R. A. Shelby did. For this kind of arrangement, the loss is very high and the bandwidth is very small, because the interaction between the rings in two directions is very large. Some authors proposed some embedded SRR to achieve 2D or 3D isotropic resonators. Because the sizes of the rings in the two dimensions are different, it leads to different macroscopic properties of the material. Besides, the arrangements of the rods haven't been considered, these structures only show negative permeability behavior.

In this paper, we proposed a new type of two dimensional left-handed metamaterial with a cross embedded form of the S-ring resonators. The rings in the two dimensions have the same size, shows more isotropic properties in the two dimensions. The transmission experiment are carried out, and the results are compared with the one realized with honeycomb form. We show a wider bandwidth and a lower loss is achieved, showing the superiority of this arrangement. The reason is that there is less coupling effect between the units in the two dimensions.

2. 2D S-RING STRUCTURES

The elementary cell of the metamaterial is based on an extended S-ring resonator, which is composed of a substrate of height $L = 49.7\,\text{mm}$ in the y direction and of thickness $d = 2\,\text{mm}$, as shown in Fig. 1. The relative permittivity of the substrate is 4.0. The extended S-shaped metallic strips are printed on both sides of the substrate, and are a mirror image of each other. The geometry is formed by linking two S-shaped structures of the same size, and the other dimensions of the sample are: $c = 2.1\,\text{mm}$, $w1 = 6.3\,\text{mm}$, $w2 = 13.5\,\text{mm}$, $h = 14.7\,\text{mm}$, $H = 17.5\,\text{mm}$.

Firstly, we arrange the one dimensional S-shaped unit cell in a traditional honeycomb structure by placing the S-shaped resonator along two orthogonal axis in a lattice, which forms a two dimensional metamaterial, as shown in Fig. 2(a). The period of each cell is 5 mm. The commercial software CST Microwave Studio has been used for the simulation of the structure, which repeats periodically in x and z directions. The electromagnetic wave is incident in z direction with an electric field polarized in y direction. The S-parameters have been extracted for the 0–10 GHz frequency range. From the results obtained using CST and shown in Fig. 3, it can be noted that there is a narrow passband around 2.34 GHz, with a loss of about 10.1 dBm. Using the retrieval method [4], we find the frequency range of 2.325–2.365 GHz is a left-handed passband, in which the effective permittivity and permeability are both negative. The result is shown in Fig. 4, from which a relative bandwidth of 1.7% could be identified.

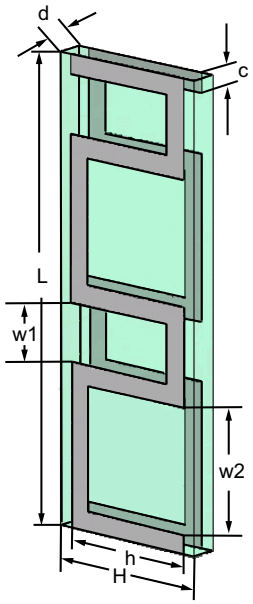


Figure 1: Dimensions characterizing the one dimensional unit cell of the metamaterial (c is constant for all metallization).

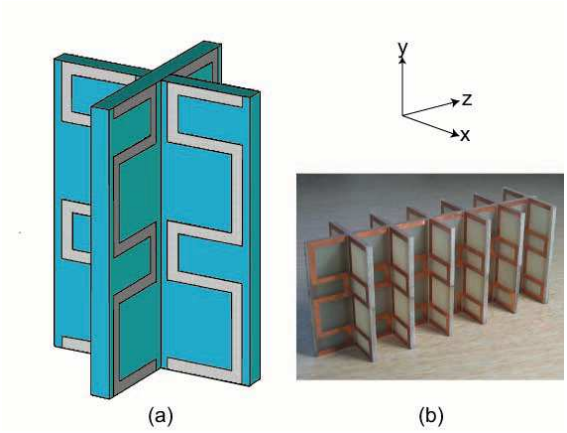


Figure 2: (a) The two dimensional metamaterial arranged as a square lattice, (b) Picture of the metamaterial used in the experimental measurements which is arranged in a honeycomb structure.

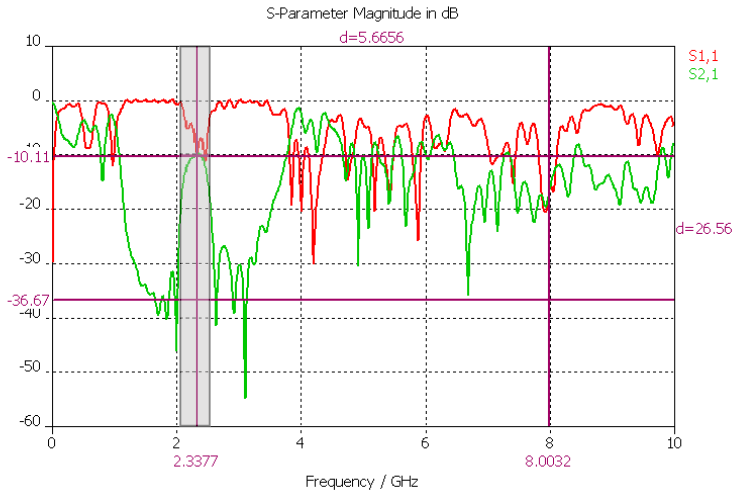


Figure 3: Simulated S-parameters of the two dimensional metamaterial arranged in the honeycomb structure.

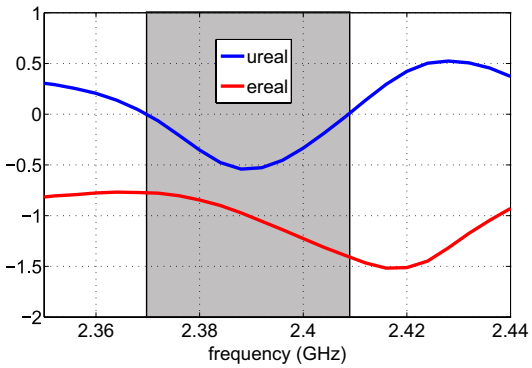


Figure 4: Effective permittivity and permeability of the 2D LHM of honeycomb structure by retrieval method using the simulated s parameters.

In comparison, we next arrange the extended one dimensional S-shaped ring in a cross embedded structure to form a two dimensional metamaterial, as shown in Fig. 5(a). The structure is repeated periodically along x and z direction with a period of 5 mm. The electromagnetic wave is incident in z direction with an electric field polarized in the y direction.

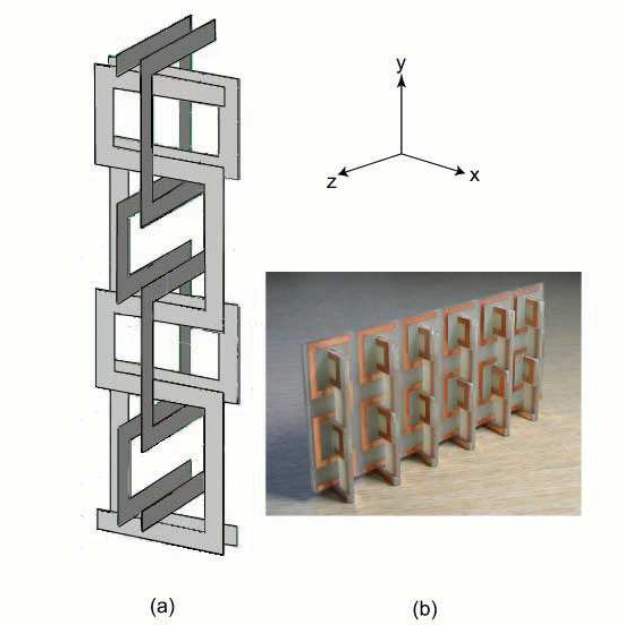


Figure 5: (a) The two dimensional metamaterial arranged in a cross embedded structure, (b) Picture of the metamaterial used in the experimental measurements which is arranged in a cross embedded structure.

The simulated S-parameters are shown in Fig. 6. Also the effective permittivity and permeability could be obtained using the retrieval method, and are shown in Fig. 7. Compared with the simulated results of the honeycomb structure, we find that the loss at the peak frequency of the passband is about 1.5 dBm, which is much lower, and the passband is much wider as well, with a relative bandwidth of 11%. The better performance of the cross embedded structure is due to the smaller interaction between the rings.

3. EXPERIMENT REALIZATION

In order to measure the transmission properties of the two dimensional metamaterials, we use a rectangular waveguide with a working frequency band from 1.7–2.5 GHz. Fig. 2(b) and Fig. 5(b) are the pictures of the samples used in the experiment. The two samples are loaded into the waveguide independently. The S parameters are recorded by an Agilent 8722ES network analyzer.

Using the s parameters measured in the experiment, the effective permittivity and permeability are retrieved, shown in Fig. 8 and Fig. 9. The relative bandwidth of honeycomb structure and

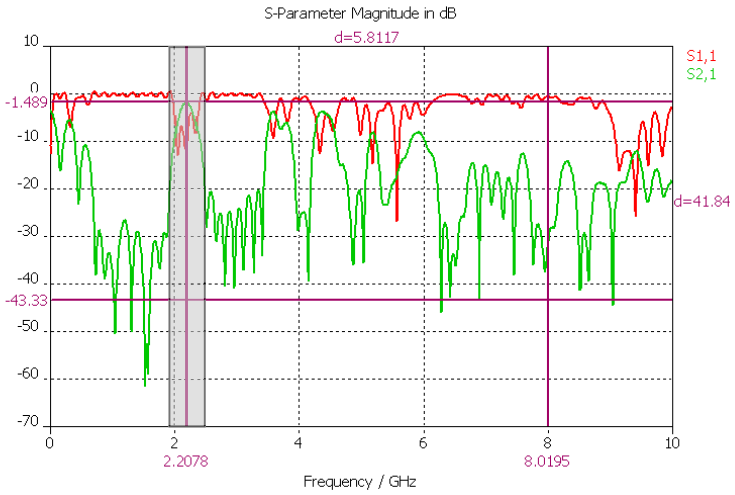


Figure 6: Simulated S-parameters of the two dimensional metamaterial arranged in a cross embedded structure.

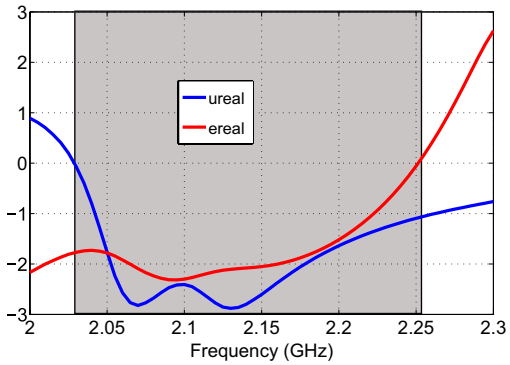


Figure 7: Effective permittivity and permeability of the 2D LHM of cross embedded structure by retrieval method using the simulated s parameters.

cross-embedded structure are 1.3% and 13% relatively. The measured bandwidth changes a little compared with the simulated result due to the gaps between the samples and the boundary of the waveguide, and a slight systematic misalignment leads to the shift in the peak frequency of the transmission band because the resonant frequency of the structure is very sensitive to small changes in parameters.

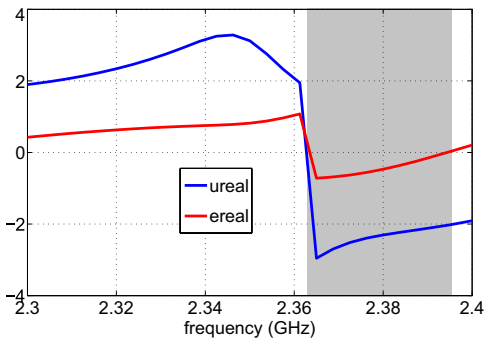


Figure 8: Effective permittivity and permeability of the 2D LHM of honeycomb structure by retrieval method using the measured s parameters.

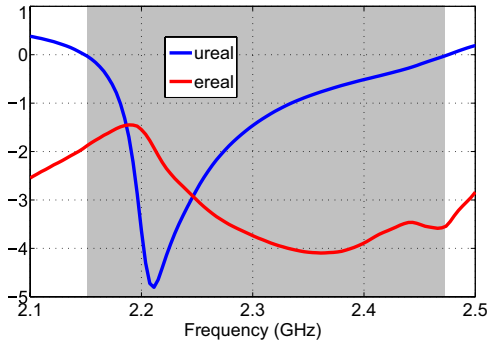


Figure 9: Effective permittivity and permeability of the 2D LHM of cross-embedded structure by retrieval method using the measured s parameters.

4. CONCLUSION

A new type of two dimensional metamaterial is presented in this paper, of which the rings are cross embedded arranged. This structure leads to a much better performance, which could be seen in

the simulation and experiment result.

ACKNOWLEDGMENT

This work is supported by NSFC 60531020, 60671003, and in part by ZJNSF R105253.

REFERENCES

1. Shelby, R. A., D. R. Smith, and S. Schultz, *Science*, Vol. 292, 77, 2001.
2. Shelby, R. A., D. R. Smith, S. C. Nemat-Nasser, and S. Schultz, *Applied Physics Lett.*, Vol. 78, 4, 2001.
3. Wang, X., Z. F. Ren, and K. Kempa, *Optics Express*, Vol. 12, 13, 2004.
4. Chen, H., J. Zhang, B. Yang, Y. Luo, L. Ran, Q. Jiang, and J. A. Kong, "Experimentally retrieve the effective parameters of metamaterial based on waveguide method," to be published.

Tunable Metamaterials Based on Nematic Liquid Crystals

Qian Zhao, **Lei Kang**, **Bo Du**, **Bo Li**, **and Ji Zhou**
State Key Lab of New Ceramics and Fine Processing, Department of Materials Science and Engineering
Tsinghua University, Beijing 100084, China

Abstract— Electrically tunable negative permeability metamaterial consisting of periodic array of split ring resonators (SRRs) infiltrated with nematic liquid crystals is demonstrated. Transmission measurement shows that the transmitted resonance dip of the metamaterial shifts towards low frequency as the applied electric field increases. The maximum shift is about 210 MHz and the shift can be continuously and reversible adjusted. Numerical simulation shows that the permeability of the metamaterial is negative near the resonance frequency and the frequency range with negative permeability can be dynamically adjusted and widen about 200 MHz by the applied electric field. An applied electric field changes the alignment of the liquid crystal directors and hence the dielectric constant of the structure, which permits dynamically tuning the negative permeability and provides a convenient means to add adaptive features to the metamaterials.

Recently the left-handed metamaterials (LHMs) with periodically arrayed split ring resonators (SRRs) creating negative permeability and wires creating negative permittivity have attracted much attention for the peculiar electromagnetic properties, such as negative refraction [1, 2], reversed Doppler shift [3], and perfect lens [4], etc. An array of SRRs can exhibit a strong resonant response to the magnetic component of an electromagnetic field, and may have a negative permeability near the resonance frequency [5]. Usually, the frequency region with negative permeability is very narrow and the frequency region with negative permittivity is relatively wider, thus the key issue of fabricating LHMs is the realization of the negative permeability. The resonant frequency of the SRRs is determined by the capacitivity and inductivity of the structure. Therefore, one can design the physical dimensions of the SRRs or the permittivity of the constituent materials to get some capacitivity and inductivity, and thus implement the magnetic resonance at such frequency. However, the above methods are static and passive, which limits the application of the metamaterials. It is desirable to realize the dynamically tunable negative permeability metamaterials by controlling the permittivity of the constituent materials.

Fortunately, the tunable electromagnetic behaviors in metamaterials have also been studied. Hou [6] et al. studied the tuning of photonic bandgaps by a field-induced structural change of fractal metamaterials. Padilla [7] et al. demonstrated the dynamical electric and magnetic metamaterial response by controlling the photo-excitation of free carriers in the substrate. Zhao [8] et al. used defect to tune the resonance of the SRRs and thus the left-handed properties. Chen [9] et al. realized a controllable left-handed metamaterial based on a varactor-loaded S-shaped resonator structure. Moreover, the tunable optics devices, such as photonic crystals have been realized by varying several external parameters, including electric field [10–12] and temperature [13], etc.

Liquid crystal has a large optical anisotropy due to its anisotropic molecular shape and alignment that is extremely sensitive to external parameters, and has been widely used to tuning the bandgaps of photonic crystal at optical frequency [14–17] and phase shifter at millimeter wave frequency [18, 19]. Therefore, liquid crystals can be used to dynamically control the dielectric properties of the metamaterials, and thus reversed electromagnetic behaviors in metamaterials such as negative permeability and negative refraction, etc. can be modulated by external field. Very recently, Khoo et al. [18] reported a theoretical analysis on the nanosphere dispersed LC to realize tunable metamaterial. In this letter, we demonstrated the electrically tuning of the metamaterials at centimeter microwave by changing the permittivity of its infiltration via the orientation of liquid crystal. The electric tuning method would also provide a convenient means to add adaptive features to the metamaterials.

Using shadow mask/etching technique, the copper SRRs are etched on a 0.25 mm-thick substrate with the permittivity of 3.3 and are arrayed to metamaterial with the lattice constant of 5.00 mm. Then the SRRs array is placed in a teflon cell filled with a nematic LC($n_o = 1.5$, $n_e = 1.78$ at 589 nm) to form the sample (Fig. 1). The dimensions of SRRs are $d_1 = 0.80$ mm, $d_2 = 1.60$ mm, $c = 0.20$ mm, $g = 0.30$ mm, and the thickness $t = 0.03$ mm. The sample is placed between two rectangular waveguide ports, and the scattering parameters are measured by an HP8720ES network analyzer. A TE_{10} mode is excited in the waveguide and the microwave propagates along the z axis

with the electric field E parallel to the x axis and magnetic field H parallel to the y axis. Two ITO electrodes connected to a dc power supply are separated by 30 mm and the dc electric field is perpendicular to the SRRs, i.e., along the y axis.

The transmissions at the normal incidence of the microwave are measured for three cases: the SRRs arrays without the liquid crystal, with the liquid crystal but no voltage supply, and with both the liquid crystal and the different voltage supply, and shown in Fig. 2. Data were taken for 50 dc electric fields from 0 to 0.35 V/μm. For clarity, only six curves are shown in Fig. 2, corresponding to the electric field: 0, 0.035, 0.070, 0.105, 0.175, 0.280 V/μm, respectively. It is observed from Fig. 2 that there is a resonance dip in the transmission curve for each case, and the transmission dip shifts about 2.14 GHz toward low frequencies with respect to the original dip at 13.22 GHz (black curve; here, the cell was empty with zero voltage) after infiltration of the liquid crystal at 11.08 GHz (red curve). The shift of the resonance dip towards low frequencies originates from the injection of the liquid crystal. It has been found that resonance dip is caused by intrinsic resonance of the SRRs and that the attachment of a dielectric slab would lower the resonance frequency due to increasing the capacitance in the SRRs [19]. Thus, the resonance dip shifts down after the empty cell is filled with the liquid crystal whose dielectric constant is larger than one. The effective dielectric constant of the infiltrated liquid crystals at microwave frequency can be extracted by comparing the transmission experiments of the SRRs-array with the simulations using CST Microwave Studio. From the experimental results that the resonance dip shifts from 13.22 GHz with no liquid crystal to 11.08 with liquid crystal filled, the effective dielectric constant of liquid crystal at the microwave frequency range can be determined as 2.7. It is noted that the liquid crystal has no refractive index dispersion over this frequency range [20].

When the electric field is applied and varied from 0 to 0.35 V/μm, the transmission dip gradually redshifts from 11.08 GHz to 10.87 GHz. The shift starts at ~ 0.01 V/μm, the threshold electric field of the liquid crystal, signaling the existence of a Freedericksz transition. A maximum tuning of 210 MHz is observed at 0.28 V/μm and the tuning appears to be independent of the dc electric field polarity. With the most rapid movement of the resonance frequency is in the electric field range between 0.01 and 0.28 V/μm. At higher electric field, the shift saturates because the nematic liquid crystal director is almost completely aligned along the electric field throughout the met materials. The tuning range is dependent on the effective refractive index difference from the zero-

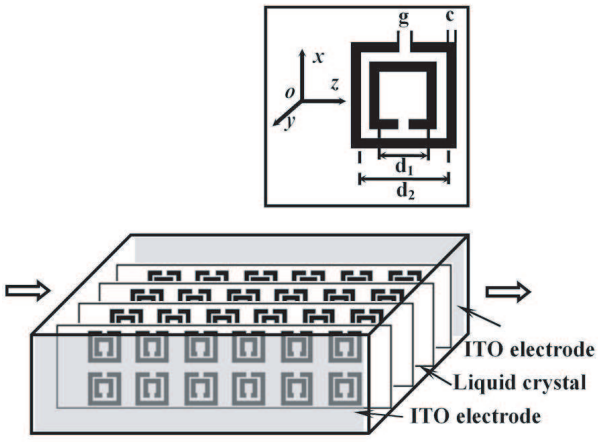


Figure 1: Schematic of the electrically tunable negative permeability metamaterial consisting of SRRs and infiltrated LC.

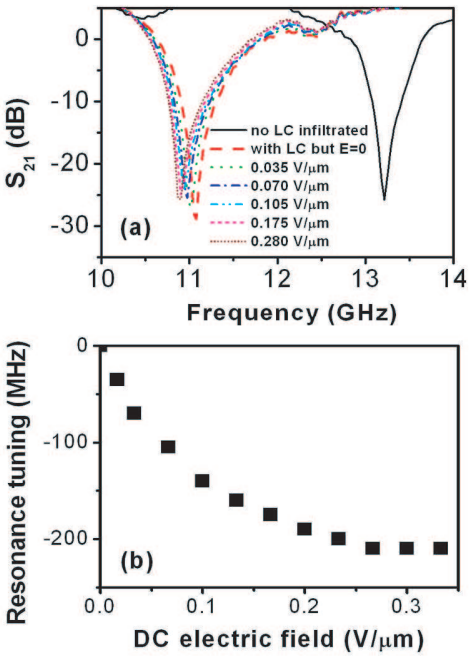


Figure 2: Demonstration of the electrically tunable negative permeability metamaterials. (a) Transmission spectra under different dc electric field. (b) Dependence of the tuning on the electric field with respect to that of zero electric field.

field state to the saturated state at the resonance frequency. Although overall tuning is limited, the demonstrated tuning range can possibly be further extended if a liquid crystal with higher birefringence is used. After removing the electric field, the resonance dip returns to its original frequency, i.e., the resonance shift is completely continuous and reversible.

The corresponding transmission phase versus frequency under different external electric field shows that the transmission phase has a transition at the resonance frequency, which indicates that the permeability may be negative. And the transmission phase can also be continuously and reversible adjusted by the electric field, which may be used as microwave phase shifter.

The effective permeability of the SRRs array metamaterials under no dc electric fields is calculated according to the expression in the Ref. [5].

$$\mu(\omega) = 1 - \frac{F\omega^2}{\omega^2 - \omega_0^2 + i\Gamma\omega}$$

where F is the fractional volume of the unit cell occupied by the interior of the SRRs, and Γ is the dissipation factor and can be determine by the geometry and materials composing the metamaterial. The parameters F and Γ of the used metamaterials should be determined as $F = 0.05$ and $\Gamma = 10^{10}/4\pi$. The calculated effective permeabilities of the metamaterials under different electric field are shown in Fig. 3. It can be seen that for the metamaterials with no liquid crystal infiltrated, the permeability is negative in the frequency range between 13.24 GHz and 13.54 GHz. When the liquid crystal is infiltrated, the negative permeability occurs in the frequency range between 11.08 GHz and 11.36 GHz, which shifts towards low frequency about 2.14 GHz due to the increase of the effective dielectric constant. When dc electric field is applied, the frequency range with negative permeability is at $10.98 \sim 11.26$ GHz and $10.86 \sim 11.14$ GHz for the dc electric field 0.067 and 0.28 V/μm, which also shifts to low frequency with respect to that of zero electric field. Under a proper electric field, the permeability can be negative in the frequency range of 0.5 GHz between 10.86 and 11.36 GHz, which is widen about 0.2 GHz with respect to that of the metamaterial without liquid crystals infiltrated. Moreover, such negative permeability can be dynamically adjusted by the external electric field.

To further verify that the electrically tunable behavior originates from the changes of liquid crystal molecular director, silicone oil is also used to infiltrate the metamaterials. The measured transmission (not shown) shows that the resonance frequency is independent on the electric field, which indicates that the electrically tunable behavior is due to the changes of LC molecular director.

It is known that the resonance frequency of SRRs is determined by the capacitance C and inductance L of the structure, i.e., $\omega_0 = 1/\sqrt{LC}$. The effective permittivity of the LC infiltrated metamaterial is controlled by an applied electric field, and thus the capacitance and the negative permeability are accordingly adjusted. The influence of anisotropic permittivity on capacitance can be analyzed using the dynamic electric field around the SRRs (Fig. 4). It can be seen that the electric field in the plane xoz (parallel to the SRRs) is along the radial of SRRs and mainly located in the gap of the inner and outer rings. Due to the stronger edge coupled effect [21] of the thinner SRRs (~ 0.03 mm), the electric field also has a y component. Therefore, the capacitance

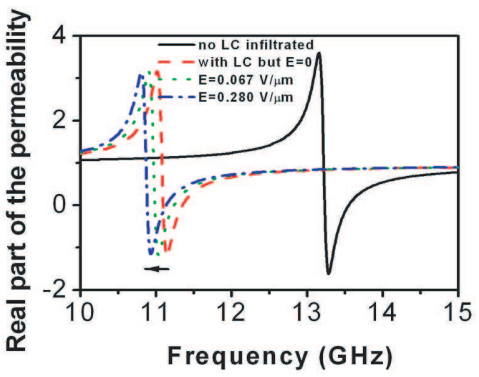


Figure 3: The calculated effective permeabilities of the metamaterials under different electric field.

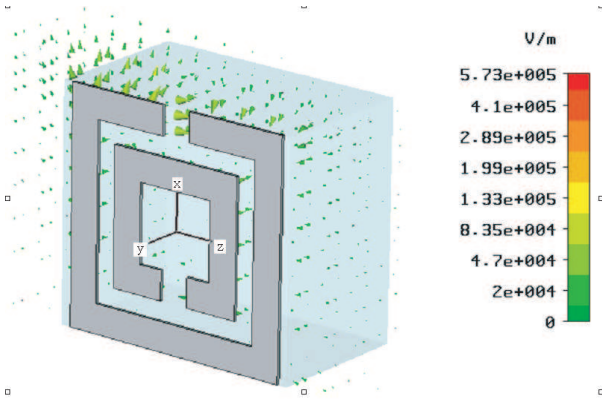


Figure 4: Simulated dynamic electric field intensity distribution of the SRRs at its resonance frequency.

of the SRRs is affected by the permittivity along the x, y, and z axis. Under no dc electric field, the LC directors are randomly oriented and the isotropic permittivity is $\varepsilon_{iso} = \frac{2}{3}\varepsilon_o + \frac{1}{3}\varepsilon_o$, where ε_o and ε_e are the ordinary and extraordinary permittivity, and ε_o is smaller than ε_e for the used LC. Application of an electric field will preferentially orient LC molecule along the field direction (y axis), which decreases ε_x and ε_z from ε_{iso} to ε_o, and increases ε_y from ε_{iso} to ε_e. And the increase of ε_y is larger than the decrease of ε_x and ε_z, the effective permittivity around the SRRs therefore increases and results in the redshift of the resonance of SRRs. It should be also noted that if SRRs becomes thicker, the edge coupled effect becomes weaker and the capacitance will mainly depend on ε_x and ε_z. Thus the application of electric field along the y axis will decrease the effective permittivity and the capacitance, which gives rise to the blueshift of magnetic resonance. The more detailed and quantitative explanation of the electrically tunable metamaterial is in study.

In summary, we demonstrated an electrically tunable negative permeability metamaterial at microwave frequencies consisting of periodic array of split ring resonators (SRRs) infiltrated with LCs. It shows that the resonance dip of SRRs can be continuously and reversible tuned by the electric field. Numerical simulation shows that the negative permeability range of the metamaterials can be adjusted by the electric field. Appling an electric field in the metamaterial changes the alignment of the liquid crystal directors and hence the dielectric constant of the structure, which permits dynamically tuning the negative permeability of metamatierlas. Because of the universality of metamaterials response over many range of frequency, the results is this paper are not limited only to microwave frequencies but may be scaled down to shorter wavelengths such as terahertz and infrared frequencies. The electrically tunable negative permeability metamaterial may be have more potential applications in the devices at terahertz or infrared frequencies, such as compact cavities, adaptive lenses, tunable mirrors, isolators, and converters.

ACKNOWLEDGMENT

This work is supported by the National Science Foundation of China under Grant Nos. 50425204, 90401012, and 60608016, the National Basic Research Program of China under subproject No. 2002CB61306, and by the Postdoctoral Science Foundation under Grant No. 20060390043.

REFERENCES

1. Houck, A. A., J. B. Brock, and I. L. Chuang, *Phys. Rev. Lett.*, Vol. 90, 137401, 2003.
2. Parazzoli, C. G., R. B. Greegor, K. Li, B. E. C. Koltenbah, and M. Tanielian, *Phys. Rev. Lett.*, Vol. 90, 107401, 2003.
3. Pendry, J. B., *Phys. Rev. Lett.*, Vol. 85, 3966, 2000.
4. Seddon, N. and T. Bearpark, *Science*, Vol. 302, 1537, 2003.
5. Pendry, J. B., A. J. Holden, D. J. Robbins, et al., *IEEE Transactions on Microwave Theory and Techniques*, Vol. 47, 2075, 1999.
6. Hou, B., G. Xu, H. K. Wong, and W. J. Wen, *Opt. Express*, Vol. 13, 9149, 2005.
7. Padilla, W. J. , A. J. Taylor, C. Highstrete, M. Lee, and R. D. Averitt, *Phys. Rev. Lett.*, Vol. 96, 107401, 2006.
8. Zhao, X. P., Q. Zhao, L. Kang, J. Song, and Q. H. Fu, *Phys. Lett. A*, Vol. 346, 87, 2005.
9. Chen, H., B. Wu, L. Ran, T. M. Grzegorczyk, and J. A. Kong, *Appl. Phys. Lett.*, Vol. 89, 053509, 2006.
10. Busch, K. and S. John, *Phys. Rev. Lett.*, Vol. 83, 967, 1999.
11. Shimoda, Y., M. Ozaki, and K. Yoshino, *Appl. Phys. Lett.*, Vol. 79, 3627, 2001.
12. Li, B., J. Zhou, L. Li, X. J. Wang, X. H. Liu, and J. Zi, *Appl. Phys. Lett.*, Vol. 83, 4704, 2003.
13. Zhou, J., C. Q. Sun, K. Pita, Y. L. Lam, Y. Zhou, S. L. Ng, C. H. Kam, L. T. Li, and Z. L. Gui, *Appl. Phys. Lett.*, Vol. 78, 661, 2001.
14. McPhail, D., M. Straub, and M. Gu, *Appl. Phys. Lett.*, Vol. 86, 051103, 2005.
15. Haurylau, M., S. P. Anderson, K. L. Marshall, and P. M. Fauchet, *Appl. Phys. Lett.*, Vol. 88, 061103, 2006.
16. Maune, B., M. LonËar, J. Witzens, M. Hochberg, T. B. Jones, D. Psaltis, and A. Scherer, *Appl. Phys. Lett.*, Vol. 85, 360, 2004.
17. Lim, K. C., J. D. Margerum, and A. M. Lackner, *Appl. Phys. Lett.*, Vol. 62, 1065, 1993.
18. Khoo, I. C., D. H. Werner, X. Liang, A. Diaz, and B. Weiner, *Opt. Lett.*, Vol. 31, 2592, 2006.
19. Markos, P. and C. M. Soukoulis, *Phys. Rev. E*, Vol. 65, 036622, 2002.
20. Yang, F. Z. and J. R. Sambles, *Appl. Phys. Lett.*, Vol. 81, 2047, 2002.
21. Marques, R., F. Medina, and R. Rafii-El-Idrissi, *Phys. Rev. B*, Vol. 65, 144440, 2002.

Bistatic Electromagnetic Scattering from Randomly Rough Surfaces

Y. Du[1], **H. Y. Chen**[1], **Y. Qi**[1], **and J. A. Kong**[2]

[1]Zhejiang University, Hangzhou 310058, China
[2]Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Abstract— In this paper we propose a bistatic model for electromagnetic scattering from a Gaussian rough surface with small to moderate heights. It is based on the integral equation formulation where the spectral representations of the Green's function and its gradient are in complete forms, a general approach similar to those used in the advanced integral equation model (AIEM) and the integral equation model for second-order multiple scattering (IEM2M). Yet this new model can be regarded as an extension to these two models on two accounts: first it has made fewer and less restrictive assumptions in evaluating the complementary scattering coefficient for single scattering, and second it contains a more rigorous analysis by the inclusion of the error function related terms for the cross and complementary scattering coefficients, which stems from the absolute phase term in the spectral representation of the Green's function. It is expected that our result for complementary scattering coefficient is more accurate and more general, even when the effect of the error function related terms is neglected. As a result, the proposed model is expected to have wider applicability with a better accuracy. The validity of this extended bistatic scattering model is demonstrated through the excellent agreements between model predictions and measurement data.

The original IEM model [1] which used simplified surface current estimate has shown to provide good predictions for forward and backward scattering coefficients. Concerns over the assumptions have prompted several modifications and variations of IEM in the literature. Regarding the spectral representation of the Green's function, the simplification was discarded and full form was restored, resulting in a modification to the complementary components. The resulting model is the so called improved IEM model (I-IEM) [2]. Additional restoration of the spectral representation of the gradient of the Green's function in its full form leads to the advanced IEM model (AIEM) [3] and the IEM2M model [4].

However, there are some technical subtleties in connection with the restoration of the full Green's function that have not been adequately reflected in these models. For example, in evaluating the average scattered complementary field over height deviation z, a split of the domain of integration into two semi-infinite ones is required due to the absolute phase term present in the spectral representation of the Green's function. This operation will lead to an expression containing the error function. Inclusion of the error function related terms is also encountered when one evaluates the incoherent powers that involve the scattered complementary field. Thus, a complete expression for the cross scattering coefficient or for the complementary scattering coefficient should have two parts: one does not contain the error function and the other includes its effect. The latter can be regarded as a correction term and an analysis of its effect is desirable. Roughly speaking, for the case where both the media above and below the rough surface are lossless, it can be shown that the correction term vanishes for the cross scattering coefficient, but not for the complementary scattering coefficient; for the case where either medium is of lossy nature, the correction term due to this lossy medium will contribute to both the cross and complementary scattering coefficients. Inclusion of the correction terms associated with the error function and an according discussion of their effect, to the best knowledge of the authors, are still missing in the literature. Supplement of this information forms one of the two contributions of the present work. The other contribution is a new treatment of the complementary incoherent power, where fewer and less restrictive assumptions are made than previous work.

In the analysis to follow immediately, we shall highlight the key development of our study, in particular on how the error function is introduced and how it is used in evaluating the bistatic scattering coefficient.

In calculating the Kirchhoff-complementary incoherent power, we need to evaluate the following quantity

$$I_F =< \exp\left\{-i\left[k_{sz}z + k_z z' - q\left|z - z'\right|\right]\right\} > \tag{1}$$

where k_{sz} and k_z are the z-components of the scattered and incident wave vectors, respectively. q is the z-related term in the spectral representation of the Green's function. Now we introduce the

transformation of variables, $z = \frac{y_1+y_2}{\sqrt{2}}$ and $z' = \frac{y_1-y_2}{\sqrt{2}}$ to obtain the decoupling

$$I_F = I_{F1} I_{F2} \tag{2}$$

where I_{F1} involves integration over the variable y_1 while I_{F2} involves integration over the variable y_2 which is expressed as

$$I_{F2} = \frac{1}{\sqrt{2\pi}\sigma\sqrt{1-\rho}} \int_{-\infty}^{+\infty} dy_2 \left\{ \exp\left\{ -i \left[\frac{k_{sz} y_2}{\sqrt{2}} - \frac{k_z y_2}{\sqrt{2}} - \sqrt{2} q \left| y_2 \right| \right] \right\} \exp\left[-\frac{y_2^2}{2\sigma^2(1-\rho)} \right] \right\} \tag{3}$$

where σ is the rms height, ρ is the surface correlation function. The quantity I_{F2} can be decomposed as

$$I_{F2} = I_{F2}^{+} + I_{F2}^{-} \tag{4}$$

where

$$I_{F2}^{+} = \frac{1}{\sqrt{2\pi}\sigma\sqrt{1-\rho}} \int_{0}^{+\infty} dy_2 \exp\left[-\frac{y_2^2}{2\sigma^2(1-\rho)} - i \left(\frac{k_{sz} - k_z}{\sqrt{2}} - \sqrt{2} q \right) y_2 \right] \tag{5}$$

and

$$I_{F2}^{-} = \frac{1}{\sqrt{2\pi}\sigma\sqrt{1-\rho}} \int_{-\infty}^{0} dy_2 \exp\left[-\frac{y_2^2}{2\sigma^2(1-\rho)} - i \left(\frac{k_{sz} - k_z}{\sqrt{2}} + \sqrt{2} q \right) y_2 \right] \tag{6}$$

Carrying out the operation yields

$$I_{F2}^{+} = \frac{1}{2} \exp\left[-\frac{1}{4} \left(k_{sz} - k_z - 2q \right)^2 (1-\rho)\sigma^2 \right] \left(1 + erf\left[\frac{\sigma\sqrt{1-\rho}}{2} \left(k_{sz} - k_z - 2q \right) \right] \right) \tag{7}$$

and

$$I_{F2}^{-} = \frac{1}{2} \exp\left[-\frac{1}{4} \left(k_{sz} - k_z + 2q \right)^2 (1-\rho)\sigma^2 \right] \left(1 + erf\left[-i\frac{\sigma\sqrt{1-\rho}}{2} \left(k_{sz} - k_z + 2q \right) \right] \right) \tag{8}$$

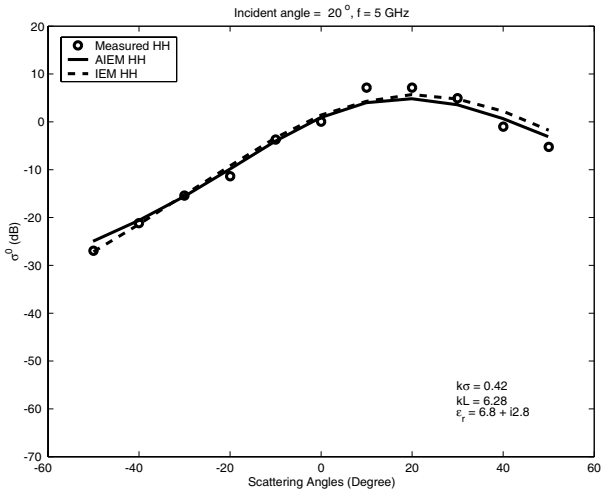


Figure 1: Comparisons between bistatic scattering models EAIEM and IEM against measurement data for Gaussian surface.

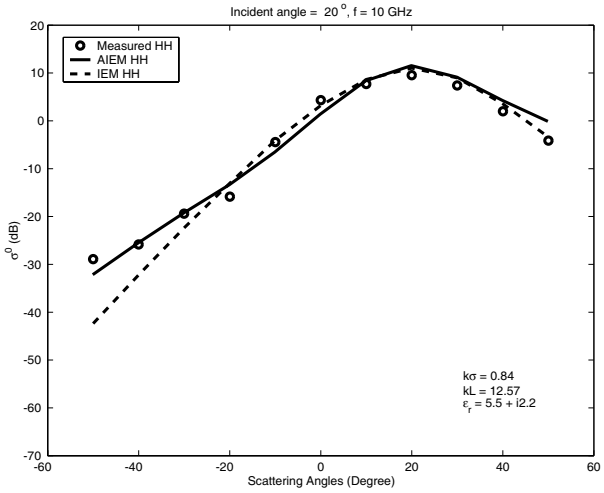


Figure 2: Comparisons between bistatic scattering models EAIEM and IEM against measurement data for Gaussian surface.

When the error function is ignored, our result for the scattering coefficient of the Kirchhoff-

complementary cross term is

$$\begin{aligned}\sigma_{qp}^{0kc} &= \frac{k^2}{4}\exp\left\{-\sigma^2\left(k_{sz}^2+k_{sz}k_z+k_z^2\right)\right\}\\ &\times\sum_{m=1}^{2}\sum_{d=-1,1}\mathrm{Re}\Big\{f_{qp}^{*}\Big[F_{qp}^{(m,d)}\left(-\mathbf{k_{s\perp}}\right)\exp\left\{-\sigma^2\left(-k_{sz}q_{m,s}d+k_zq_{m,s}d+q_{m,s}^2\right)\right\}\\ &\times\sum_{n=1}^{\infty}\frac{1}{n!}\left[\sigma^2\left(k_z+q_{m,s}d\right)\left(k_{sz}+k_z\right)\right]^n W^{(n)}\left(\mathbf{k_{s\perp}}-\mathbf{k_{i\perp}}\right)\\ &+F_{qp}^{(m,d)}\left(-\mathbf{k}_{i\perp}\right)\exp\left\{-\sigma^2\left(-k_{sz}q_{m,i}d+k_zq_{m,i}d+q_{m,i}^2\right)\right\}\\ &\times\sum_{n=1}^{\infty}\frac{1}{n!}\left[\sigma^2\left(k_{sz}-q_{m,i}d\right)\left(k_{sz}+k_z\right)\right]^n W^{(n)}\left(\mathbf{k_{s\perp}}-\mathbf{k_{i\perp}}\right)\Big]\Big\} \qquad (9)\end{aligned}$$

which includes both the upward wave ($\ell = 1$) and downward ($\ell = -1$) in both the media above the surface ($m = 1$) and below the surface ($m = 2$). In comparing our results with those appear in the literature, we find that it is identical to the result obtained in [3, 4]. When the error function related terms are to be retained, we can regard them as correction terms, which we express in (10).

$$\begin{aligned}\Delta\sigma_{qp}^{0kc} &= \frac{k^2}{4}\exp\left\{-\sigma^2\left(k_{sz}^2+k_{sz}k_z+k_z^2\right)\right\}\\ &\times\sum_{m=1}^{2}\sum_{d=-1,1}\mathrm{Re}\Big\{f_{qp}^{*}\Big[\exp\big\{-\sigma^2\big(-k_{sz}q_{m,s}d+k_zq_{m,s}d+q_{m,s}^2\big)\big\}\,erf\Big[i\frac{\sigma}{2}\left(k_{sz}-k_z-2q_{m,s}d\right)\Big]\\ &\times F_{qp}^{(m,d)}\left(-\mathbf{k_{s\perp}}\right)\sum_{n=1}^{\infty}\frac{1}{n!}\left[\sigma^2\left(k_z+q_{m,s}d\right)\left(k_{sz}+k_z\right)\right]^n W^{(n)}\left(\mathbf{k_{s\perp}}-\mathbf{k_{i\perp}}\right)\\ &+\exp\left\{-\sigma^2\left(-k_{sz}q_{m,i}d+k_zq_{m,i}d+q_{m,i}^2\right)\right\}erf\left[i\frac{\sigma}{2}\left(k_{sz}-k_z-2q_{m,i}d\right)\right]\\ &\times F_{qp}^{(m,d)}\left(-\mathbf{k}_{i\perp}\right)\sum_{n=1}^{\infty}\frac{1}{n!}\left[\sigma^2\left(k_{sz}-q_{m,i}d\right)\left(k_{sz}+k_z\right)\right]^n W^{(n)}\left(\mathbf{k_{s\perp}}-\mathbf{k_{i\perp}}\right)\\ &+i\exp\left\{-\sigma^2\left(-k_{sz}q_{m,s}d+k_zq_{m,s}d+q_{m,s}^2\right)\right\}\frac{\sigma\left(k_{sz}+k_z\right)}{\sqrt{\pi}}\\ &\times F_{qp}^{(m,d)}\left(-\mathbf{k_{s\perp}}\right)\exp\left\{\left[\frac{\sigma}{2}\left(k_{sz}-k_z-2q_{m,s}d\right)\right]^2\right\}\\ &\times\sum_{n=1}^{\infty}\frac{1}{n!}\left[\sigma^2\left(k_z+q_{m,s}d\right)\left(k_{sz}+k_z\right)\right]^n W^{(n+1)}\left(\mathbf{k_{s\perp}}-\mathbf{k_{i\perp}}\right)\\ &-i\exp\left\{-\sigma^2\left(-k_{sz}q_{m,i}d+k_zq_{m,i}d+q_{m,i}^2\right)\right\}\frac{\sigma\left(k_{sz}+k_z\right)}{\sqrt{\pi}}\\ &\times F_{qp}^{(m,d)}\left(-\mathbf{k}_{i\perp}\right)\exp\left\{\left[\frac{\sigma}{2}\left(k_{sz}-k_z-2q_{m,i}d\right)\right]^2\right\}\\ &\times\sum_{n=1}^{\infty}\frac{1}{n!}\left[\sigma^2\left(k_{sz}-q_{m,i}d\right)\left(k_{sz}+k_z\right)\right]^n W^{(n+1)}\left(\mathbf{k_{s\perp}}-\mathbf{k_{i\perp}}\right)\\ &+i\exp\left\{-\sigma^2\left(-k_{sz}q_{m,s}d+k_zq_{m,s}d+q_{m,s}^2\right)\right\}\frac{\sigma\left(k_{sz}+k_z\right)}{\sqrt{\pi}}\\ &\times F_{qp}^{(m,d)}\left(-\mathbf{k_{s\perp}}\right)\exp\left\{\left[\frac{\sigma}{2}\left(k_{sz}-k_z-2q_{m,s}d\right)\right]^2\right\}W\left(\mathbf{k_{s\perp}}-\mathbf{k_{i\perp}}\right)\\ &-i\exp\left\{-\sigma^2\left(-k_{sz}q_{m,i}d+k_zq_{m,i}d+q_{m,i}^2\right)\right\}\frac{\sigma\left(k_{sz}+k_z\right)}{\sqrt{\pi}}\end{aligned}$$

$$\times F_{qp}^{(m,d)}\left(-\mathbf{k}_{i\perp}\right)\exp\left\{\left[\frac{\sigma}{2}\left(k_{sz}-k_z-2q_{m,i}d\right)\right]^2\right\}W\left(\mathbf{k_{s\perp}}-\mathbf{k_{i\perp}}\right)\Bigg]\Bigg\} \tag{10}$$

After much algebra, we obtain the scattering coecient for the complementary term. The results, in both error function exclusive and error function inclusive forms, are quite complicated and lengthy and will not be reported here. The details will be provided in a follow-up paper.

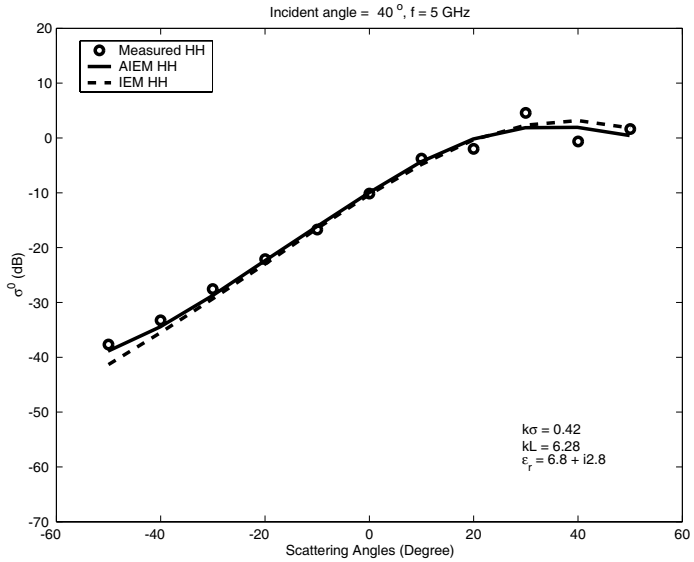


Figure 3: Comparisons between bistatic scattering models EAIEM and IEM against measurement data for Gaussian surface.

To validate the proposed EAIEM model, we compare its simulation results with measurement data for dielectric rough surfaces. The measurement data are taken from [6] where rough surfaces with Gaussian height and Gaussian power spectrum are concerned. The surface roughness parameters are correlation length $L = 6\,\text{cm}$ and rms height $\sigma = 0.4\,\text{cm}$. Figs. 1–3 present the model predictions of EAIEM and IEM against the measurement data. It is seen that for all these three cases, the model predictions by EAIEM agree very well with the measurement data, and outperform that of IEM for certain scattering angles.

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. Fung, A. K., Z. Li, and K. S. Chen, "Backscattering from a randomly rough dielectric surface," *IEEE Trans. Geosci. Remote Sensing*, Vol. 30, 356–369, Mar. 1992.
2. Hsieh, C. Y. and A. K. Fung, "Application of an extended IEM to multiple surface scattering and backscatter enhancement," *J. Electromagnetic Waves and Applications*, Vol. 13, 121–135, 1999.
3. Chen, K. S., et al., "Emission of rough surfaces calculated by the integral equation method with comparison to three-dimensional moment method simulations," *IEEE Trans. Geosci. Remote Sensing*, Vol. 41, 90–101, Jan. 2003.
4. Alvarez-Perez, J., "An extension of the IEM/IEMM surface scattering model," *Waves in Random Media*, Vol. 11, 307–329, 2001.
5. Licheri, M., et al., "On the scattering from natural surfaces: the IEM and the improved IEM," *Proceeding of IGARSS '01*, July 2001.
6. Macelloni, G., et al., "Experimental validation of surface scattering and emission models," *IEEE Trans. Geosci. Remote Sensing*, Vol. 38, 459–469, 2000.

Monitoring of Satellite Thermal Plareau in Relation to Concentration of Infrared Beams out of Sea Surface Waves

S. Nakamura
Kyoto University, Japan

Abstract— In order to have a physical understanding of a satellite thermal plateau found on the sea surface at a satellite monitoring, a model of an infrared beam concentration at a satellite out of the water wave facet on the sea surface.

In order to have a physical understand a satellite thermal plateau found on the sea surface at a satellite monitoring, a model of an infrared beam concentration at the satellite out of water surface wave facet is introduced in this work.

A problem is on a satellite detected some beam concentrating infrared beams out of a sea surface water wave facet. Assuming that the sea surface could be taken as a black body in a scope of thermodynamics to radiate a flux of some infrared beams, a sensor of the satellite might detect an anomalous signal. The interested beam could form a caustic (or a focus) at the sensor as an effect of a specific concave wave facet on the sea water surface when each beam is radiated normal to the sea water surface. Thermodynamics tells us that a radiation follows the Stefan-Boltsmann's law, so that, the relation between an integrated intensity B of the radiation and the water surface temperature T (in Kelvin unit) is well formulated to show the following relation for a mall variations of $\mathrm{d}B$ and $\mathrm{d}T$, as follows:

$$\mathrm{d}B/B = 4\mathrm{d}T/T \tag{1}$$

This relation is independent of any frequency of the interested infrared beam radiated out of the sea surface as a black body.

When beams out of the sea surface water wave facet form a focusing point at the sensor mounted on a satellite, the value of $\mathrm{d}B$ should be infinity so that the value of the $\mathrm{d}T$ is also infinity. In this case, the distance between the sensor mounted on the satellite and the wave facet is same to the altitude of the satellite above the sea surface.

When the beams form a caustic to hit the interested sensor of the satellite, and the value of $\mathrm{d}T$ is observed on a sea surface thermal pattern after a satellite monitoring, a degree of the caustic for the concentration of the infrared beams can be evaluated.

When $\mathrm{d}T = 5\,\mathrm{K}$ and $T = 273\,\mathrm{K}$, then, we have 0.018 as $\mathrm{d}T/T$. This evaluation may be a helpful result that the value 0.018 by part could be caused the existence of the wave facets which are always found on the sea water surface. Any one of the facets should be concave upward. Then, this may give us an understanding of a discrepancy on the sea surface thermal pattern with a temperature parameter after the satellite monitoring of a local sea surface thermal pattern. This may a suggesting for adjusting the thermal pattern obtained after the satellite monitoring system since the starting date of satellite monitoring. The author unfortunately such a thermodynamical understanding has never been payed for attention at the correction of the sea surface thermal pattern obtained by the satellite monitoring.

When $\mathrm{d}T = 30\,\mathrm{K}$, the evaluated value of $\mathrm{d}T/T$ is approximately 0.0989 which means the radius of the water surface curvature is near to the altitude of the satellite above the sea surface. This case is occasionally found under some conditions of the conditions related to meteorology on the interested sea surface.

For example, this case appears in the coastal zone in the northwestern Pacific. The thermal pattern of the sea surface looks like to be a thermal plateau when a distant storm in the subtropical zone is seen and the storm generated disturbances in form of the sea surface water waves are propagating out of the storm area to hit the coastal zone. This thermal plateau is found just under a foot print of a satellite with a clear sky condition. An uniformly atmospheric condition could be similar to find the storm induced sea surface waves on the apparently high temperature field taken to be as a thermal plateau. In the northwestern Pacific, the existence of the typhoon is effective.

The other case is for the cold season in the specific monsoon area. In winter of the northern hemisphere, a developing cold front of the atmosphere near the ground surface could be the most effective at finding the thermal plateau in the coastal zone in the northwestern Pacific.

With the above noted result, we should not yet take it as a convenient way for a detecting the sea surface wave field. At present, there are several factors controlling this appearance, and this problem has to be studied for a practical application to a demonstration of where the thermal plateau is appeared and of how about extent is possible to see a wave field by the satellite monitoring.

Monitoring Satellite Thermal Pinnacle in Relation to Spacial Spectrum of Sea Surface Waves

S. Nakamura
Kyoto University, Japan

Abstract— In order to realize a satellite thermal pinnacle and a satellite thermal Plateau found on the sea surface, a simple model of an infrared radiation out of the water wave facet on the sea water surface is introduced.

In order to realize a satellite thermal pinnacle and a satellite thermal plateau found on the sea surface, a model of an infrared radiation out of water surface wave facet is introduced in this work.

First, a frame of a modeling in this work is introduced. When a flat sea surface is on the co-ordinate x horizontally, a position of a satellite above the sea surface can be expressed as $(x,\, z) = (0,\, H)$ in the orthogonal system of the co-ordinates with the axis z for the axis to be positive upward. In this work, for convenience, a two dimensional problem is considered. This frame can be taken as an approximation at considering a small sea area under a satellite in a polar orbital motion, then, it can be taken to be acceptable to consider one pixel size of 4 km square in a satellite thermal pattern to be a flat plane in the foot print comparing to the radius of the Earth. If the sea surface on the earth everywhere could be taken to be as a black body, then, it can be a case of thermodynamics which tells us a radiation out of the sea surface as a beam of an electromagnetic wave in the infrared band.

When a sensor mounted on the satellite is directed vertically downward, the sensor detects the beam radiated vertically out of the flat sea surface.

Assuming an arbitrary function of the sea surface,

$$F = F(x,\, z;\, t), \tag{1}$$

then,

$$H = H\cos\theta + F, \quad \text{for } x = x, \text{ at an arbitrary time } t. \tag{2}$$

This could be a most simplified model of a satellite thermal pinnacle.

As the tangent of F for the position $(X,\, Z) = (x,\, z)$ at time t is to be $F' = \partial F/\partial x$, the orthogonal of the tangent is described as

$$(Z - z) = -(1/F')(X - x). \tag{3}$$

This normal line is taken to be corresponding to the direction of the beam radiated at $(x,\, z)$. When this normal line hits the point $(X,\, Z) = (0,\, H)$, the sensor catches the beam as a signal from the sea surface. At any other cases of $Z \neq H$, the hitting beam can not be expected. When this pattern can be taken as a model for a beam radiated out of the sea surface, the beam can be found only at a single pixel in a thermal pattern which was reduced after directly receiving of an interested satellite's signals related to the sea surface thermal pattern.

The function F introduced above can be transformed mathematically to express it in a form of spectral function $S_s = S_s(\omega;\, t)$ at a fixed position x. That is,

$$S_s = \int_0^\infty F(x,\, z;\, t)\, \exp(i\omega t)\, \mathrm{d}\omega, \tag{4}$$

where, the notation ω is for frequency.

This spectral expression has been widely used for studying wave developing process at a fixed position at a time t. Though, this spectral form is not effective for a spacial pattern of the waves or of a wave facet on the sea surface.

Then, it should be introduced a spacial spectral function S, which is expressed as

$$S = S(k;\, t), \quad \text{at} \quad t = t,$$

and, introducing a notation k for wave number,

$$S = \int_0^\infty F(x,\, z;\, t)\, \exp(ikx)\, \mathrm{d}k, \quad \text{at} \;\; t = t, \tag{5}$$

Then,

$$F = \int_0^\infty S(k;\, t)\, \exp(-ikx)\, \mathrm{d}x, \quad \text{at} \;\; t = t. \tag{6}$$

Substituting (6) into (3), it is obtained that

$$(Z - z) = (X - x)\partial/\partial x \left[\int_0^\infty S(k;\, t)\, \exp(-ikx)\, \mathrm{d}x \right]^{-1}. \tag{7}$$

When this normal line to the tangent at $(x,\, z)$ hits the position $(x,\, z) = (0,\, H)$,

$$(Z - H) = (X)\partial/\partial x \left[\int_0^\infty S(k;\, t)\, \exp(ikx)\, \mathrm{d}x \right]^{-1}. \tag{8}$$

This shows a single beam hitting at the point $(x,\, z) = (0,\, H)$ for an arbitrary form of the spectral function $S(k;\, t)$.

When this is taken as a model for a infrared beam radiated to hit at a sensor for the sea surface thermal pattern, a single pixel in a sea surface thermal pattern could be as a thermal pinnacle of an impulse form. Nevertheless, this model is hard to be even a simplified model for help to see any one of the sea surface thermal pinnacles actually found in the satellite sea surface thermal pattern.

Following what tells us the geomery for a plane, the radius of curvature in this work is for the wave facet on the sea surface. That is, expressed as follows referring to the sensor's position $(x,\, z) = (0,\, H)$ which stated above. Then,

$$(1/R) = (\mathrm{d}\theta/\mathrm{d}s), \tag{9}$$

where, the notation R is the radius of curvature of the wave facet, and the notation $\mathrm{d}\theta$ is for the angle length of the wave facet at the sensor. The length segment $\mathrm{d}s$ is;

$$(\mathrm{d}s)^2 = (\mathrm{d}x)^2 + (\mathrm{d}z)^2, \tag{10}$$

where, $\mathrm{d}z$ and $\mathrm{d}x$ are for the wave facet at the position $(X,\, Z)$. The relation of X and Z is described as shown in a form of (8).

With what the author has studied above, The relation between X and Z is obtained in a simple form, and spacial spectral function at a time t is a function of wave number k in this case. Thence, the expression (8) should be rewritten for the problem on the energy flux or energy transfer of the beam out of the facet concentrates just near at the sensor where the beam is focusing as an electromagnetic wave with a consideration. The factor of interferometry is simply effective at decaying of the beam intensity, then,the author considers that this interferometry problem is outside of this work. The SAR system (Synthetic Aperture Rader) is for a faint thermal difference so that the author takes it also outside of his interest in this work.

Data-derived SEA for Time Domain EMI Sensing of UXO

K. Sun[1], **K O'Neill**[1,2], **B. E. Barrowes**[2], **F. Shubitidze**[1], **I. Shamatava**[1]
J. P. Fernández[1], **and K. D. Paulsen**[1]

[1]Thayer School of Engineering, Dartmouth College, Hanover NH, 03755, USA
[2]USA ERDC Cold Regions Research and Engineering Laboratory
72 Lyme Road, Hanover NH, 03755, USA

Abstract— Electromagnetic induction (EMI) is a prominent technique in Unexploded Ordnance (UXO) detection and discrimination research. Existing idealized forward models for the EMI response can be defeated by both the material and geometrical heterogeneity of realistic UXO. We have developed a new, physically complete modeling system referred to as the Standardized Excitations Approach (SEA). The SEA accounts for all the effects from these heterogeneities including their interactions within the object, and is applicable in both the near and far fields. According to the SEA, the excitation field is decomposed into fundamental modes, and the response of a given target to each fundamental mode (denoted as a fundamental solution) is obtained beforehand and saved in a library. In this way, the target response to an arbitrary excitation field can be calculated via a simple superposition of these fundamental solutions.

The model parameters (i.e., the fundamental solutions) of a given object are extracted from a sufficiently detailed set of measurement data. These parameters will be specific to each EMI instrument. The parameter extraction process was developed previously for the frequency domain using the GEM-3 EMI instrument. In this paper, we apply this SEA to time domain using the EM-63 instrument as an example. The receiver coil of the EM63 is a 0.5 m by 0.5 m square loop and can not be approximated by a point receiver. Therefore, in the model, the data is interpreted as the integration of the secondary field over the receiver loop. The objects we consider are all Body of Revolution (BOR) type objects. We exploit the fact that the calculated SEA model parameters also exhibit specific behavior because the target is a BOR Thus, the algorithm is improved by enforcing symmetric properties and zero total magnetic charge, which makes the algorithm more robust and more efficient. Preliminary results show that this approach works well for this time domain EMI instrument. After optimization, this model may be fast enough for implementation in inversion processing algorithms.

1. INTRODUCTION

The detection and discrimination of UXO based on EMI techniques can be classified as an inverse problem which requires many evaluations of a fast forward model. Existing forward models are either too slow for inversion or insufficient for representing the response of complicated objects like UXO. We developed a fast and accurate forward model, the standardized excitation approach (SEA) [1, 2], in which we represent the excitation field with superposition of a set of fundamental modes. The responses of the given target to each fundamental mode are termed as fundamental solutions, which are solved for beforehand and saved in a library. The target response to an arbitrary excitation will be simply a superposition of the fundamental solutions. The fundamental solutions can be obtained from a set of thorough lab measurement data (termed as solution definition data). We have reported on the SEA technique as applied to the frequency domain elsewhere [2]. In this paper, we further improve the technique and apply it in time domain, using EM63 sensor as an example.

2. DECOMPOSING EM63 SENSOR FIELD INTO SPHEROIDAL MODES

We choose spheroidal modes as the fundamental modes of primary field. The primary field of the EM63 sensor depends both on space and time. Since the frequency is low relative to the spatial scale we studied, the problem can be approximated as magnetoquasistatic, i.e., the field has the same phase within the entire space under consideration. Let the wave form of the transmitter be $T(t)$, then the primary field at any position $\mathbf{r}$ is $\psi^{pr}(\mathbf{r})\,T(t)$. The spatially dependent term $\psi^{pr}(\mathbf{r})$ can be decomposed into fundamental modes, so that

$$\psi^{pr}(\mathbf{r}) = \frac{H_0 d}{2} \sum_{m}^{\infty} \sum_{n=m}^{\infty} \sum_{p=0}^{1} b_{pmn} P_n^m(\eta) P_n^m(\xi) T_{pm}(\phi) \tag{1}$$

where (η, ξ, ϕ) are the standard spheroidal coordinates [3], d is the inter-focal distance, P_n^m are Associated Legendre functions, $T_{pm}(\phi)$ is $\cos(m\phi)$ for $p = 0$ and is $\sin(m\phi)$ for $p = 1$. Calculating magnetic potential of a sensor field is usually time consuming. We obtain the mode coefficients from magnetic field instead [1].

$$H_\xi(\mathbf{r}) = -\frac{1}{h_\xi}\frac{\partial\psi(\mathbf{r})}{\partial\xi} = -\frac{1}{h_\xi}\frac{H_0 d}{2}\sum_m^\infty\sum_{n=m}^\infty\sum_{p=0}^1 b_{pmn}P_n^m(\eta)\frac{dP_n^m(\xi)}{d\xi}T_{pm}(\phi) \quad h_\xi = \frac{d\sqrt{\xi^2-\eta^2}}{2\sqrt{\xi^2-1}} \tag{2}$$

In order to employ orthogonality, we define

$$f_\xi = -h_\xi H_\xi = \frac{H_0 d}{2}\sum_m^\infty\sum_{n=m}^\infty\sum_{p=0}^1 b_{pmn}P_n^m(\eta)\frac{dP_n^m(\xi)}{d\xi}T_{pm}(\phi)$$

so that

$$b_{pmn} = \int_{-1}^1 P_n^m(\eta)\int_0^{2\pi} T_{pm}(\phi)f_\xi d\phi d\eta \Bigg/ \left(\frac{\alpha\pi H_0 d}{2}\frac{dP_n^m(\xi)}{d\xi}\frac{(n+m)!}{(n-m)!}\frac{2}{2n+1}\right) \tag{3}$$

3. FORWARD MODEL

We represent the target response under each fundamental mode excitation $\frac{H_0 d}{2}P_n^m(\eta)P_n^m(\xi)T_{pm}(\phi)T(t)$ with a set of magnetic charges $\tilde{\rho}_{pmn}(\mathbf{r}',\mathrm{t})$. In other words, the scattered field at any given position $\mathbf{r}$ and at time t is $\mathbf{H}^s_{pmn}(\mathbf{r},\mathrm{t}) = \int\limits_{V_{\text{source}}} dV'[\tilde{\rho}_{pmn}(\mathbf{r}',\mathrm{t})\ \mathbf{G}(\mathbf{r},\mathbf{r}')]$, where V_{source} is the volume containing $\tilde{\rho}_{pmn}(\mathbf{r}',\mathrm{t})$, as shown in Fig. 1.

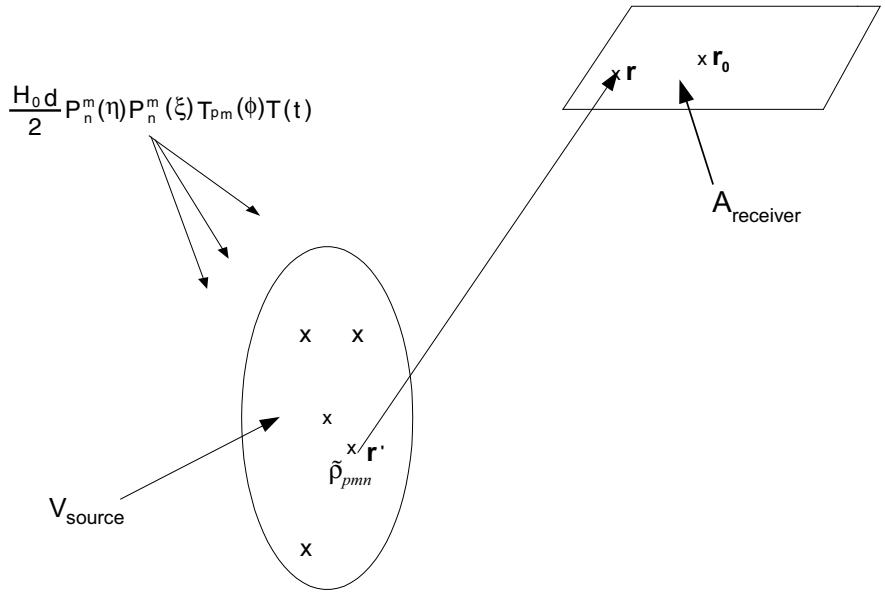


Figure 1: Sensor-target geometry. Target response of each fundamental mode can be represented by a set of magnetic charges $\widetilde{\rho}_{pmn}(\mathbf{r}', \mathbf{t})$.

The EM63 receivers are 0.5 m by 0.5 m square loops with area A_{receiver}. In the recording process, the system integrates the normal component of $\mathbf{H}^s$ over A_{receiver}, then differentiates it with respect to time, probably amplifies it, etc. All those latter processes are indicated by F_{Rx}. Suppose it can be separated into time dependent operation F_t and space dependent operation F_r, which are all

linear. If we denote the receiver center as $\mathbf{r}_0$, the received signal, $S(r,t)$, will be:

$$\begin{aligned} S(\mathbf{r_0},t) = F_{Rx}\left[\mathbf{H^s}\left(\mathbf{r},t\right)\right] &= F_rF_t\left[\mathbf{H^s}\left(\mathbf{r},t\right)\right] \\ &= \sum_{pmn} b_{pmn}(\mathbf{r_0}) \int\limits_{V_{\text{source}}} dV' F_rF_t[\tilde{\rho}_{pmn}(\mathbf{r'},\mathbf{t})\mathbf{G}(\mathbf{r},\mathbf{r'})] \\ &= \sum_{pmn} b_{pmn}(\mathbf{r_0}) \int\limits_{V_{\text{source}}} dV' F_t[\tilde{\rho}_{pmn}(\mathbf{r'},\mathbf{t})]F_r[\mathbf{G}(\mathbf{r},\mathbf{r'})] \end{aligned} \tag{4}$$

Next, define $\rho_{pmn}\left(\mathbf{r}',t\right) = F_t\left[\tilde{\rho}_{pmn}\left(\mathbf{r}',t\right)\right]$, then

$$S(\mathbf{r_0},\mathbf{t}) = \sum_{pmn} b_{pmn}(\mathbf{r_0}) \int\limits_{V_{\text{source}}} dV' \rho_{pmn}(\mathbf{r'},\mathbf{t})F_r[\mathbf{G}(\mathbf{r},\mathbf{r'})] \tag{5}$$

For EM63 system, the space dependent operation F_r integrates over the receiver loop area A_{receiver} and takes the normal component (i.e., n_{receiver} component only)

$$F_r\left[\mathbf{G}\left(\mathbf{r},\mathbf{r'}\right)\right] = \int\limits_{A_{\text{receiver}}(\mathbf{r}_0)} \mathbf{G}\left(\mathbf{r},\mathbf{r'}\right)\cdot\mathbf{n}_{\text{receiver}}\left(\mathbf{r_0}\right)dA \tag{6}$$

$$S(\mathbf{r_0},t) = \sum_{pmn} b_{pmn}(\mathbf{r_0}) \int\limits_{V_{\text{source}}} dV' \rho_{pmn}(\mathbf{r'},\mathbf{t}) \int\limits_{A_{\text{receiver}}(\mathbf{r_0})} \mathbf{G}(\mathbf{r},\mathbf{r'})\cdot\mathbf{n}_{\text{receiver}}(\mathbf{r_0})dA \tag{7}$$

For a Body of Revolution (BOR), the sources are several azimuthal rings of magnetic equivalent charges, with magnitude $\rho_{i,pmn}(t)T_{pm}$ for the ith ring. The ϕ dependency of the charge rings has been taken care of by the term $T_{pm}(\phi)$. The first term $\rho_{i,pmn}$ is the same for $p=0$ and $p=1$, i.e., $\rho_{i,0mn}(t) = \rho_{i,1mn}(t)$, which can be denoted as $\rho_{i,mn}(t)$.

$$S(\mathbf{r_0},t) = \sum_{pmn} b_{pmn}\left(\mathbf{r_0}\right) \sum_i \rho_{i,mn}(t) \int\limits_{A_{\text{receiver}(\mathbf{r_0})}} \mathbf{G}_{i,pm}(\mathbf{r})\cdot\mathbf{n}_{\text{receiver}}(\mathbf{r_0})dA \tag{8}$$

where

$$\mathbf{G}_{i,pm}(\mathbf{r}) = \int\limits_0^{2\pi} \frac{T_{pm}(\phi')}{4\pi|\mathbf{r}-\mathbf{r'}|^3}(\mathbf{r}-\mathbf{r'})r_i d\phi' \tag{9}$$

r_i is the radius of ith charge ring.

Define $INTG_{ipm}\left(\mathbf{r_0}\right) = \int\limits_{A_{\text{receiver}}(\mathbf{r_0})} \mathbf{n}_{\text{receiver}}\left(\mathbf{r_0}\right)\cdot\mathbf{G}_{i,pm}\left(\mathbf{r}\right)dA$

$$S(\mathbf{r_0},t) = \sum_{pmm}\sum_i b_{pmn}(\mathbf{r_0})\rho_{i,mn}(t)INTG_{ipm}(\mathbf{r_0}) \tag{10}$$

4. SOLVING MODEL PARAMETERS FROM SD DATA

At any given position $\mathbf{R_j}$, the modeled signal is

$$S^m(\mathbf{R_j},t) = \sum_{pmm}\sum_i b_{pmn}(\mathbf{R_j})\rho_{i,mn}(t)INTG_{ipm}(\mathbf{R_j}) \tag{11}$$

$$INTG_{ipm}\left(\mathbf{R_j}\right) = \int\limits_{A_{\text{receiver}}(\mathbf{R_j})} \mathbf{n}_{\text{receiver}}\left(\mathbf{R_j}\right)\cdot\mathbf{G}_{i,pm}\left(\mathbf{r}\right)dA \tag{12}$$

Obtain the sources $\rho_{i,mn}(t)$ by matching model data and measurements $S^m = S^{\mathbf{d}}$

$$\sum_{imn}(b_{0mn}(\mathbf{R_j})INTG_{i0m}(\mathbf{R_j}) + b_{1mn}(\mathbf{R_j})INTG_{ilm}(\mathbf{R_j}))\rho_{i,mn}(t) = S^d(\mathbf{R_j},t)$$

This is a set of linear equations but is usually ill-conditioned. The solutions are obtained using Tikhonov regularization [4].

5. EXAMPLE RESULTS

Consider data from UXO model 81 mm as an example of applying the SEA method in the time domain. The measurement were done with EM63, for a given objection location and orientation, the sensor travels along a square grid. The SD data includes 5 scenarios: 100 cm under the sensor with dip angle 0°, 45° and 270°. And 60 cm under the sensor with dip angles 0^0 and 270^0. The resolved fundamental solutions are then tested with data at distance 60 cm and dip angle 45°. Fig. 2 showed the comparison of modeled data and measurement data at the first time channel. The good matching of modeled data and measurement data suggested that the model is accurate.

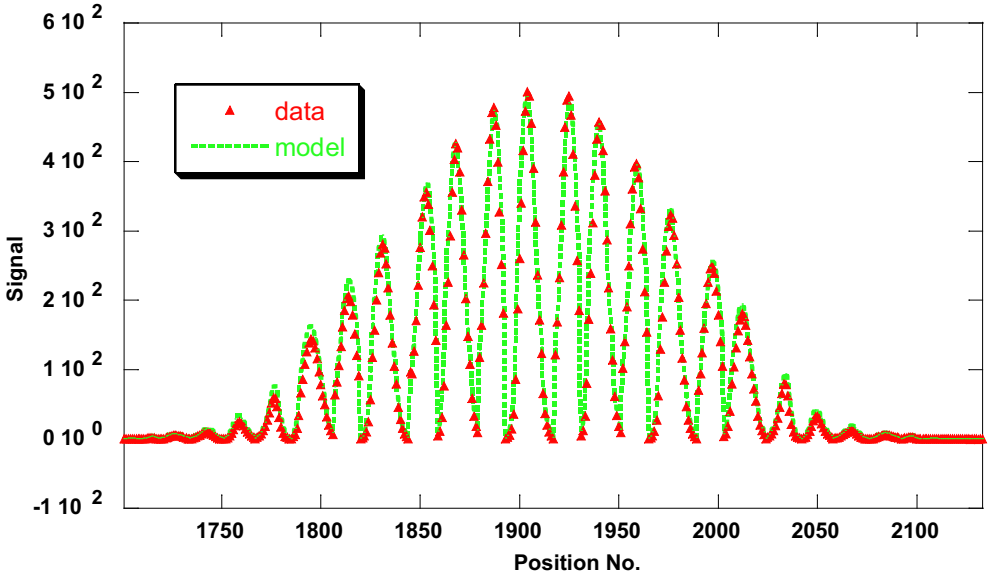


Figure 2: Comparison of modeled data and measurement data.

ACKNOWLEDGMENT

This work was sponsored in part by the Strategic Environmental Research and Development Program and US Army CoE ERDC BT25 and AF25 programs. The authors would like to thank Sky research for providing measurement data.

REFERENCES

1. Shubitidze, F., K. O'Neill, I. Shamatava, K. Sun, and K. D. Paulsen, "A fast and accurate representation of physically complete EMI response by a heterogeneous object," *IEEE Trans. GeoScience and Remote Sesning*, Vol. 43, 1736–1750, August 2005.
2. Sun, K., K. O'Neill, F. Shubitidze, I. Shamatava, and K. D. Paulsen, "Fast data-derived fundamental excitation models and implementation in UXO discrimination," *IEEE Trans. Geosciences and Remote Sensing*, Vol. 43, No. 11, 2573–2583, 2005.
3. Flammer, C., *Spheroidal Wave Functions*, Stanford, CA, Stanford Univ. press, 1957.
4. Marquardt, D., "An algorithm for least-squares estimation of non-linear parameters," *SIAM J. Appl. Math.*, Vol. 11, 431–441, 1963.

Bound States and Resonance States of H^- and He Embedded in Debye Plasma Environments

Sabyasachi Kar and Y. K. Ho
Institute of Atomic and Molecular Sciences, Academia Sinica, P.O. Box 23-166, Taipei 106, Taiwan, China

Abstract— The doubly excited S-, P- resonance states of H^- and the doubly-excited S-, P-, D-resonance states of He embedded in Debye plasma environments have been investigated for by calculating the density of resonance states using stabilization method. The ground states of H^- and the bound excited states of He atom in plasmas for various Debye lengths have also been estimated in the framework of Rayleigh-Ritz variational method and by using a quasi-random process. Highly correlated wave functions have been used to represent the correlation effects between the charged particles. Some of our calculated results on the bound states and resonance states are highlighted in this paper and the detail results will be presented at the conference.

The investigation on the effect of external environment on the atomic processes has gained considerable interest in recent years [1–15, references therein] with the renewed advancement in the experimental studies of highly charged ions, atomic systems. One may tentatively assume that the effects of charge neutral environments are due to plasma and may employ the Debye model to derive an effective screened Coulomb potential for the atomic system. The concept of Debye screening has been derived from the weakly coupled plasmas in which coupling constant are much smaller than 1. In the present work, we have investigated the ground states and $^1S^e$, $^{1,3}P^o$ resonance states of H^-, and the bound excited $^{1,3}S$, $^{1,3}P$, $^{1,3}D$ states and $^{1,3}S^e$, $^{1,3}P^o$, $^{1,3}D^e$, $^{1,3}P^e$, $^{1,3}D^o$ resonance states of helium atom in plasmas for various Debye lengths. We have used the effective screening of the form $\exp(-r/\lambda_D)/r$. A particular value of the Debye length λ_D corresponds to the range of the plasma conditions, as λ_D is a function of electron density and electron temperature. For the two-component plasmas near the thermodynamic equilibrium, the Debye length λ_D can be written as $\lambda_D = \left[4\pi(Z^*+1)e^2 n_e/k_B T_e\right]^{-1/2}$, where $Z^*$is the effective charge, n_e is the number density of the electron, k_B is the Boltzmann constant and T_e is the electron temperature. We have employed the wave function for the bound-excited S-, P-, D-states and $^{1,3}S^e$, $^{1,3}P^o$ resonance states calculations for the proposed systems of the form

$$\Psi = (1 \pm P_{12}) \sum_{i=1}^{N} C_i r_1^L P_L(\cos\theta_1) \exp\left[(-\alpha_i r_1 - \beta_i r_2 - \gamma_i r_{12})\omega\right], \tag{1}$$

where α_i, β_i, γ_i are the non-linear variation parameters to be chosen by using a quasi-random process [16], ω is a scaling constant, and P_{12} is the permutation operator on the indices 1 and 2. The stabilization method [17–19] that is a very simple and powerful method to extract resonance parameters has been used to calculate resonance energies, E_r and widths, Γ. To obtain the ground state of H^- and the bound excited-state energies (shown in Table 1) of He we first set $\omega = 1$. For the $^{1,3}D^e$, $^{1,3}P^e$, $^{1,3}D^o$ resonance states calculation, we have considered the following CI-type wave functions:

$$\Psi(r_1, r_2) = A \sum_{l_a l_b} \sum_{ij} C_{a_i, b_j} \eta_{a_i}(r_1) \eta_{b_j}(r_2) \times Y_{L l_a l_b}(r_1, r_2), \tag{2}$$

where

$$\eta_{a_i}(r) = r^{n_{a_i}} e^{-\alpha_{a_i}\omega r} \tag{3}$$

is a Slater-type orbital, C's are the coefficient to be determined, A is the antisymmetization operator, Y is the two-body spherical harmonics, i, j and a, b refer to channel indices. The electron-electron correlation is included by considering pseudostates. For the $^{1,3}D^e$, $^{1,3}P^e$, $^{1,3}D^o$ resonance states calculations, we vary ω and the eigenvalues for each given ω are calculated, and from which a stabilization plot is constructed.

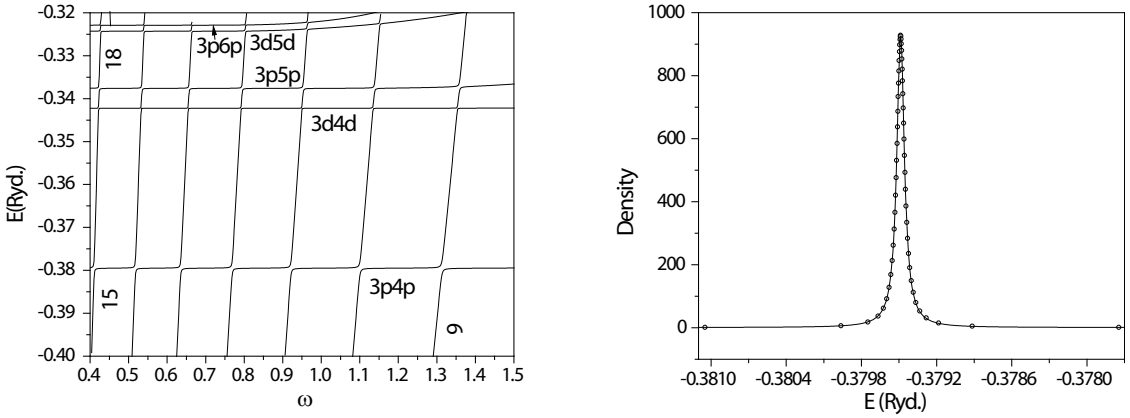
For all the resonance states calculations the variational parameter ω in the wave function will act as the reciprocal range of the 'soft" wall [17–19]. Varying ω from 0.3–1.0 for $^{1,3}S^e$, $^{1,3}P^o$ states and from 0.3–1.7 for $^{1,3}D^e$, $^{1,3}P^e$, $^{1,3}D^o$ states, we have computed the energy levels $E(\omega)$ for each Debye length by diagonalizing the atomic Hamiltonian with the above basis function and then

Table 1: Bound excited states energies of He under Debye screening.

λ_D	$1s^2\,{}^1S$	$1s2s\,{}^1S$	$1s2p\,{}^1P$	$1s3d\,{}^1D$	$He^+(1S)$
∞	−2.90372437700 −2.90372437703*	−2.14597404604 −2.14597404605*	−2.12384308647 −2.12384308650*	−2.055619 −2.055620*	−2.0000
50	−2.84418057552	−2.08725861976	−2.06510552556	−1.997868	−1.9603
10	−2.61485294693	−1.87503634005	−1.8527035426	−1.8073	−1.8073

* Best results

construct the stabilization diagrams (as shown in Figures 1(a) and 3(a)) by plotting $E(\omega)$ versus ω. If there is a resonance at energy E, the stabilized or slowly decreasing energy levels will appear in the stabilization plateau. The resonance parameters have been estimated by calculating the density of states in the vicinity of the stabilized energies in the stabilization plateau and then fitted the calculated density of states using a Lorentzian form.



(a) (b)

Figure 1: The stabilization plots for the ${}^1P^e$ states in (a) and the fitting (solid line) of the calculated density (circles) of states in (b) for the lowest ${}^1P^e$ resonances corresponding to 15th energy level in (a) for $\lambda_D = 30$.

To extract the resonance energy E_r and the resonance width Γ, we calculate the density of resonance states for a single energy level using the formula,

$$\rho_n(E) = \left| \frac{E_n(\omega_{i+1}) - E_n(\omega_{i-1})}{\omega_{i+1} - \omega_{i-1}} \right|^{-1}_{E_n(\omega_i)=E}, \tag{4}$$

where the index i is the i-th value for ω and the index n is for the n-th resonance. After calculating the density of resonance states $\rho_n(E)$ using formula (4), we fit it to the following Lorentzian form that yields resonance energy E_r and a total width Γ, with

$$\rho_n(E) = y_0 + \frac{A}{\pi} \frac{(\Gamma/2)}{(E - E_r)^2 + (\Gamma/2)^2}, \tag{5}$$

where y_0 is the baseline offset, A is the total area under the curve, E_r is the center of the peak, and Γ denotes the full width of the peak of the curve at half height.

Varying the Debye length λ_D from infinity to small values, different resonance parameters have been obtained for various Debye parameters. The lowest ${}^1S^e$ and ${}^{1,3}P^o$ resonances of H^-, a total of three ${}^1S^e$, five ${}^3S^e$, nine ${}^{1,3}P^o$, seven ${}^1D^e$, two ${}^3D^e$ of He atom below the $n = 2$ He^+ thresholds, and a total of five ${}^{1,3}P^e$, six ${}^{1,3}D^o$ resonances of He atom below the $n = 3$ He^+ thresholds have been estimated by calculating the density of states using stabilization method. The stabilization plots corresponding to $\lambda_D = 30$ for the five ${}^1P^e$ resonance states is presented in Figure 1 and the best fitting (least chi-square and the square of the correlation coefficient very close to 1) of the density of states for the lowest ${}^1P^e$ resonance (as shown in Figure 1(a)) are shown in Figure 1(b). The

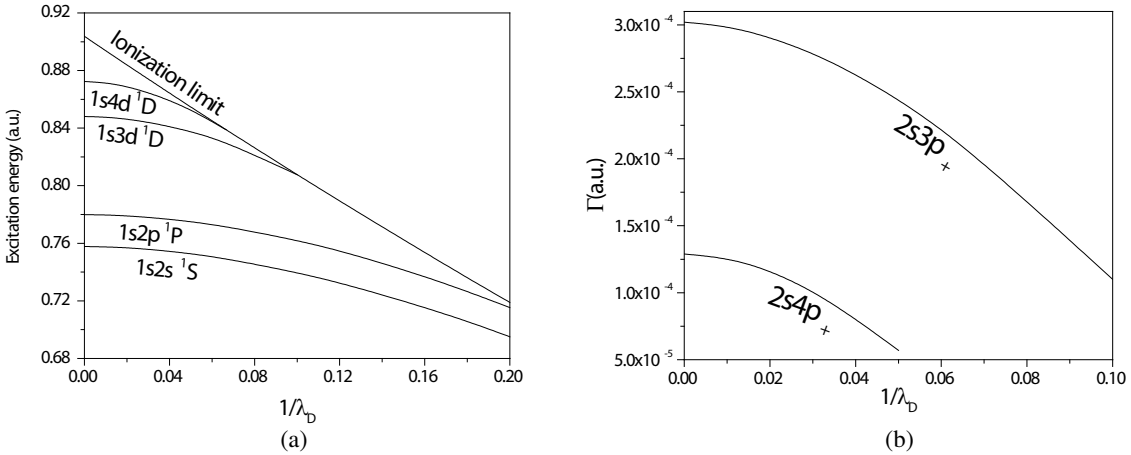


Figure 2: Excitation energies in (a) of the few singlet states of He and the resonance widths for the ${}^1P^o$ states of He for various Debye shielding parameters (b).

Figure 1(b) yields the resonance parameters (E_r, Γ) as (-0.37949, 0.0000735) Ryd. for the ${}^1P^e$ states corresponding to λ_D=30. Figure 2 shows our estimated excitation energies (in Figure 2(a)) for few singlet states of He and the resonance widths (as shown in Figure 2(b)) for the $2s3p_+$ $2s4p_+$ ${}^1P^o$ states of He. The stabilization diagram corresponding to the Debye length $\lambda_D = 15$ for the lowest ${}^1D^e$ states is presented in Figure 3(a) and the best fitting corresponding to 9th energy level using the Lorentzian form (5) is presented in Figure 3(b). From the Figure 3(b), we have obtained the resonance parameter (E_r, Γ) as (-1.03160, 0.00447) Ryd. for the lowest ${}^1D^e$ states corresponding to $\lambda_D = 15$. We have also observed the effect of plasma with and without electron-electron screening for S-, P-states of H^- and D^e states of He. We have also calculated the bound meta-stable ${}^{1,3}P^e$ and ${}^{1,3}D^o$ states of He below the $n = 2$ He^+ thresholds. The detail results will be presented at the conference. In addition, the wavelengths for the photo-absorption of helium atom from its ground state to the ${}^1P^o$ resonance states for different Debye lengths are will also be presented. We hope our findings will provide useful information to the research communities of atomic physics, plasma physics and astrophysics.

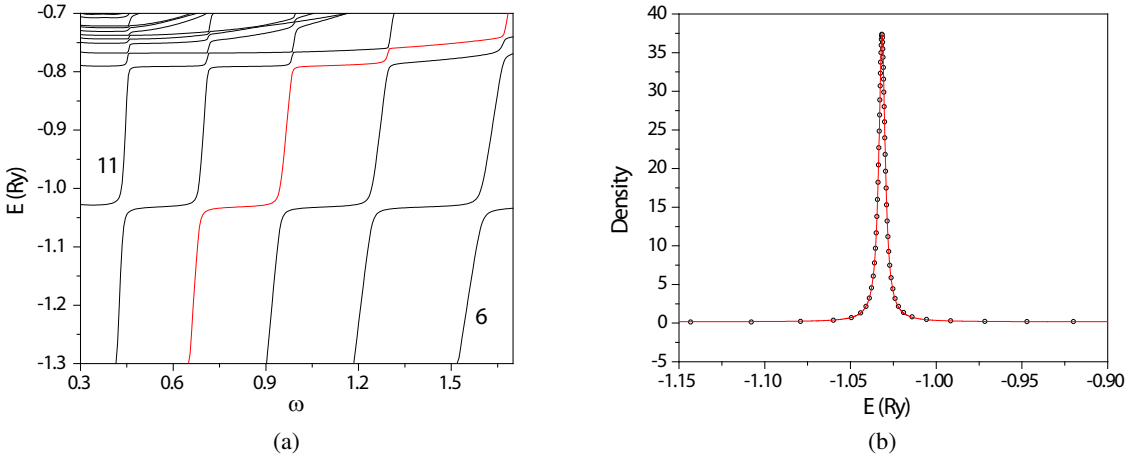


Figure 3: (a) The stabilization plot for the lowest ${}^1D^e$ state of the plasma-embedded He atom for $\lambda_D = 15$. (b) The calculated density (circles) and the fitted Lorentzian (solid line) from corresponding to the best fit for the 9th energy levels in (a).

ACKNOWLEDGMENT

The work is supported by the National Science Council of R.O.C..

REFERENCES

1. Wang, Z. and P. Winkler, "Pair function calculations for two-electron systems in model plasma environments," *Phys. Rev. A*, Vol. 52, No. 1, 216–220, 1995.
2. Zhang, L. and P. Winkler, "Negative hydrogen and helium in a variety of Debye plasmas," *Int. J. Quantum Chem.: Quantum Chem. Sym.*, Vol. 30, 1643–1650, 1996.
3. Winkler, P., "Detachment energies for a negative hydrogen ion embedded in a variety of Debye plasmas," *Phys. Rev. E*, Vol. 53, No. 5, 5517–5519, 1996.
4. Lopez, X., C. Sarasola, and J. M. Ugalde, "Transition energies and emission oscillator strengths of helium in model plasma environments," *J. Phys. Chem. A*, Vol. 101, 1804–1807, 1997.
5. Dai, S.-T., A. Solovyova, and P. Winkler, "Calculations of properties of screened He-like systems using correlated basis functions," *Phys. Rev. E*, Vol. 64, No. 1, 016408, 2001.
6. Saha, B., T. K. Mukherjee, P. K. Mukherjee, and G. H. F. Diercksen, "Variational calculations for the energy levels of confined two-electron atomic systems," *Theor. Chem. Acc.*, Vol. 108, 305–310, 2002.
7. Kar, S. and Y. K. Ho, "Auto-ionizing ${}^1S^e$ states of H^- in Debye plasmas environments," *Phys. Rev. E*, Vol. 70, No. 6, 066411-5, 2004.
8. Kar, S. and Y. K. Ho, "Doubly-excited $2s^2\,{}^1S^e$ resonance states of helium embedded in Debye plasmas," *Chem. Phys. Lett.*, Vol. 402, No. 4–6, 544–548, 2005.
9. Kar, S. and Y. K. Ho, "Electron affinity of hydrogen atom and a resonance state of the hydrogen negative ion embedded in Debye plasmas," *New J. Phys.*, Vol. 7, 141, 2005.
10. Okutsu, H., T. Sako, K. Yamanouchi, and G. H. F. Diercksen, "Electronic structure of atoms in laser plasmas: a Debye shielding approach," *J. Phys. B*, Vol. 38, 917–927, 2005.
11. Kar, S. and Y. K. Ho, "Doubly-excited $2s2p\,{}^{1,3}P^o$ resonance states of helium in dense plasmas," *Phys. Rev. A*, Vol. 72, No. 1, 0107003 (R), 2005.
12. Sil, A. N. and P. K. Mukherjee, "Effect of Debye plasmas on the doubly excited states of highly stripped ions," *Int. J. Quantum Chem.*, Vol. 102, 1061–1068, 2005.
13. Kar, S. and Y. K. Ho, "Bound-states of helium atom in dense plasmas," *Int. J. Quantum Chem.*, Vol. 106, 814–822, 2006.
14. Kar, S. and Y. K. Ho, "Doubly-excited inter-shell P-wave resonances of helium in weakly coupled plasmas," *J. Phys. B*, Vol. 39, 2445–2453, 2006.
15. Kar, S. and Y. K. Ho, "Bound D-states of helium atom under Debye screening," *Int. J. Quantum Chem.*, Vol. 107, in press, 2007.
16. Frolov, A. M., "Multi-box strategy for constructing highly accurate bound-state wave function for three-body systems," *Phys. Rev. E*, Vol. 64, 036704, 2001.
17. Mandelshtam, V. A., T. R. Ravuri, and H. S. Tailor, "Calculations of the density of resonance states using the stabilization method," *Phys. Rev. Lett.*, Vol. 70, 1932–1935, 1993.
18. Tan, S. S. and Y. K. Ho, "Determination of resonance energy and width by calculation of the density of states using the stabilization method," *Chin. J. Phys.*, Vol. 35, No. 6-I, 701–707, 1997.
19. Kar, S. and Y. K. Ho, "S-wave resonances in e^+-He scattering below the Ps ($n = 2$) excitation threshold," *J. Phys. B*, Vol. 37, 3177–3186, 2004.

The Use of the Virial Theorem and Sum-rules in Atomic Structure Calculations

Krishna Lamichhane and Peter Winkler

University of Nevada, Reno, NV 89557, USA

Abstract— Dilatation transformations have been used sucessfully to obtain eigenvalues and eigenfunctions of metastable states such as resonances in electron-atom scattering.The application of dilatation theory requires in general the expansion of the eigenfunctions and energies of dilated Hamiltonians in powers of the dilatation parameter. These expansions yield a set of sum rules which supply stationary and stability conditions which exact solutions will satisfy automatically. For approximate wave functions these expansions provide tools for optimization in a given parameter space. The first member in the set of sum rules is the quantum virial theorem which is particularly valuable to obtain the correct balance of potential and kinetic energy. Applied to resonances these expectation values are both complex numbers. Accurate calculations of properties of few electron systems are of interest for astrophysical plasma diagnostics. So far, most calculations have been performed on isolated atoms and molecules. In Debye-Hückel theory [4], the dilatation technique can be conveniently extended to allow for the inclusion of a model plasma environment.

Correlated wave functions for two-electron systems have widely been applied to atomic structure calculations in screening environments modeled by the Debye-Hückel theory, e.g., [3, 6]. The customary expansion of the eigensolutions as linear combinations of a finite number of correlated basis functions depends crucially on the size and choice of the basis set. In particular, moderately sized basis sets of Slater type orbitals require a careful, non-linear optimization of the exponents. The quantum virial theorem and a set of related sum rules provide tools not only to assess the quality of approximate calculations when the exact result is not known *a priori* but also to optimize the electronic wave function in a large space of nonlinear parameters. The present work studies the extension of the dilatation technique to ionic systems which are embedded in a Debye Plasma.

The satisfaction of the virial theorem is warranted for physical reasons so that the potential and kinetic energy contributions are well balanced. In its complex version, the quantum virial theorem and the corresponding complex sum rules provide in addition a means to optimize resonance energies and widths without having to strictly orthogonalize the eigenvectors to all lower solutions of the same symmetry. In complex, non-Hermitian quantum mechanics this is not only not possible but also not wanted.

Debye screening, originally developed for strong electrolytes is now used as a convenient model for plasmas at or near thermal equilibrium. In Debye plasmas, the interaction between the electrons and the atomic nucleus are screened as well as the electron-electron interaction. Two-electron systems offer both relatively simple spectra and the potential for highly accurate calculations which makes them valuable tools for plasma diagnostics.

Of course, electronic structure calculations of metastable states have been performed using a variety of different techniques: An extension into the complex plane is the hallmark of the dilatation transformation approach [1, 5, 7, 10–13] — also called rotated coordinate methods — as well as the Siegert approach [14], while other approaches solve the same problem entirely with real numerics [6]. In the present work we focus on the bound and resonance state calculations by augmenting the dilatation transformation approach with the implementation of the quantum virial theorem and the second order sum rule.

The dilatation transformation consists mainly in the replacement of all radial variables $\vec{r}$ by $i\theta\vec{r}$ in the Hamiltonian operator of the system. By this procedure the computational method is extended onto relevant parts of the complex energy plane which allows for the calculation of decaying states (i.e., resonances) as eigenstates of the transformed Hamiltonian which is no longer Hermitean [8]. The corresponding eigenvalues are complex-valued with its imaginary part related to the lifetime of the state. The departure from Hermiticity has consequences for the computational strategy. We lose, in particular, the well established rule of variational calculations: The lower the calculated energy the better is the wave function. One aims instead at a high degree of stability of the resonance eigenvalue against changes of the rotation angle θ. In order to calculate a resonance

eigenvalue it is required that this angle has a certain non-zero value but further rotation should not affect the eigenvalue. For the exact wave function this is satisfied. Approximate wave functions, however, do change upon further rotation. The computational goal has now become to determine the value of θ at which the eigenvalue is most stable. In the early days of complex rotation this region of stability was found by graphically following the trajectory of the resonance eigenvalue in the complex energy plane. As illustrative as this procedure may be, computationally more satisfactory is the use of the virial theorem and the sum rules which arise when one expands the rotated energy in terms of an infinitesimal *additional* rotation. For the Coulomb potential this was done first in our articles from the years 1977–78. The generalization to general potentials is given in [2], for the particular choice of Debye screening it is given in [3, 9].

In computational practice, one has also introduced complex valued rotation angles θ and effects this way not only the complex rotation but also some kind of real scaling. Hence, the dilatation technique which originally had been introduced for resonances can also be beneficial in bound state calculations. Bound states should not be affected by the rotation with a real rotation angle θ, but this is true only for exact wave functions. Finite basis set expansions are not immune to those rotations but they do benefit from the real scaling part. In bound state calculations, the satisfaction of the sum rules assures an increased degree of stationarity and stability of the result and provides at the same time a measure for the goodness of the wavefunction which is anyhow desirable to have.

REFERENCES

1. Brändas, E. and P. Froehlich, *Phys. Rev.*, Vol. A16, 2201, 1977.
2. Cederbaum, L. S. and P. Winkler, *Theor. Chim. Acta.*, Vol. 88, 257, 1994.
3. Dai, S.-T., A. Solovyova, and P. Winkler, *Phys. Rev. E*, Vol. 64, 016408, 2001.
4. Debye, P., *Polar Molecules*, Dover Publications, New York, 1929.
5. Doole, G., J. Nuttall, and R. Stagat, *Phys. Rev.*, Vol. A10, 1612, 1974.
6. Kar, S. and Y. K. Ho, *Phys. Rev.*, Vol. E70, 066411, 2004.
7. Moiseyev, N., P. R. Certain, and F. Weinhold, *Mol. Phys.*, Vol. 36, 1613, 1978.
8. Reinhardt, W. P., *Ann. Rev. Phys. Chem.*, Vol. 33, 223, 1982.
9. Wang, Z., P. Winkler, B. T. Pickup, and N. Elander, *Chemical Physics*, Vol. 135, 247, 1989.
10. Winkler, P., *Z. Physik*, Vol. A 283, 149, 1977.
11. Winkler, P. and R. Yaris, *J. Phys. B: Atom. Molec. Phys.*, Vol. 11, 1481, 1978.
12. Winkler P. and R. Yaris, *J. Phys. B: Atom. Molec. Phys.*, Vol. 11, 4257, 1978.
13. Yaris, R. and P. Winkler, *J. Phys. B: Atom. Molec. Phys.*, Vol. 11, 1475, 1978.
14. Yaris, R., R. Lovett, and P. Winkler, *Chem. Phys.*, Vol. 43, 29, 1979.

Line-broadening due to Plasma Fluctuations

Peter Winkler
Department of Physics, University of Nevada, Reno, NV 89557, USA

Abstract— A wide variety of plasmas are efficiently modelled by applying the Debye-Hückel model to account for the screening of the various electrostatic interactions in the system. In the present study, the model has been applied to a system which is close to but not quite at thermal equilibrium. It has been shown that small temporal fluctuations due to remaining ion motion can be interpreted as small local fluctuations within the system which may lead to sizeable contributions to line broadening.

1. INTRODUCTION

The observation of width and shape of spectral lines of atomic transitions from atoms or ions embedded in a plasma has early on been used as an important tool for the temperature and density diagnostics of plasmas [16]. The relevance of such observations has stimulated a large number of studies of the physical processes that contribute to features of spectral lines which allow for the extraction of a wealth of plasma information. A selection of relevant publications in the area is given by the following list [1, 8, 5, 11, 6, 15]. The present study draws attention to a contribution to line broadening which so far has not yet been documented and which will be called fluctuation line broadening in this study. While recent work of ours dealt with a plasma situation for which this effect — though larger by three orders of magnitude than the natural line width — is very small compared to other broadening mechanisms, we focus here on plasmas close to the pressure ionization limit. In this region fluctuation line broadening is much larger. Our study is based on weakly coupled plasmas which are customarily described by the Debye-Hückel screening model [3]. This particular screening model is convenient to obtain numerical results. All screening models predict a decrease in binding with increasing plasma effects. However, different electronic states respond differently to increasing plasma conditions, and these differences, while shifting all transition energies downward, leave footprints of their varying amounts. The effects are not noticeable in a plasma in which each radiating atom or ion experiences exactly the same screening condition. With some amount of temporal or local fluctuations of the screening environment, however, a related distribution of transition energies results. The omission of one sizable contribution leads to incorrect assignments for temperature, density or pressure.

A number of important experimental plasmas fall into the realm of Debye-Hückel plasmas: Gas-discharge plasmas, plasmas of thermonuclear fusion experiments, plasmas in the solar chromosphere and plasmas in interstellar space. These are weakly coupled plasmas with coupling constants much less than unity. The importance of properly including Debye screening in the evaluation of astrophysical plasma data has been emphasized by Spatschek [15] and, more recently, by Saha et al. [13].

The system we envision in the present study is a neutral plasma with positive ions embedded in an electron gas. In this system we encounter one rapid process (the thermalization of the free electrons) and a noticeably slower process (the thermalization of the positive ions which, typically, may have one or more bound electrons attached). As it is customarily done, we single out one of the positive ions as a central test system and attach the origin of the coordinate system to its center of mass. This test charge Q repels all other positive ions, and attracts both its own bound electrons and intruders from the surrounding electron gas. Thus, around our test charge the electron density n_e increases and the ion density n_i decreases: The test ion gathers a shielding cloud. A particularly simple example are the astrophysically abundant hydrogen plasmas which contain in addition to electrons and protons also neutral hydrogen and, as a rule, even an abundance of negative hydrogen ions. Emission lines from pure hydrogen plasmas are entirely due to the neutral hydrogen component. Consequently, the theory has first been applied to the bound states of neutral hydrogen atoms [12, 10]. The application to electron scattering resonances in a plasma has been developed following the latter of these references [17]. Before that, Scheibner et al. had recognized the need to screen the proton-proton interaction for scattering events occurring in astrophysical plasmas [14]. The extension to two-electron systems (in particular, the negative hydrogen ion) has been reported in [18]. Highly accurate computations of bound and resonant properties of two-electron systems are appearing with increasing frequency, e.g., [13, 7].

The effect of screening the electron-nucleus interaction is often expressed as "lowering of the continuum threshold", reflecting the fact that atoms and ions are less bound in a plasma environment than in vacuum and expressing also the observed disappearance of the upper members of spectral series [4]. In addition to these features, the Debye screening model — whenever it is applicable [15] — accounts also for the general shift of transition energies toward smaller values with increasing screening.

The Debye length, given in atomic length units u, is defined as $D = \left(\frac{T}{4\pi n_0 e^2}\right)^{1/2} \approx 740\sqrt{\frac{T \text{ in eV}}{n_0 \text{ in cm}^{-3}}}$, where the reduced temperature T is often close to the ion temperature. Since the Debye length is a function of both temperature T and electron density n_0, a single value of D represents a wide range of different plasma conditions.

In the investigation of plasmas consisting of atomic ions and electrons, it is essential to modify the interaction potential between all particles. In this work, we have replaced all Coulomb interactions by screened interactions of the Debye type. These modifications affect one-electron properties as well as properties involving more than one electron, in particular, the electron-electron interactions.

In addition to the natural line width which is usually quite negligible, the most important plasma line broadening effects are

(a) Resonance broadening caused by the interaction with the radiation field leading to induced absorption and emission thus decreasing the lifetimes of electronic states.

(b) Collisional broadening caused by collisional excitation and de-excitation and may also greatly reduce the lifetime of electronic states.

(c) Doppler broadening caused by the motion of radiating ions in a plasma leads to line widths which are usually much greater than the natural line width.

(d) Stark broadening [9] assumes the presence of strong, fluctuating micro-fields which add to the broadening of lines in addition to the splitting into several possible sublevels when the spherical symmetry of the environment is broken, for instance, by external fields.

If we assume now that the thermalization of the free electrons occurs so fast that even when the ions have not yet quite reached thermal equilibrium the electrons can and will adjust momentarily to the slow changes that are introduced by the motion of the ions and achieve temporary equilibrium temperature and density. In this model we will encounter slow variations of the screening parameter D which lead to small differences in the energies of the electronic eigenstates. As the time rate of such changes of the screening environment is assumed to be sufficiently slow, the changes by themselves will not introduce transitions between electronic states but lead to a broadening of the resulting spectral lines. The use of a time dependent theory is in this case irrelevant. An alternative to observing slow time variations of the screening potential experienced by a single central ion consists in viewing the whole ensemble of ions at once and find their respective screening parameters distributed about an average value D_0.

Employing correlated wave functions as they have been studied in [2], we have determined dipole transition energies of several allowed transitions in helium which is embedded in plasmas of differing strengths. While the calculation of the helium energies requires quite sophisticated wave functions the resulting energy differences T can be fitted to simple exponential expressions of the form $T(D) = E_2 - E_1 = -ae^{-D/b} + c$, where the fitting interval has been restricted to values of $D < 20\ u$. (u is the atomic length unit, $1u = 0.53 \times 10^{-10}$ m). The region above is of limited interest because the function does not change enough to induce noticeable line broadening [19].

For the lowest optical transition the transition energy T as a function of the screening length D is given by the formula above with the following parameters $a = 0.99607$ eV, $b = 1.58367\ u$, $c = 0.76528$ eV. When we assume that the Debye lengths of our sample fluctuate by $\pm\varepsilon$ about a mean value D, then there results an uncertainty in the transition energy $T(D \pm \varepsilon)$ given by $\Delta T = 2ae^{-\frac{D}{b}} \sinh\frac{\varepsilon}{b}$. This constitutes the contribution to line broadening exclusively due to fluctuations in the sample. Whether this contribution is relevant or can be safely neglected depends on the magnitude of other line broadening effects present. For the helium atom in plasmas the calculated width under reasonable physical conditions can well amount to several eV.

The Debye-Hückel approach to screening in plasmas has been known to explain the decrease of binding of electronic states with increased screening — an effect that is often addressed as "lowering of the continuum threshold". It has also long been known that the theory explains the shift of transition lines toward lower energies with increased screening. Here, it has been shown that it may as well result in a sizeable contribution to the line width. Other features observed in studies of line profiles have been left out in the present work but can be included. If the

experimental conditions are such that the effects of fluctuation broadening are comparable to the other contributions one better includes the former in the diagnostic process.

REFERENCES

1. Breene, R. G., *The Shift and Shape of Spectral Lines*, Pergamon, London, 1961.
2. Dai, S.-T., A. Solovyova, and P. Winkler, *Phys. Rev. E*, Vol. 64, 016408, 2001.
3. Debye, P., *Polar Molecules*, Dover Publications, New York, 1929.
4. D'yachkov, L. G., G. A. Kobzev, P. M. Pankratov, and J. Quant, *Spectrosc. Radiat. Transfer*, Vol. 44, 123–128, 1990.
5. Griem, H., *Spectral Line Broadening in Plasmas*, Academic Press, New York, 1974.
6. Hutchinson, I. H., *Principles of Plasma Diagnostics*, Cambridge University Press, 1984.
7. Kar, S. and Y. K. Ho, *Phys. Rev. E*, Vol. 70, 066411, 2004.
8. Kuhn, H. G., *Atomic Spectra*, Academic Press, New York, 1962.
9. Mancini, R. C., D. P. Kilcrease, L. A. Woltz, and C. F. Hooper, *Computer Physics Comm.*, Vol. 63, 314, 1991.
10. Nauenberg, M., *Phys. Rev. A*, Vol. 8, 2217, 1973.
11. Nicholson, D. R., *Introduction to Plasma Theory*, John Wiley & Sons, New York, 1983.
12. Rogers, F. J., H. C. Graboske, and D. J. Harwood, *Phys. Rev. A*, Vol. 1, 1557, 1970.
13. Saha, B., P. K. Mukherjee, and G. H. F. Diercksen, Astronom, Astrophys, Vol. 396, 337, 2002.
14. Scheibner, K., J. Weisheit, and N. Lane, *Phys. Rev. A*, Vol. 35, 1252, 1987.
15. Spatschek, K. H., Theoretische Plasmaphysik, Teubner Studienbücher, B. G. Teubner, Stuttgart, 79ff, 1990.
16. Tonks, L. and I. Langmuir, *Phys. Rev.*, Vol. 33, 195, 1929.
17. Wang, Z. and P. Winkler, *J. Quantum Chem.*, Vol. S 23, 89–94, 1989.
18. Wang, Z. and P. Winkler, *Phys. Rev A*, Vol. 52, 216, 1995.
19. Zhang, L. and P. Winkler, *Chem. Phys.*, Vol. 329, 338–342, 2006.

Stochastic Perturbation of Parabolic Law Optical Solitons

Anjan Biswas
Department of Applied Mathematics and Theoretical Physics, Delaware State University
Dever, DE 19901-2277, USA

Abstract— The stochastic perturbation of optical solitons due to parabolic law nonlinearity is studied in this paper with the aid of soliton perturbation theory. The corresponding Langevin equation is derived and it is proved that the soliton propagates with a fixed mean velocity.

1. INTRODUCTION

The dynamics of solitons propagating through optical fibers is governed by Nonlinear Schrödinger Equation(NLSE) [2]. It is well known that the pulse dynamics are not accurately described the pure NLSE. It is therefore necessary, from practical consideration, to take into account the perturbation terms.

Besides the deterministic type perturbations one also needs to take into account, from realistic situations, stochastic type perturbation. These effects can be accounted from three basic sources [2]:

1. Stochasticity associated with the chaotic nature of the initial pulse due to partial coherence of the laser generated radiation.
2. Stochasticity due to random nonuniformities in the optical fibers like the fluctuations in the values of dielectric constant the random variations of the fiber diameter and more.
3. The chaotic field caused by a dynamic stochasticity might arise from a periodic modulation of the system parameters or when a periodic array of pulses propagate in a fiber optic resonator.

Thus, stochasticity is inevitable in optical soliton communications. In this paper the NLSE with parabolic law nonlinearity is going to be studied in presence of deterministic as well as stochastic perturbation terms. The dimensionless form of NLSE with parabolic law nonlinearity is given by

$$iq_t + \frac{1}{2}q_{xx} + \left(|q|^2 + \nu|q|^4\right)q = 0 \tag{1}$$

whose solution is given by

$$q(x,t) = \frac{A}{[1 + a\cosh\{B(x - \bar{x}(t))\}]^{\frac{1}{2}}} e^{i(-\kappa x + \omega t + \sigma_0)} \tag{2}$$

where

$$\frac{d\bar{x}}{dt} = \upsilon = -\kappa \tag{3}$$

$$\omega = \frac{A^2}{4} - \frac{\kappa^2}{2} \tag{4}$$

$$B = \sqrt{2}A \tag{5}$$

$$a = \sqrt{1 + \frac{4}{3}\nu A^2} \tag{6}$$

Here, A is the amplitude of the soliton, B is the width, κ is the frequency, ω is the wave number, $\bar{x}$ is the center position of the soliton and σ_0 is the center of phase of the soliton.

Considering the effects of perturbation on the propagation of solitons through optical fibers, (1) is modified to [1]

$$iq_t + \frac{1}{2}q_{xx} + \left(|q|^2 + \nu|q|^4\right)q = i\epsilon R \tag{7}$$

where

$$R = \delta|q|^{2m}q + \beta q_{xx} - \gamma q_{xxx} + \lambda\left(|q|^2 q\right)_x + \nu\left(|q|^2\right)_x q + \sigma(x,t) \tag{8}$$

For the perturbation terms, $\delta < 0$ is the nonlinear damping coefficient, β is the bandpass filtering term. Also, λ is the self-steepening coefficient for short pulses, ν is the higher order dispersion coefficient and γ is the coefficient of the third order dispersion.

The amplifiers, although needed to restore the soliton energy, introduces noise originating from amplified spontaneous emission (ASE). To study the impact of noise on soliton evolution, the evolution of the mean free velocity of the soliton due to ASE will be studied in this paper. In case of lumped amplification, soliton are perturbed by ASE in a discrete fashion at the location of the amplifiers. It can be assumed that noise is distributed all along the fiber length since the amplifier spacing satisfies $z_a \ll 1$. In (8), $\sigma(x,t)$ represents the Markovian stochastic process with Gaussian statistics and is assumed that $\sigma(x,t)$ is a function of t only so that $\sigma(x,t) = \sigma(t)$. Now, the complex stochastic term $\sigma(t)$ can be decomposed into real and imaginary parts as [3]

$$\sigma(t) = \sigma_1(t) + i\sigma_2(t) \tag{9}$$

is further assumed to be independently delta correlated in both $\sigma_1(t)$ and $\sigma_2(t)$ with

$$\langle \sigma_1(t) \rangle = \langle \sigma_2(t) \rangle = \langle \sigma_1(t)\sigma_2(t') \rangle = 0 \tag{10}$$

$$\langle \sigma_1(t)\sigma_1(t') \rangle = 2D_1\delta(t-t') \tag{11}$$

$$\langle \sigma_2(t)\sigma_2(t') \rangle = 2D_2\delta(t-t') \tag{12}$$

where D_1 and D_2 are related to the ASE spectral density. In this paper, it is assumed that $D_1 = D_2 = D$. Thus,

$$\langle \sigma(t) \rangle = 0 \tag{13}$$

and

$$\langle \sigma(t)\sigma(t') \rangle = 2D\delta(t-t') \tag{14}$$

In soliton units, one gets [2],

$$D = \frac{F_n F_G}{N_{ph} z_a} \tag{15}$$

where F_n is the amplifier noise figure, while

$$F_G = \frac{(G-1)^2}{G \ln G} \tag{16}$$

is related to the amplifier gain G and finally N_{ph} is the average number of photons in the pulse propagating as a fundamental soliton.

2. MATHEMATICAL ANALYSIS

The NLSE with parabolic law nonlinearity has three integrals of motion which are energy (E), linear momentum (M) and the Hamiltonian (H). However, in this paper, only the first two of them will serve our purpose. They are respectively by [1, 2]

$$E = \int_{-\infty}^{\infty} |q|^2 dx = \frac{A^2}{aB} F\left(1, 1, \frac{3}{2}; \frac{a-1}{2a}\right) B\left(1, \frac{1}{2}\right) \tag{17}$$

$$M = \frac{i}{2} \int_{-\infty}^{\infty} (qq_x^* - q^* q_x)\, dx = \frac{\kappa A^2}{aB} F\left(1, 1, \frac{3}{2}; \frac{a-1}{2a}\right) B\left(1, \frac{1}{2}\right) \tag{18}$$

where $F(\alpha, \lambda; \gamma; z)$ is the Gauss' Hypergeometric function and $B(l, m)$ is the beta function. Using these integrals of motion, one can write

$$\begin{aligned}\frac{dA}{dt} &= \frac{\epsilon\delta A^{2m+1}}{2^m a^{m+1}} F\left(m+1, m+1, m+\frac{3}{2}; \frac{a-1}{2a}\right) B\left(m+1, \frac{1}{2}\right) \\ &+\frac{\epsilon\sqrt{2}\beta A}{a}\left[\kappa^2 F\left(1,1,\frac{3}{2};\frac{a-1}{2a}\right) B\left(1,\frac{1}{2}\right) - 2A^2 F\left(3,1,\frac{5}{2};\frac{a-1}{2a}\right) B\left(1,\frac{3}{2}\right)\right] \\ &+\epsilon a^2 A\sqrt{2}\left[\sigma_1 \int_{-\infty}^{\infty} \frac{\cos\phi}{(1+a\cosh\tau)^{\frac{1}{2}}} dx + \sigma_2 \int_{-\infty}^{\infty} \frac{\sin\phi}{(1+a\cosh\tau)^{\frac{1}{2}}} dx\right] \qquad (19)\end{aligned}$$

$$\begin{aligned}\frac{d\kappa}{dt} &= -\frac{\epsilon\beta\kappa A^2}{2} \frac{F\left(3,2,3;\frac{a-1}{2a}\right)}{F\left(1,1,\frac{3}{2};\frac{a-1}{2a}\right)} = \frac{B(2,1)}{B\left(2,\frac{1}{2}\right)} \\ &-\frac{\epsilon\sqrt{2}}{AE}\int_{-\infty}^{\infty}\left[\frac{aB(\sigma_2\cos\phi - \sigma_1\sin\phi)\sinh\tau}{(1+a\cosh\tau)^{\frac{3}{2}}} - \frac{2\kappa(\sigma_1\cos\phi + \sigma_2\sin\phi)}{(1+a\cosh\tau)^{\frac{1}{2}}}\right] dx \qquad (20)\end{aligned}$$

where $\tau = B(x - \bar{x})$ and $\phi = -\kappa x + \omega t + \sigma_0$. If the terms with σ_1 and σ_2, in (19) and (20), are suppressed, the resulting dynamical system has a stable fixed point, namely a *sink*. Now, linearizing the dynamical system about this fixed point gives, after simplification [3]

$$\frac{dA}{dt} = -\epsilon\left(A^{2m+1} - \frac{\xi}{\bar{A}}\right) \qquad (21)$$

$$\frac{d\kappa}{dt} = -\epsilon\left[\kappa - \zeta(1 + A - \kappa)\right] \qquad (22)$$

where $\bar{A}$ is the fixed point of the amplitude while

$$\xi = \sigma_1 \int_{-\infty}^{\infty} \frac{\cos\phi}{(1+a\cosh\tau)^{\frac{1}{2}}} dx + \sigma_2 \int_{-\infty}^{\infty} \frac{\sin\phi}{(1+a\cosh\tau)^{\frac{1}{2}}} dx \qquad (23)$$

and

$$\zeta = \int_{-\infty}^{\infty}\left[\frac{aB(\sigma_2\cos\phi - \sigma_1\sin\phi)\sinh\tau}{(1+a\cosh\tau)^{\frac{3}{2}}} - \frac{2\kappa(\sigma_1\cos\phi + \sigma_2\sin\phi)}{(1+a\cosh\tau)^{\frac{1}{2}}}\right] dx \qquad (24)$$

Equation (21) and (22) are called the *Langevin* equations which will now be analyzed to compute the soliton mean drift velocity of the soliton. If the soliton parameters are chosen such that ζA is small, then (22) yields [2, 3]

$$\frac{d\kappa}{dt} = -\epsilon[\kappa - \zeta(1-k)] \qquad (25)$$

One can solve (25) for κ and eventually the mean drift velocity of the soliton can be obtained. Assuming that σ is a Gaussian stochastic variable and using the initial condition $\kappa(0) = 0$ yields

$$\langle\kappa(t)\rangle = -\frac{D}{1-D}\left\{1 - e^{-\epsilon(1-D)t}\right\} \qquad (26)$$

which leads to

$$\lim_{t\to\infty}\langle\upsilon(t)\rangle = \lim_{t\to\infty}\langle\kappa(t)\rangle = \frac{D}{1-D} \qquad (27)$$

Thus, for large t, $\langle\upsilon(t)\rangle$ approaches a constant value provided $D < 1$. Thus, the soliton mean with drift approaches a constant for a large time.

3. CONCLUSIONS

In this paper, the dynamics of optical solitons with parabolic law nonlinearity in presence of perturbation terms, both deterministic as well as of stochastic, are studied. The Langevin equation were derived and the corresponding parameter dynamics was studied. The mean drift velocity of the soliton was obtained.

REFERENCES

1. Biswas, A., "Quasi-stationary non-Kerr law optical soliton," *Optical Fiber Technology*, Vol. 9, Issue 4, 224–259, 2003.
2. Biswas, A. and S. Konar, *Introduction to Non-Kerr Law Optical Solitons*, CRC Press, Boca Raton, FL, 2007.
3. Elgin, J. N., "Stochastic perturbations of optical solitons," *Optics Letters*, Vol. 18, No. 1, 10–12, 1993.

Dynamic Response of Discrete Fiber Raman Amplifiers to Multi-channel Randomly Variable Packet Traffic

Johann Gest and Lawrence R. Chen
Department of Electrical and Computer Engineering, McGill University, Montreal, QC H3A 2A7 Canada

Abstract— We investigate the dynamic response and statistical nature of gain transients in multi-channel discrete fiber Raman amplifiers that arise from random packet-to-packet input power variations. We assess the efficiency of gain-clamping for controlling such transients.

1. INTRODUCTION

Discrete fiber Raman amplifiers (DFRAs) are a potential enabling technology for increasing the capacity of dynamic optical networks such as the agile all-photonic network (AAPN) [1]. DFRAs have several advantages including, among others, providing amplification at bandwidths generally not available with doped fiber amplifiers. However, they suffer from power transients in the same way as Erbium-doped fiber amplifiers (EDFAs) whereby variations in the total input power, usually resulting from channel add/drop, will cause gain or power transients and hence deteriorate system performance [2]. All-optical gain-clamping (AOGC), which consists in creating a lasing signal outside the bandwidth of the channels being amplified, has been investigated as a means to clamp the amplifier gain and mitigate power transients.

Although the dynamic response of DFRAs has been investigated [3], these have focused largely on circuit-switched network scenarios within the context of channel add/drop. On the other hand, there have been no studies on the statistical nature of the gain variations that arise due to random power fluctuations. In this paper, we study theoretically the gain transients that arise when a multi-channel DFRA is fed by randomly variable packet-to-packet traffic, i.e., when the input power of the different channels varies in a random fashion on a packet-to-packet basis. Such power variations arise when the packets traverse different paths in a network and experience different loss/amplification or transmission impairments. We also assess the efficiency of AOGC for controlling such transients and compare the performance of unclamped vs. gain-clamped DFRAs for different packet durations, amplifier operating regimes, and fluctuations in the input power.

2. SIMULATIONS

The model used to simulate the dynamic response of the DFRAs takes into account all pump-to-pump, signal-to-signal, and pump-to-signal interactions between forward and backward propagating channels; it is solved numerically using the average power analysis technique [4].

In our simulations, we consider a DFRA comprising 5 km of highly nonlinear fiber (with an attenuation coefficient of $0.61\,\mathrm{dB{\cdot}km^{-1}}$ at 1550 nm and an effective Raman gain spectrum with a peak value of $\approx 6\,\mathrm{km^{-1}{\cdot}}W^{-1}$ at a frequency shift of 14.5 THz) that is backward pumped by three pumps at 1435 nm, 1450 nm and 1480 nm. In the case of the gain-clamped DFRA, two fiber Bragg gratings located at 1625 nm, which induces the least gain tilt relative to the unclamped case [5], are used. The pump powers are adjusted so that both unclamped and gain-clamped amplifiers provide the same gain (16 dB) when they are operated in the small-signal regime.

We consider 64 input channels spaced by 50 GHz from 1560 nm to 1585.6 nm. One of the 64 channels is set as a non-varying probe channel while the other 63 channels have time-varying input powers. The input power of each of the 63 channels is varied on a packet-to-packet basis randomly and independently according to a Gaussian distribution. The mean of the distribution is set by the operational regime of the amplifier: for operation in the small signal regime, the per channel input power (PCIP) is −30 dBm and near saturation, the PCIP is −20.9 dBm. The standard deviation (s) of the distribution is set at 1% and 5% of the PCIP. The packets are assumed to be aligned in time which corresponds to a worst-case scenario. We then examine the gain transients for the probe channel located at midband ($\lambda_m = 1575.2$ nm). Each simulation consists of 5 packets, i.e., the input power for each of the 63 channels varies 5 times. We set the packet durations at 25 µs and 100 µs. In order to study the statistical behaviour of the amplifier performance, we perform 500 simulations for each case considered, i.e., for a given operational regime, a given input power distribution, and a given packet duration.

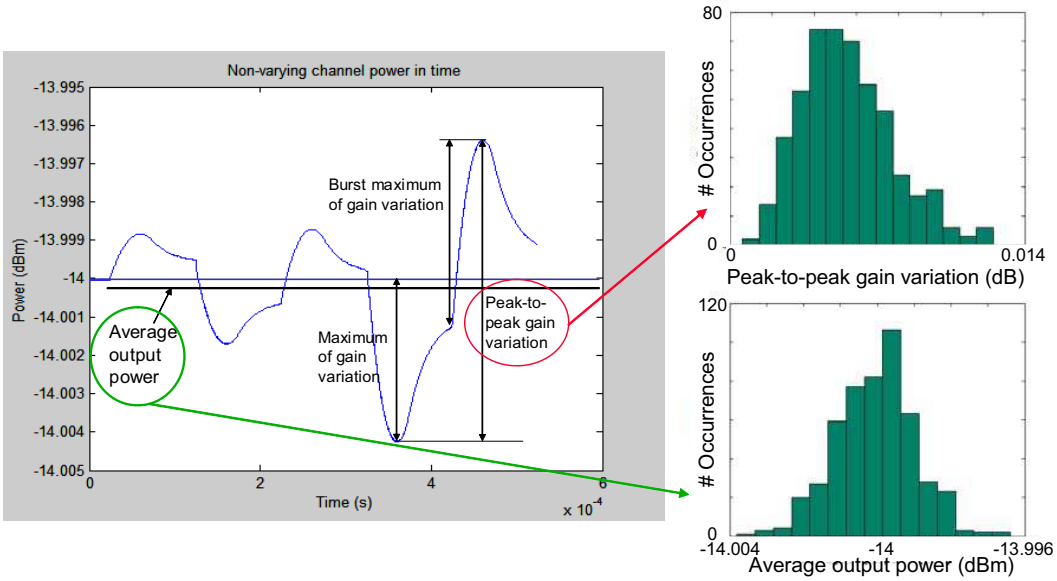


Figure 1: Left: Typical simulation result showing the variations in the output power of the probe channel due to fluctuations in the input power of the other 63 channels over 5 packets. Right: Histograms for the two comparison parameters obtained over the 500 simulations.

3. RESULTS AND DISCUSSION

Figure 1 shows the variations in the probe channel over 5 packets from a typical simulation. We are interested in the following two comparison parameters: (1) the output power of the probe channel averaged over the 5 packets (which characterizes the average gain of the DFRA) and (2) the peak-to-peak gain variation (ΔG_{pp}). Ideally, the variations in the other 63 channels should not have an impact on the probe channel in that the amplifier should provide the required gain and ΔG_{pp} should be minimal. We use the 500 simulations to produce histograms for these two comparison parameters from which we extract the average value and variance.

The results are summarized in Tables 1 and 2. As expected, the mean and the variance of both comparison parameters increase with s with one exception: in the small-signal regime, the mean value of the probe channel average output power is independent of s. This mean value is also the same for both unclamped and gain-clamped amplifiers so that both amplifiers are capable of providing the required gain in the presence of input power fluctuations. The mean and the variance

Table 1: Mean and variance of the distribution of the comparison parameters for both DFRAs operated in small-signal and near saturation with a packet duration of 25 µs.

			Unclamped DFRA		Gain-clamped DFRA	
Operational regime	Standard deviation (s)	Comparison parameters	Mean	Variance	Mean	Variance
Small-signal	1%	Average output power (dBm)	-14	1.35e-8	-14	0
		Peak-to-peak gain variation (dB)	5.67e-4	4.40e-8	4.22e-4	2.67e-8
	5%	Average output power (dBm)	-14.00	3.10e-7	-14.00	3.64e-8
		Peak-to-peak gain variation (dB)	0.0027	1.15e-6	0.0021	6.69e-7
Near saturation	1%	Average output power (dBm)	-6.29	2.59e-7	-5.09	8.68e-8
		Peak-to-peak gain variation (dB)	0.0028	1.03e-6	0.0032	1.42e-6
	5%	Average output power (dBm)	-6.29	7.15e-6	-5.09	2.27e-6
		Peak-to-peak gain variation (dB)	0.0145	3.14e-5	0.0151	3.21e-5

Table 2: Mean and variance of the distribution of the comparison parameters for both DFRAs operated in small-signal and near saturation with a packet duration of 100 µs.

			Unclamped DFRA		Gain-clamped DFRA	
Operational regime	Standard deviation (s)	Comparison parameters	Mean	Variance	Mean	Variance
Small-signal	1%	Average output power (dBm)	-14	2.33e-8	13.99	0
		Peak-to-peak gain variation (dB)	8.10e-4	8.54e-8	5.26e-4	4.65e-8
	5%	Average output power (dBm)	-14.00	4.92e-7	-14.00	1.73e-8
		Peak-to-peak gain variation (dB)	0.0040	2.36e-6	0.0027	1.16e-6
Near saturation	1%	Average output power (dBm)	-6.29	6.94e-7	-5.09	4.05e-8
		Peak-to-peak gain variation (dB)	0.0038	1.58e-6	0.0037	1.90e-6
	5%	Average output power (dBm)	-6.29	9.64e-5	-5.09	4.09e-6
		Peak-to-peak gain variation (dB)	0.0195	5.02e-5	0.0182	4.95e-5

of the comparison parameters also increase when the DFRAs operate near saturation.

When the packet duration is 25 µs and the DFRAs are operated in the small-signal regime, AOGC is efficient to control the gain transients. With $s = 5\%$, the mean of ΔG_{pp} is 0.0021 dB for the gain-clamped DFRA in comparison to 0.0027 dB for the unclamped DFRA. Although the difference between both amplifiers is small in terms of the mean values, AOGC induces a smaller variance in the distribution of ΔG_{pp} and in fact, the variance in the case of the unclamped amplifier is almost double that in the case of the gain-clamped amplifier. When the DFRA operates near saturation with the same packet duration, AOGC succeeds in reducing the mean and the variance of the average output power distribution but it induces larger and more scattered peak-to-peak gain variations independently of s. For example, with $s = 5\%$, the mean and the variance of ΔG_{pp} are 0.0151 dB and $3.21 \cdot 10^{-5}$ dB^2 respectively, for the gain-clamped DFRA compared to 0.0145 dB and $3.14 \cdot 10^{-5}$ dB^2 for the unclamped DFRA. When the packet duration is increased to 100 µs, AOGC is effective at reducing both the mean value and the scattering of the average output power and ΔG_{pp} independently of the operational regime. We also notice that the difference in the amplifiers performances increases with s, except the variance of ΔG_{pp} that decreases with s.

4. SUMMARY

In summary, we have investigated the statistical nature of gain transients in multi-channel DFRAs that arise from random packet-to-packet input power variations. We consider the worst case scenario for a 64 channel system in which the input power of 63 of the channels fluctuates randomly. We compare the performance of both unclamped and gain-clamped amplifiers for different operational regimes and packet durations. The performance of both amplifiers is typically the same, though small variations exist depending on the operational regime or packet duration. Thus, we do not believe that fluctuations in the input power of the signals to a DFRA that would be typical in an AAPN would have a significant impact on amplifier and overall system performance.

ACKNOWLEDGMENT

This research was supported in part by the NSERC Canada and industrial partners, through the Agile All-Photonic Networks (AAPN) Research Program.

REFERENCES

1. Mason, L., A. Vinokurov, N. Zhao, and D. V. Plant, "Topological design and dimensioning of agile all-photonic networks," *Computer Networks*, Vol. 50, 268–287, 2006.
2. Chen, C.-J. and W.-S. Wong, "Transient effects in saturated raman amplifiers," *Electron. Lett.*, Vol. 37, No. 6, 371-372, 2001.
3. Karásek, M. and M. Menif, "Channel addition/removal response in raman fiber amplifiers:

modeling and experimentation," *IEEE/OSA J. Lightwave Technol.*, Vol. 20, No. 9, 1680-1687, 2002.

4. Gest, J. and L. R. Chen, "Analysis of gain transients in cascades of discrete fiber raman amplifiers," *Proceedings SPIE*, Vol. 5970, 2005.
5. Zambelis, T., N. Gryspolakis, and L. R. Chen, "Design of all-optical gain-clamped discrete fiber raman amplifierns," *Proceedings of the Lasers and Electro-Optics Society Annual Meeting*, Tucson, USA, October 2003.

A Faster Approximation of Forward Scattering by Gradient-index Particles

Xiangzhen Li, **Xiang'e Han**, **Renxian Li**, **and Huifen Jiang**
Xidian University, 710071, China

Abstract— By means of geometrical optics we present an approximation algorithm with which to accelerate the computation of scattering intensity distribution within a forward angular range (0°–60°) for gradient-index particles illuminated by acollimated ted incident beam. The incident angle of the reflected light is derived from the scattering angle to improve the approximation precision. This method proves effective for transparent particles with size parameters larger than 75 but fails to give good approximation results at scattering angles at which refractive rays are absent. With the given index model, the geometrical optics approximation is effective only for forward small angles, typically less than 10° or so when the refractive index difference of a particle is greater than a certain value.

1. INTRODUCTION

Gradient-index (GRIN) particles have found many applications in modern industries. For instance [1], shell and the continuous refractive index models are used for the crystalline lens; a waveguide model is considered for human phot-Oreceptors; the atmosphere of the earth has a refractive index that decrease with height because the density decreases at higher altitudes, and the GRIN models are used for the optical fiber widely. To probe them optically, one needs to calculate the theoretical scattering intensities of particles of different sizes.

Studies [2] of approximation methods of the scattering light pattern for the homogeneous or coated particles have been carried out by several researchers. By means of a geometrical-optics approximation (GOA), Glantschnig and Chen derived a formula with which to calculate scattering intensity in the forward angular range for water droplets. A. Belafhal et al. [3] studied light scattering by absorbing and nonabsorbing spheroidal particles in the Wentzel-Kramers-Brillouin approximation and developed a new mathematical description of light scattering. Min et al. [4] calculated the optical cross sections of size shape distributions of spheroids, and Grynko and Shkuratov [5] studied the scattering characteristics of semitransparent particles faceted with various shapes. Xu and Cai [6] combined FD theory with geometrical optics and proposed the whole geometrical approximation method to accelerate the computation of the light's intensity for homogeneous absorbent particles and coated particles. However, geometrical-optics approximation of forward scattering by GRIN particles is never be researched. In this paper a more general method of GOA is developed that can be applied in the calculation of GRIN particles.

2. GEOMETRICAL-OPTICS APPROXIMATION OF GRIN PARTICLES

From a geometrical-optics viewpoint, scattered light is considered to be the superposition of diffracted, refracted, and reflected fractions. That is,

$$s = s_{diffraction} + s_{refraction} + s_{reflection} \tag{1}$$

Xu [7] considered that 95 percent of the scattering energy is produced by reflection and first refraction, therefore other emerging rays can be neglected. As indicated by Fig. 1, m_0, m_1 are the refractive indices of the center and surface. In this paper the refractive index of the environment m_s equals 1.0.

In 1637 Descartes [8] described geometrically the relationship between incident and scattering angles for a uniform sphere:

$$\theta = 2\tau - 2(N-1)\tau' \tag{2}$$

where τ represents the complement of the impact angle, τ' is the complement of the internal refracted angle, and N is the number of light-sphere interactions (see Fig. 1). If the refractive index of the sphere is not uniform, the Descartes law (Eq. (1)) is no longer applicable.

For a sphere with spherically symmetric refractive index $m(\widetilde{r})$ immersed in a uniform medium, the relationship between incident and scattering angles for a spherically symmetric particle can

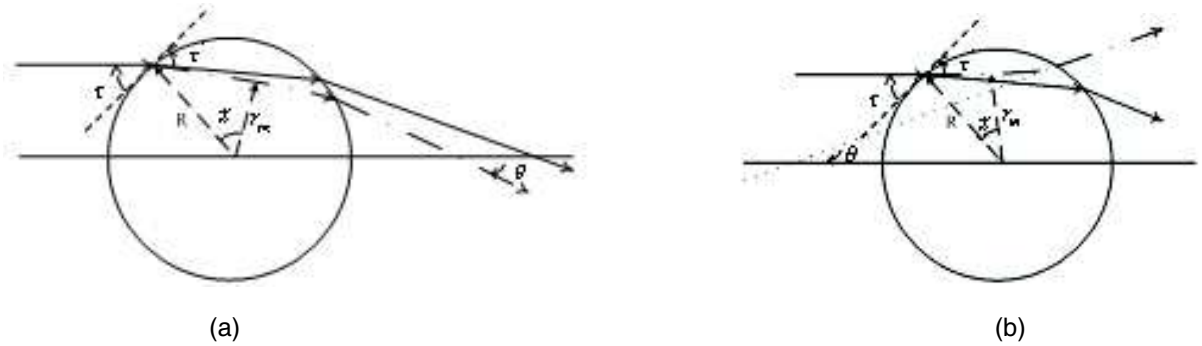


Figure 1: Comparison of the light path inside a sphere with a uniform refractive index (solid curves) and the light path inside a sphere with a spherically symmetric refractive index $m(\widetilde{r})$ (dashed curves). (a) $m_0 > m_1$, (b) $m_0 < m_1$.

not be described by the Descartes law (Eq. (2)). the differential equation for the light rays can be written in a nondimensional form [9]:

$$\chi(r) = \chi_0 + \int_{rc}^{r} \frac{e}{r\sqrt{m(\widetilde{r})^2\widetilde{r}^2 - e^2}} dr \tag{3}$$

Where $e = m_s \cos(\tau)$ is a constant, χ_0 equals zero under ordinary conditions, r is a position of one point of a light ray, and $\widetilde{r} = r/R$ is the nondimensional sphere radius. In 2005 Maria Rosaria Vetrano found geometrically the relationship between incident and scattering angles for a spherically symmetric particle:

$$\theta = 2\tau - 2(N-1)[\chi(r=R) - \chi(r=rc)] \tag{4}$$

For a spherically symmetric particle with radius R, the dimensionless parameter is set: $\alpha = 2\pi R/\lambda$, where λ is the wavelength of incident light, and rc represents the minimal distance of approach of the light ray from the center of the sphere. Combined with the relationship (Eq. (4)), the phase shift and the scattering intensity of a spherically symmetric particle can be derived.

The dimensionless scattering intensity is defined by [10]

$$i_{1,2} = \alpha^2 k_{1,2}^2 D \tag{5}$$

Where $k_{1,2}$ is introduced to characterize the fraction of the incident intensity contained in emerging rays

$$\begin{array}{ll} k_{1,2} = \varepsilon_{1,2} & \text{for } N = 1 \\ k_{1,2} = (1 - \varepsilon_{1,2}^2)(-\varepsilon_{1,2})^{N-2} & \text{for} N = 2, 3, 4 \ldots \end{array} \tag{6}$$

D is usually called divergence or gain, denoting the influence of particle shape on angular intensity distribution. It can be expressed as

$$D = \frac{\sin\tau\cos\tau}{\sin\theta\,|d\theta/d\tau|} \tag{7}$$

Where $\varepsilon_{1,2}$ are Fresnel reflection coefficients, defined as

$$\begin{aligned} \varepsilon_1 &= \frac{\sin\tau - m_1\sin\tau'}{\sin\tau + m_1\sin\tau'} \\ \varepsilon_2 &= \frac{m_1\sin\tau - \sin\tau'}{m_1\sin\tau + \sin\tau'} \end{aligned} \tag{8}$$

Phase shifts [10] $\psi_{1,2}$, which are due to reflection and correspond to perpendicular and parallel polarized components, are expressed by the Fresnel reflection coefficients. Xu and Cai [6] used a one-dimensional piecewise cubic spline interpolation to get the amplitudes and phases of the rays. However, this method may engender deviation. In this paper, the same scattering angles should be applied to calculate the amplitudes and phases of the rays. Hence, the incident angle of the reflected light should be derived from the scattering angle.

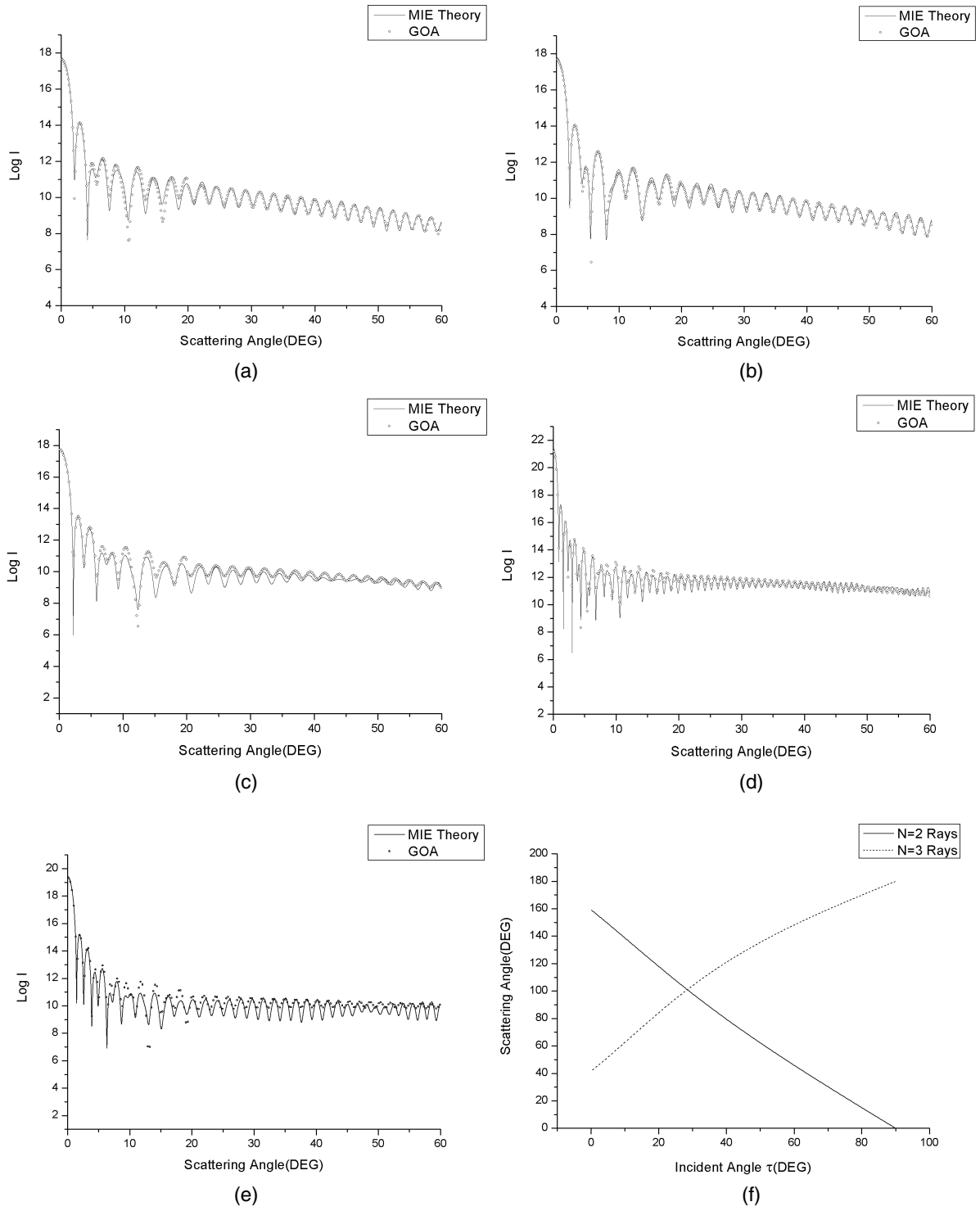


Figure 2: Comparison of intensity of calculation by the MIE theory and by the GOA method. (a) R = 10 um, m0 = 1.49, m1 = 1.47. (b) R = 10 μm, m0 = 1.47, m1 = 1.49. (c) R = 10 μm, m0 = 1.5, m1 = 1.1. (d) R = 25 μm, m0 = 1.5, m1 = 1.1. (e) R = 10 μm, m0 = 2.1, m1 = 1.1. (f) R = 10 μm, m0 = 2.1, m1 = 1.1.

Following Van de Hulst's discussion of the phase change that is due to optical length and focal lines, the combined phases $\sigma_{reflection1,2}$ and $\sigma_{refraction1,2}$ in the foregoing cases can finally be obtained:

$$\begin{aligned} \sigma_{reflection1,2} &= \pi/2 + 2\alpha \sin(\tau_{reflection}) + \psi_{1,2} \\ \sigma_{refraction1,2} &= 3\pi/2 + 2\alpha \sin(\tau) - \frac{2\pi L}{\lambda} \end{aligned} \tag{9}$$

Where L is the optical path length of light and $\tau_{reflection} = \theta/2$ is the incident angle of the reflected light.

With the scattering angles and the combined phases, amplitude functions for the two polarizations can be written as [10]:

$$s_{1,2} = \sqrt{i_{1,2}} \exp(I\sigma_{1,2}) \tag{10}$$

The final amplitude functions can be finally obtained [6]:

$$\begin{array}{ll} s = s_{reflection} + s_{refraction} + s_{diffraction} & \theta \in [0^\circ,\ 20^\circ] \\ s = s_{reflection} + s_{refraction} & \theta \in [20^\circ,\ 60^\circ] \end{array} \tag{11}$$

Where $s_{diffraction} = \alpha^2 J_1(\alpha \sin\theta)/(\alpha \sin\theta)$ and J_1 is the first-order Bessel function.

Assuming that the particle is illuminated by an unpolarized monochromatic beam with wavelength λ and incident intensity I_0, the scattering intensity $I(\theta)$ at the distance f can finally be obtained by

$$I(\theta) = \frac{\lambda^2 I_0}{8\pi^2 f^2}(|s_1(\theta)|^2 + |s_2(\theta)|^2) \tag{12}$$

In this paper the index $\lambda^2 I_0/(8\pi^2 f^2)$ is defined as 1and the wavelength of light λ equals 1.0.

3. COMPARISON OF THE GEOMETRICAL-OPTICS APPROXIMATION WITH MIE THEORY

Refractive index is defined as $m(\widetilde{r}) = m_0(1 - \frac{m_0^2 - m_1^2}{m_0^2}\widetilde{r}^2)^{\frac{1}{2}}$. $\chi(r)$ can be denoted as analytical solution by use this refractive index. Five types of transparent particles were calculated by the GOA method. The results were compared with those obtained by use of the MIE theory (Fig. 2). Besides, for the difference $\Delta m = |m_0 - m_1| > 1$, a portion of $N = 2$ rays is beyond $60^\circ \sim 180^\circ$ and a portion of $N = 3$ rays is beyond $40^\circ \sim 60^\circ$ (Fig. 2(f)). The intensities calculated by the GOA on this condition may have remarkable approximation errors(Fig. 2(e)).

With the radius that varied from $15\,\mu m \sim 1000\,\mu m$, the refractive indices of the center and the surface were 1.33 and 1.30, respectively, numerical calculations were performed on an Intel Pentium 3.0 GHz PC with 1.0 G RAM to test the calculation efficiency of the two methods. As illustrated in Fig. 3, the calculation time of the rigorous MIE theory increased rapidly with increasing radius. With case B, the calculation time of the GOA theory varied from 2.4 ms to 120 ms.

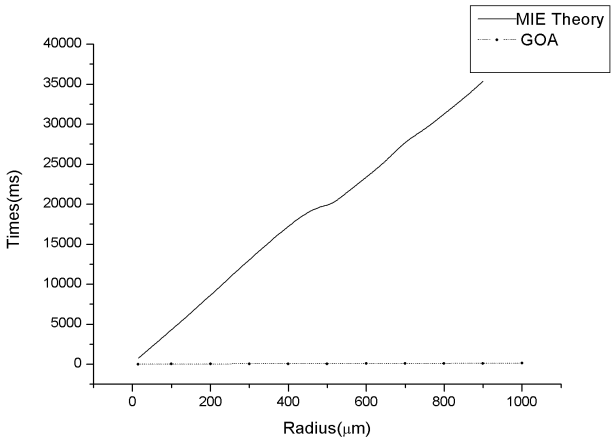


Figure 3: Comparison of speed of calculation by the MIE theory and by the GOA method.

4. CONCLUSIONS

In this paper we have treated light scattering as the superposition of diffraction, reflection, and refraction. An approximation algorithm — the GOA method — was developed to calculate the scattering light pattern within the forward direction (0°–60°) for non homogeneous particles. With the typical continuous refractive index model, the calculation time of the GOA theory are compared with MIE theory and CPU time is greatly decreased. But it fails to describe light scattering at the scattering angles from which the refractive rays are absent . Neither can it work well at large scattering angles form which a portion of $N = 3$ rays is beyond $30^\circ \sim 60^\circ$.

REFERENCES

1. Gomez-Reino, C., M. V. Perez, and C. Bao, *Gradient-Index Optics Fundamentals and Applications*, Springer, New York, 2002.
2. Xu, F., X. Cai, and K. Ren, "Geometrical-optics approximation of forward scattering by coated particles," *Appl. Opt.*, Vol. 43, 1870–1879, 2004.
3. Belafhal, A., M. Ibnchaikh, and K. Nassim, "Scattering amplitude of absorbing and nonabsorbing spheroidal particles in the WKB approximation," *J. Quant. Spectrosc. Radiat. Transfer*, Vol. 72, 385–402, 2002.
4. Min, M., J. W. Hovenier, and A. de Koter, "Scattering and absorption cross sections for randomly oriented spheroids of arbitrary size," *J. Quant. Spectrosc. Radiat. Transfer*, 79–80, 939–951, 2003.
5. Grynko, Y. and Y. Shkuratov, "Scattering matrix calculated in geometric optics approximation for semitransparent particles faceted with various shapes," *J. Quant. Spectrosc. Radiat. Transfer*, Vol. 78, 319–340, 2003.
6. Xu, F., X. S. Cai, and J. Shen, "Geometric approximation of light scattering in arbitrary diffraction regime for absorbing particles: application in laser particle sizing," *Acta Opt. Sin.*, Vol. 23, 1464–1469, 2003.
7. Xu, F., K. F. Ren, and X. S. Cai, "Extension of geometrical-optics approximation to on-axis Gaussian beam scattering," *Appl.Opt.*, Vol. 45, No. 20, 4990–4999, 2006.
8. Descartes, R., *Discourse on Method, Optics, Geometry, and Meteorology*, revised ed., translated by P. J. Olscamp, Hackett, Indianapolis, Ind., 2001.
9. Vetrano, M. R., J. P. A. Johannes van Beeck, and M. L. Riethmuller, "Generalization of the rainbow Airy theory to nonuniform spheres," *Appl. Opt.*, Vol. 30, 658–660, 2005.
10. Van de Hulst, H. C., *Light Scattering by Small Particles*, Dover, New York, 1981.

An optimized Scheme for Optical Phased Array Beam Steering Controlled by Wavelength

Ying Zhao, **Xiaozhou Yang**, **Qin Cai**, **and Weiwei Hu**

National Laboratory on Local Fiber-Optic Comm. Networks and Advanced Optical Comm. Systems
Peking University, Beijing 100871, China

Abstract— An efficiency-optimized scheme for Optical phased array (OPA) beam steering controlled by wavelength is proposed. In this scheme, an OPA is designed based on optical waveguides with specific interelement center-to-center spacing and length difference, by which the number of elements (phase shifters) that result in high Signal-to-Noise ratio (SNR) can be dramatically decreased. To suppress the sidelobes, we worked out a method based on gradient techniques to optimize the interelement spacing and length difference. And we mathematically prove the SNR in the case of optimized scheme is unchanged when beam steering controlled by wavelength.

1. INTRODUCTION

Optical phased arrays, which have been scaled down for use at optical wavelengths, are direct functional analogs of microwave phased-array antennas and share the same fundamental beam-forming concepts as their microwave precursors. OPA provide an elegant means for the inertialess, high-resolution random-access beam steering that is required by numerous applications, including laser radar and laser communication. Before us, various kinds of phase shifters have been studied and tested, such as $LiNbO_3$ [1], AlGaAs [2, 3], liquid crystals [4], PLZT [5] and etc. However, some difficulties impede the technology of OPA beamsteering to high performance and practical application, one of which is the coupling problem because the interelement spacing of phased array should be sufficiently close ($d \leq \lambda/2$ is needed for an efficient phased array, where d is the interelement spacing of the array and λ is the optical wavelength) to eliminate sidelobes in the far-field radiation pattern.

We proposed an idea [6] that the phased-array optical deflectors are randomly arranged with nonuniform spacing, which can solve this difficulty of $d \leq \lambda/2$ in the optical domain. Moreover, we have proved feasibility of using integrated optical waveguides [6] and photonic crystals [7] as the phase shifters of OPA.

However, this idea has a disadvantage that it can effectively suppress the sidelobes only when the number of the elements is quite large, (To have a suppression ratio of 10 dB, it needs average 100 to 200 elements at the range of spacing from 20 µm to 40 µm) and it is difficulty to fabricate and, more importantly, it may not meet the condition of temporal coherence. In this paper, we worked out a computing method to optimize the interelement spacing and length difference to curtail the number of elements used in the array for practical application and to ensure the possibility of temporal coherence.

This paper is organized as follows. In Section 2, we give the optimized result which is a group of numbers relevant to the interelement spacing and length difference. In Section 3, we mathematically prove the Signal-to-Noise ratio (SNR) is constant with change of radiation angle. And we demonstrate the method of optimization we used.

2. OPTIMIZED RESULT

In the following, firstly, we review the idea of nonuniform spaced array which necessarily establish the relation between the interelement spacing d_i and the length difference Δl_i ($\Delta l_i = l_{i+1} - l_i$) between the adjacent elements $i+1$ and i, and numerically calculate radiative angle of the phased array against optical wavelength.

In this theoretical model shown in Fig. 1, the distances between elements are not equal, i. e., $d_1 \neq d_2 \neq ... \neq d_{N-1}$, N is number of elements. And, the length differences between adjacent elements are nonuniform. A tunable laser source is used in the scheme, in which the wavelength λ varying from $\lambda_0 - \Delta\lambda_{\max}$ to $\lambda_0 + \Delta\lambda_{\max}$ and $d_i \gg \lambda$ are assumed. For a one-dimensional array of emitters and under the Fraunhofer approximation, we can get the maximum intense beam, the

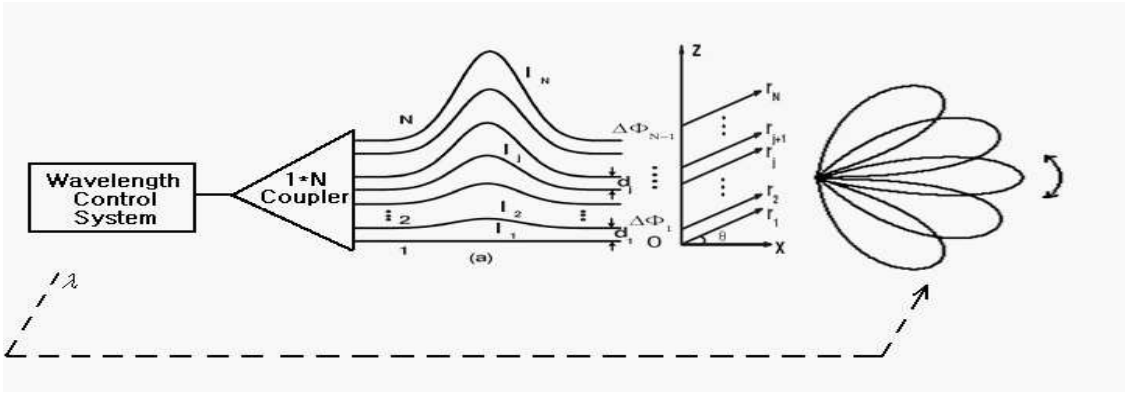


Figure 1: A scheme for nonuniform spaced array fabricated by optical waveguide.

so-called main lobe, which results from the constructive interference, at the direction [6]

$$\theta = \arcsin \frac{k_i \cdot \Delta\lambda}{d_i} \tag{1}$$

We define $k_i \cdot \Delta\lambda_{\max} = d_i$, and $n \cdot \Delta l_i = k_i \cdot \lambda_0 + d_i \sin 0^\circ = k_i \cdot \lambda_0$. Where k_i is an integer which determines the interelement spacing d_i and length difference Δl_i, n is the effective refractive index of the waveguide, and $i = 1, 2, \ldots, N-1$. Then $\theta = \arcsin(\Delta\lambda/\Delta\lambda_{\max})$. And in directions other than θ, the combined intensity is relatively low.

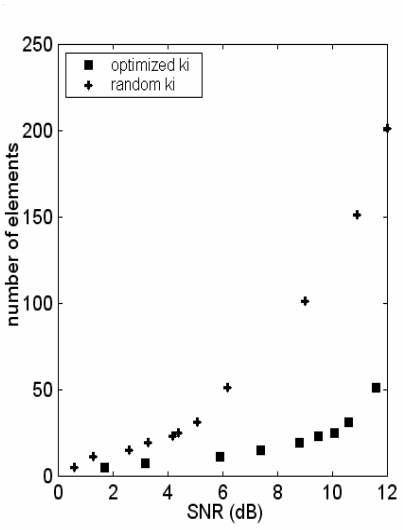


Figure 2: Comparison of number of elements of optimized k_i with that of random k_i. The element spacing ranges from 17.040 um to 66.704 um. And the direction of optical beam is controlled by wavelength from 1.1 rad to -1.1 rad..

Now we demonstrate that when k_i is a special group of numbers, the sidelobes will be dramatically suppressed further. By an optimized method based on the gradient techniques, we get some groups of optimized numbers, k_i, which remarkably decrease the number of elements used in array at the same Signal-to-Noise ratio (SNR) in contrast to the condition that k_i are random numbers.

The SNR is defined as $SNR = \dfrac{\text{maximal sidelobe power}}{\text{mainlobe power}}$.

The above results show through optimization, the SNR, which is approximately 10 dB for $N = 25$, can be effectively increased compared with the result that the SNR is approximately 10 dB for $N = 120$ when k_i is random number. To achieve the same ratio of suppression of sidelobes, the number of elements we need with optimized k_i is 1/3 to 1/4 of that with random k_i. This result has great significance because with relatively fewer elements used, the SNR can also achieve the level

we expect. So it provides an opportunity for fabrication for practical use with integrated optical waveguide and more importantly, fewer elements mean smaller scale of OPA, which is demanded to meet the condition of temporal coherence because finite wave-trains may not be coherent in an OPA of large scale.

Table 1: Optimized k_i of different number of elements.

Elements number	**Optimized** k_i									
5	3645	3848	2910	4169						
7	4041	3702	3441	2373	3841	3364				
11	3471	3676	3293	3772	2998	3176	2716	3226	3087	3395
15	2285	2306	2377	2921	2828	3083	3012	2679	2369	2508
	2659	2586	3230	2578						
19	1570	2449	1889	1795	1684	1775	1838	2011	1519	1428
	1908	2309	2389	1719	1590	2272	1637	2110		
23	1951	1707	1296	1896	1146	1950	2043	1250	1889	1478
	1775	2043	1558	1999	1874	1065	1304	1103	1371	1964
	1867	1237								
25	2062	1567	1275	1912	1432	1141	1542	1788	1404	1726
	1752	1380	1301	1734	2109	1526	1882	1993	2077	2178
	1691	1616	1212	1600						
31	2134	2031	1910	1513	1769	2267	1864	1862	2407	1588
	1879	1604	1373	1467	2126	1974	1704	2269	2403	2435
	2175	2212	2243	2222	1667	1740	1425	2475	1725	2061

The optimized k_i are given from the left form. The above results are derived regardless of the pattern of aperiodic spacing. And $\lambda_0 = 1.55\,\mu\text{m}$, $\Delta\lambda_{\max} = 0.016\,\mu\text{m}$, with the difference of optical path length $n \cdot \Delta l_i$ ranging from $1650.75\,\mu\text{m}$ to $6461.95\,\mu\text{m}$ and the element spacing ranging from $17.040\,\mu\text{m}$ to $66.704\,\mu\text{m}$.

In addition, when using these optimized k_i in experiments, the fabrication error has an influence on the SNR. If we hypothetically add all independent random fabrication errors with the same level to the element spacing respectively, we can see with increase of the error level, the standard variance of SNR enlarge significantly. When the error level is 200 nm, the SNR decrease by approximately 2 dB and the standard variance enlarge to about 1 dB. Even though the SNR diminish and the standard variances enlarge with the increase of error, this result is much better than that when k_i is random number [6] because with the random error increasing, the length difference and interelement spacing tend to become random gradually.

3. MATHEMATICAL THEORY ABOUT THIS MODEL

Next we mathematically prove the SNR in the case of optimized k_i is unchanged when optical wavelength varies. For a one-dimensional array of emitters, under the Fraunhofer approximation the radiation is given by [7]

$$E(\theta) = E_1 \cdot \left\{ 1 + \sum_{i=1}^{N-1} \exp\left[-j\frac{2\pi}{\lambda}(n\sum_{j=1}^{i}\Delta L_j - \sum_{j=1}^{i} d_j \cdot \sin\theta) \right] \right\}$$

where $E_1 = \exp\left[-j\frac{2\pi}{\lambda}(nL_1 + r_1)\right]$, $d_j = -k_j\Delta\lambda_{\max}$, and $\Delta L_j = \frac{k_j \cdot \lambda_0}{n}$, so that, the intensity

$$I(\theta) = I_1 \cdot \left|1 + \sum_{i=1}^{n-1} \exp\left[-j\frac{2\pi}{\lambda}\sum_{j=1}^{i} k_i\left(\lambda_0 - \sin\theta \cdot \Delta\lambda_{\max}\right)\right]\right|^2 \tag{2}$$

where $I_1 = E_1^2$. Given $\Delta\lambda = \lambda - \lambda_0$ changing from $\Delta\lambda_1$ to $\Delta\lambda_2$, we can find

$$\theta_1|_{\Delta\lambda_1} = f(\theta_2|_{\Delta\lambda_2}) = \arcsin\left[\frac{\lambda_0 \cdot (\Delta\lambda_1 - \Delta\lambda_2) + \Delta\lambda_{\max} \cdot \sin\theta_2 \cdot (\lambda_0 - \Delta\lambda_1)}{(\lambda_0 - \Delta\lambda_2) \cdot \Delta\lambda_{\max}}\right] \tag{3}$$

s.t. $I(\theta_1)|_{\Delta\lambda_1} = I(\theta_2)|_{\Delta\lambda_2}$. In order to prove SNR is unchanged, we need to prove the local maximum of sidelobes in these two cases ($\Delta\lambda = \Delta\lambda_1$, $\Delta\lambda = \Delta\lambda_2$) are equal and corresponding. Given θ_1 is the local maximal point of $I(\theta)|_{\Delta\lambda_1}$, from (4), we should prove the corresponding θ_2 also is the local maximal point of $I(\theta)|_{\Delta\lambda_2}$. In another word, we should prove $\exists\varepsilon > 0$, $\forall\theta_2 - \varepsilon < \theta < \theta_2 + \varepsilon$, s.t. $I(\theta)|_{\Delta\lambda_2} \le I(\theta_2)|_{\Delta\lambda_2}$.

Because θ_1 is the local maximal point of $I(\theta)|_{\Delta\lambda_1}$, So that $\exists\varepsilon_1 > 0$, $\forall\theta_1 - \varepsilon_1 < \theta_1^* < \theta_1 + \varepsilon_1$, s.t. $I(\theta_1^*)|_{\Delta\lambda_1} \le I(\theta_1)|_{\Delta\lambda_1}$. From (3), we have

$$\sin\theta_1^* = \frac{\lambda_0 \cdot (\Delta\lambda_1 - \Delta\lambda_2) + \Delta\lambda_{\max} \cdot \sin\theta_2^* \cdot (\lambda_0 - \Delta\lambda_1)}{(\lambda_0 - \Delta\lambda_2) \cdot \Delta\lambda\max}$$

written as $\theta_1^* = f(\theta_2^*)$, so that $\exists\varepsilon_2 > 0$, $\forall\theta_2 - \varepsilon_2 < \theta_2^* < \theta_2 + \varepsilon_2$, s.t. $\theta_1 - \varepsilon_1 < f(\theta_2^*) < \theta_1 + \varepsilon_1$. $I(f(\theta_2^*))|_{\Delta\lambda_1} = I(\theta_2^*)|_{\Delta\lambda_2}$, and $\theta_1 - \varepsilon_1 < f(\theta_2^*) < \theta_1 + \varepsilon_1$, so that $I(\theta_2^*)|_{\Delta\lambda_2} = I(f(\theta_2^*))|_{\Delta\lambda_1} \le I(\theta_1)|_{\Delta\lambda_1}$, and because $I(\theta_1)|_{\Delta\lambda_1} = I(\theta_2)|_{\Delta\lambda_2}$, so $\exists\varepsilon_2 > 0$, $\forall\theta_2 - \varepsilon_2 < \theta_2^* < \theta_2 = +\varepsilon_2$, $I(\theta_2^*)|_{\Delta\lambda_2} \le I(\theta_2)|_{\Delta\lambda_2}$.

So, θ_2 also is the local maximum of $I(\theta)|_{\Delta\lambda_2}$. Similarly, when $\Delta\lambda$ changes from $\Delta\lambda_2$ to $\Delta\lambda_1$, all the local maximum of $I(\theta)|_{\Delta\lambda_2}$ can find the corresponding local maximum of $I(\theta)|_{\Delta\lambda_1}$. So, every local maximum can find its equally corresponding local maximum when wavelength changes which means it is impossible to appear new local maximum with change of λ. So, the SNR is unchanged when λ varies.

In the following, we demonstrate the method of optimization we used. We try to find a group of k_i that can make the SNR achieve the maximum, but there is a difficulty that the SNR cannot be written as a mathematical analytic form because SNR(k_i) is not a conservative function considering varies of the position of sidelobes. So although there are many numerical methods such as The Fletcher-Powell Optimization Technique and Least-pth Optimization Technique [8], they are not useful in this scheme because these methods need an analytic form of the function of SNR. We worked out a numerical method based on the Gradient Technique that does not need the SNR to be an analytic form.

In this mathematical method, we assume that the SNR is a function which changes with the variation of the N parameters $\delta k_1, \delta k_2, \ldots, \delta k_{N-1}$, where N is the number of the elements. When the changes of these N parameters are sufficiently small, then follows that

$$S/N(\boldsymbol{k} + \delta\boldsymbol{k}) \approx S/N(\boldsymbol{k}) + (\nabla S/N)^T\delta\boldsymbol{k} \tag{4}$$

where $\boldsymbol{k}$ is a vector having components $k_1, k_2, ..., k_{N-1}$, and $S/N(\cdot)$ is the function of SNR. The gradient is a vector which can be written as $\nabla S/N = (\frac{\partial S/N}{\partial k_1}, \frac{\partial S/N}{\partial k_2}, \ldots, \frac{\partial S/N}{\partial k_{N-1}})$.

Since $\delta\boldsymbol{k}$ is a vector, it can be described by two quantities: direction and magnitude. In order to find a relative maximum of SNR we will use the $\nabla S/N$'s information to choose $\delta\boldsymbol{k}$ to point in the direction of positive gradient. (Because the SNR cannot be written as a mathematical analytic form, we use $\left(\frac{S/N(k_1+\delta k_1)-S/N(k_1)}{\delta k_1}, \frac{S/N(k_2+\delta k_2)-S/N(k_2)}{\delta k_2}, \ldots, \frac{S/N(k_{N-1}+\delta k_{N-1})-S/N(k_{N-1})}{\delta k_{N-1}}\right)$ to approximately estimate $\nabla S/N$. In the optimized method, we first choose an initial set of $\boldsymbol{k}$, which results in a value of SNR. Next the gradient $\nabla S/N$ is calculated at point $\boldsymbol{k}$. The parameter changes are then chosen to be in the direction of positive gradient; that is $\delta\boldsymbol{k} = \alpha\nabla S/N$. Where α is a

positive constant that can be adjusted according to different accuracy. Then the new parameters can be calculated from the old parameters:

$$\boldsymbol{k} + \alpha \nabla S/N \rightarrow \boldsymbol{k} \tag{5}$$

By doing these till the $\nabla S/N$ approximately equal to zero, we can find a relative maximum of the SNR.

However, often many different initial parameters produce numerous local maximum. Choosing the initial parameters can get different path and achieve different local maximum. In this case, the value of SNR we get may be a local maximum and we can only choose as much as possible of paths to further increase the SNR. In computing, we choose many groups of initial k_i randomly and increase the SNR as much as possible.

4. CONCLUSIONS

In conclusion, an optimized method of OPA made by waveguides and controlled by wavelength was schemed to reduce the number of elements in the system. By doing this, we can ensure the condition of temporal coherence can be met and during the practical use, it can reduce the difficulty of fabrication. We theoretically proved that the SNR remains constant when the optical wavelength changes and we demonstrated a method based on gradient technique to work out the optimized k_i when the function of SNR cannot be written as an analytic form. These results show that a wide-angle optical beamsteering with high SNR and high velocity can be obtained by a relative simple device. This scheme will pave a way for the practical reality of OPA.

ACKNOWLEDGMENT

The authors thank the National Natural Science Foundation of China (NSFC) for the support (No. 60477002).

REFERENCES

1. Ninomiya, Y., "Ultrahigh resolving electrooptic prism array light deflectors," *IEEE J. Quant. Electron*, 791–795, 1973.
2. Wight, D. R., et al., "Novel phased array optical scanning device implemented using GaAs/AlGaAs technology," *Appl. Phys. Lett.*, Vol. 59, 899–901 ,1991.
3. Vasey, F., et al., "Spatial optical beam steering with an AlGaAs integrated phased array," *Appl. Opt.*, Vol. 32, 3220–3232, 1993.
4. Mcmanamon, P. F., et al., "Optical phased array technology," *Proc. IEEE*, Vol. 84, 268–298, 1996.
5. Talbot, P. J., et al., "Design and simulation of PLZT-based eletro-optic phased array scanners," *Opt. Memory Neural Net.*, Vol. 3, 111, 1994.
6. Xiao, F., et al., "Optical phased array beam steering controlled by wavelength," *Applied Optics*, Vol. 44, No. 26, 5429–5433, 2005.
7. Wang, X., et al., "Wavelength controlled optical phased array using photonic crystal," *Proceedings of 20th Congress of the International Commission for Optics (ICO)*, 148, Changchun, China, August 2005.
8. Daniels, R. W., *An Introduction to Numerical Methods and Optimization Techniques*, North-Holland, New York, 1978.

A Structure Design for the Expansion of Fiber Element Radiation Scope in Optical Phased Array

Yanyun Yang, **Cheng Hong**, **Weiwei Hu**, **and Anshi Xu**

National Laboratory on Local Fiber-Optic Comm. Networks and Advanced Optical Comm. Systems
Peking University, Beijing 100871, China

Abstract— An optical waveguide structure, which increases the far field radiation angle two times than that of conventional fiber, is designed and analyzed for the research of optical phased array. The structure improves the scanning scope of the optical fiber phased array effectively through designing the shape and the refractive index profile of the single mode fiber. Far field characteristic of this structure is analyzed using FD-BPM and the Fraunhofer diffraction theory.

1. INTRODUCTION

Optical phased arrays are direct functional analogs of the well known microwave phased array antennas that make possible the agile, inertialess steering of microwave beams. There are many application areas that can benefit from the performance of optical phased arrays, such as laser radar and Laser communications.

In the previous research of our laboratory, a wavelength controlled optical phased array has been designed [1], in which different lengths L_i of fiber waveguides result in different phases Φ_i at the end of the array and the distances d_i between two adjacent elements are not equal. Figure 1 shows the scheme of wavelength controlled optical phased array technique.

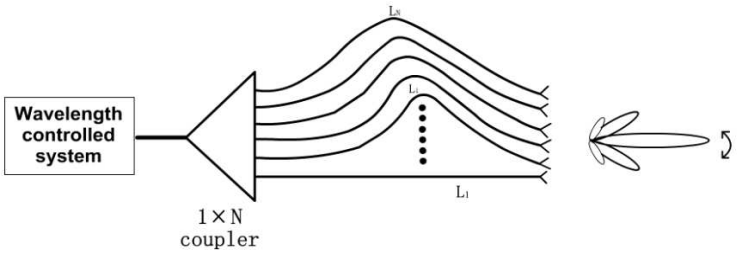


Figure 1: A scheme for wavelength controlled optical phased array.

In the optical fiber phased array, the far field intensity can be written in the form $I(\theta) = S(\theta)F(\theta)$, where $S(\theta)$ is a structure factor determined by the array configuration, and $F(\theta)$ is a factor determined by the properties of single fiber diffractive function. The scanning scope of the optical phased array is determined by $F(\theta)$ to a great extent. Therefore, to expand the scanning scope of the array, the far field divergent angle for single fiber has to be increased. In this article, we present a special waveguide structure which widens the far field radiation angle effectively, through tapering standard single mode fiber and increasing the refractive index difference simultaneously.

In Section 2 numerical method FD-BPM is outlined and the diffraction theory for far field calculation is presented. The novel tapered single-mode fiber is numerically and theoretically analyzed and the simulation results are presented in the next section. Finally the general conclusions are drawn.

2. FD-BPM AND FAR FIELD CALCULATION

There have been already many methods to analyze tapered fiber such as the finite difference time-domain method (FDTD), beam propagation method (BPM) and coupled-mode theory. FDTD is a numerical method [2], which results in an accurate analysis at the expanse of high consuming of memory of the computer and extensive computational effort. Coupled-mode theory can also be used to describe the propagation characteristics of tapered single-mode fiber, available in the literature [3], however it is inconvenient to calculate the far field radiation angle. BPM has been successfully used to analyze a wide spectrum of guided-wave structures [4], and the accuracy and applicability have been studied extensively.

Here we solve the BPM paraxial wave equation by involving numerical technique finite difference approximation (FD-BPM) to analyze the propagation characteristics in the new waveguide structure.

In the presence of the one-dimensional cross-sectional index profile $n(x,z)$ and the paraxial limit, with the assumption of $\left|\frac{\partial^2 u}{\partial z^2}\right| \ll \left|2\bar{k}\frac{\partial u}{\partial z}\right|$, Helmholtz equation can be reduced to the following paraxial wave equation:

$$\frac{\partial u}{\partial z} = \frac{j}{2\bar{k}}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \left(k^2 - \bar{k}^2\right)\right) u \tag{1}$$

which is the basic BPM equation. Here $k = k_0 n(x,y,z)$, $\bar{k}$ is a constant number to represent the average phase variation of the field ϕ and is referred to as the reference wave number $\bar{k} = k_0\bar{n}$.

Involving the finite difference method to Equation (1), the numerically discrete equation can be obtained:

$$\frac{u_i^{n+1} - u_i^n}{\Delta z} = \frac{j}{2\bar{k}}\left[\frac{\delta^2}{\Delta x^2} + \left(k(x_i, z_{n+1/2})^2 - \bar{k}^2\right)\right]\frac{u_i^{n+1} + u_i^n}{2} \tag{2}$$

Here u_i^n denote the field at transverse grid point i and longitudinal plane n, and $\delta^2 u_i = u_{i+1} + u_{i-1} - 2u_i$, $z_{n+1/2} = z_n + \Delta z/2$.

The transparent boundary condition (TBC) [5] has been used in the FD-BPM computation.

With the electronic field data $\psi(r)$ at the end face of the new structure from FD-BPM calculation, the far field can be obtained through Fraunhofer diffraction theory:

$$E(\theta) = 2\pi \int_0^\infty \psi(r) J_0(kr\sin\theta) r dr \tag{3}$$

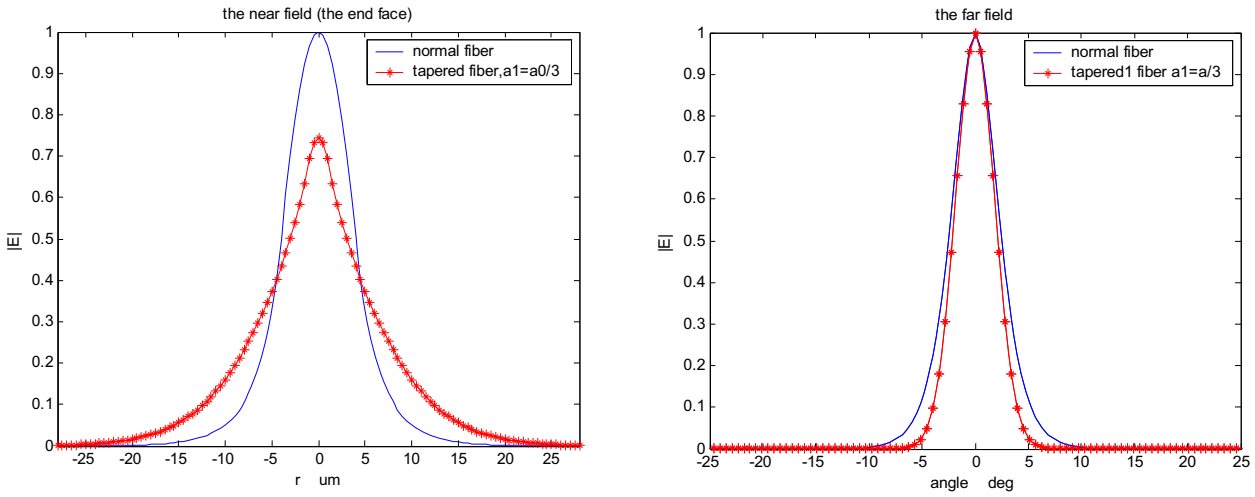
3. SIMULATION RESULTS

The light energy in cladding P_{clad} is inversely proportional to the normalized frequency V, $\frac{P_{clad}}{P} = \frac{4\sqrt{2}}{3V}$, that is to say, the energy in the core is proportional to V. More energy converges in the core with V as large as possible under the single mode condition.

According to the Fraunhofer diffraction theory, the spot for diffraction has to be smaller to acquire a larger far field radiation angle than the standard single-mode fiber. The special structure must have a smaller mode field diameter (MFD) or a smaller radius.

To get a smaller radius, if we just taper the standard single-mode fiber only, the normalized frequency V will be reduced. Along with the mode coupling, more power is transferred to the cladding. Hence, at the end face the MFD becomes larger than that of the standard single-mode fiber on the contrary and far field divergent angle gets smaller. Figures 2(a) and (b) illustrate the electronic field at the near end face and the far field separately.

To get smaller MFD, more power should be restrained in the core with the radius diminishing.



(a) (b)

Figure 2: (a) $|E|$ at the end face, (b) $|E|$ at the far field.

Since the power energy in the core is proportional to V, the fact that V should be as large as possible has to be warranted under the single mode condition. As it is known $V = a\frac{2\pi}{\lambda}\sqrt{n_1^2 - n_2^2} = a\frac{2\pi}{\lambda}n_1\sqrt{2\Delta}$, and the background index n_1 stays a constant 1.465 normally. Therefore, to keep V as a constant, Δ has to be increased as the radius diminishes. Figure 3 illustrates the typical structure we designed (the radius decreases linearly from a_1 to a_0 and Δ increases quadratically in the tapering area from Δ to Δ_1). Here a_1 is the diameter at the end face, a_0 is that of the conventional single mode fiber, L_2 is the length of the tapering area and Δ_1 is the index difference at the end face.

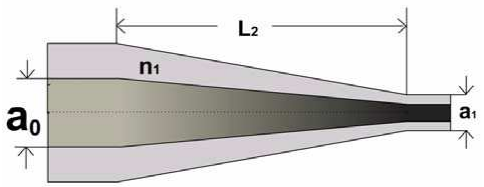


Figure 3: The fiber core of the designed structure.

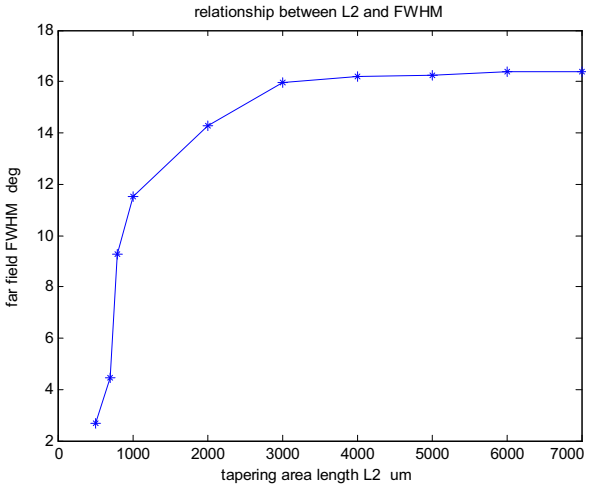
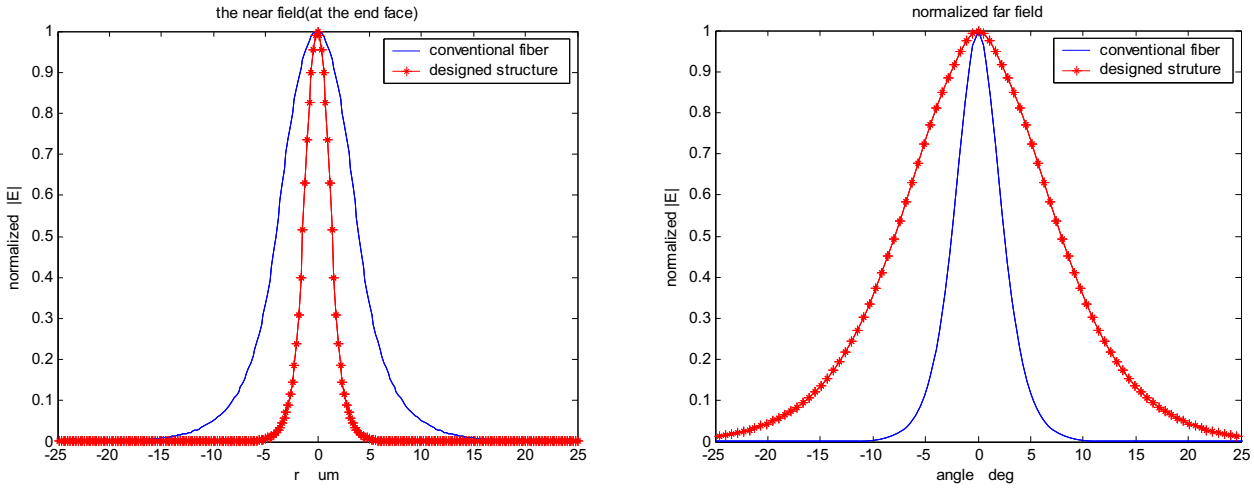


Figure 4: Influence of L2 on the far field FWHM.

The length of tapering area L2 contributes greatly to the far field intensity Full Width Half Maximum (FWHM). As it is showed in Figure 4, the far field FWHM increases while L2 grows longer and maintains almost stable after L2 reaching 3000 µm. With a_1 fixed at 3 µm, shorter L2 means swift change in the shape, which results in energy transfer from fiber core to the cladding. The mode coupling would lead to expansion of MFD, which is responsible for the less radiation angle and far field FWHM. In the computational simulation, we designed $L_2 = 3700\,\mu m$, $a1 = 3\,\mu m$, $a0 = 9\,\mu m$ and $\Delta_1 = 9 * \Delta$. There is little variation of the far field FWHM among different tapering area shapes linear, quadratic and exponential. In this article, the radius decreases linearly in view of the convenience of manufacturing techniques.

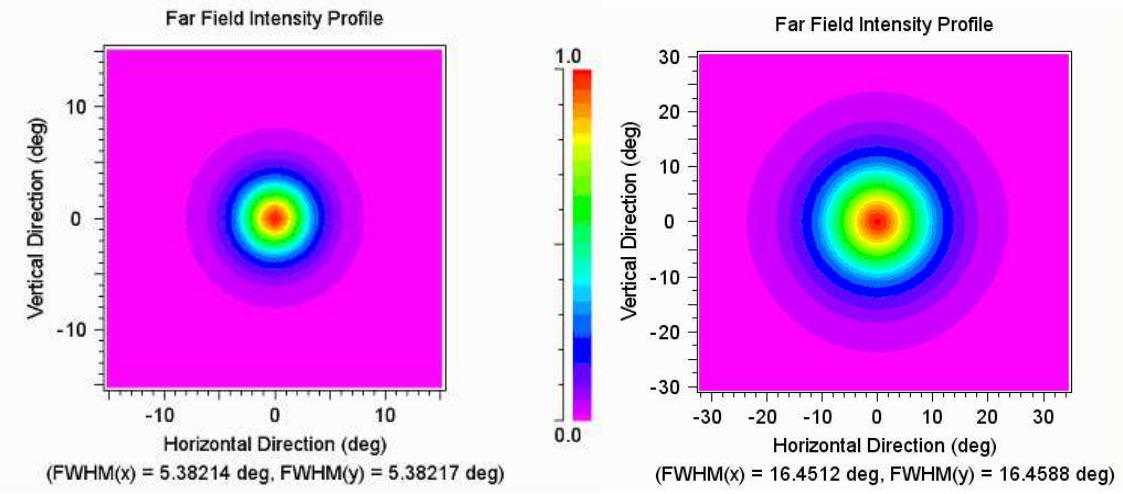
Figures 5(a) and (b) illustrate the sharp contrast of the near and far field radiation angle between the new waveguide structure and standard single mode fiber respectively. This is a two-dimensional intensity profile, since the fiber is symmetrical. From the figure, it is clear that this novel structure



(a) (b)

Figure 5: Two-dimensional normalized $|E|$ in (a) the near field and (b) far field of the designed structure.

has a much smaller MFD and larger radiation angle. ($\Delta x = 0.1\,\mu$m, $\Delta y = 0.1\,\mu$m $\Delta z = 1\,\mu$m, in the FD-BPM simulation)



(a) (b)

Figure 6: Far field intensity profile. (a) The standard single mode fiber, (b) The designed structure.

Figures 6(a) and (b) illustrate the far field intensity profile of the standard single mode fiber and this new structure respectively. It is clear that the Full Width Half Maximum (FWHM) of the conventional fiber is just 5.382 deg, while FWHM of the structure we designed reaches 16.451 deg which is nearly three times of that of the conventional fiber.

4. SUMMARY AND CONCLUSIONS

The far field intensity of the special structure waveguide we designed has been calculated using the FD-BPM and Fraunhofer diffraction theory. It is found that the far field FWHM of this new structure increases greatly, reaching two times greater than that of the standard single mode fiber. The radiation angle expansion can effectively widen the steering scope of optical phased array. This novel structure has other applications in optical sensing, waveguide coupler, etc.

ACKNOWLEDGMENT

The authors thank the National Natural Science Foundation of China (NSFC) for the support (No. 60477002).

REFERENCES

1. Xiao, F. and W. Hu, "Optical phased array beam steering controlled by wavelength," *Applied Optics*, Vol. 44, No. 26, 5429–5433, 2005.
2. Nakamura, H. and T. Saiki, "FDTD simulation of tapered structure of near-field fiber probe," *Computer Physics Communications*, Vol. 142, 464–467, 2001.
3. Fielding, A. J., "Experimental observation of mode evolution in single-mode tapered optical fibers," *Journal of Lightwave Technology*, Vol. 17, No. 9, 1649–1656, 1999.
4. Chung, Y. and N. Dagli, "An assessment of finite difference beam propagation method," *J. Quantum Electron*, Vol. 26, No. 8, 1335–1339, 1990.
5. Hadley, G. R., "Transparent boundary condition for the beam propagation method," *J. Quantum Electron*, Vol. 28, No. 1, 363–370, 1992.
6. Guo, F.-Y. and B. Lin, "Characteristic beam parameters in diffracted field of fiber end face," *Journal of Zhejiang University (Engineering Science)*, Vol. 38, No. 3, 281–285, 2004.
7. Yevick, D. and B. Hermansson, "Efficient beam propagation techniques," *J. Quantum Electron*, Vol. 26, No. 1, 109–112, 1990.
8. Wang, X. and W. Hu, "Wavelength controlled optical phased array using photonic crystal," *Proceedings of the SPIE*, Vol. 6027, 147–155, 2006.

A Diagonal Split-cell Model for the High-order Symplectic FDTD Scheme

Wei Sha, **Xianliang Wu**, **and Mingsheng Chen**
Key Laboratory of Intelligent Computing & Signal Processing
Anhui University, Hefei 230039, China

Abstract— A high-order symplectic FDTD (SFDTD) scheme using the diagonal split-cell model is presented to analyze electromagnetic scattering of the curved three-dimensional perfectly conducting objects. One the one hand, for the undistorted cells, the fourth-order accurate spatial difference is employed. On the other hand, for the completely distorted cells, the treatment of the curved surfaces is based on the diagonal split-cell model. Finally, for the partially distorted cells, the interpolation strategy is proposed to keep the field components continuous. The numerical experiments suggest that the diagonal SFDTD scheme can obtain more accurate results than both the staircased SFDTD scheme and the traditional diagonal FDTD method. Furthermore, in view of the high numerical stability, the improved symplectic scheme does not need to decrease time increment to comply with the stability criterion.

1. INTRODUCTION

As the most standard algorithm, the traditional finite-difference time-domain (FDTD) method [1], which is second-order accurate in both space and time, has been widely applied to the electromagnetic computation and simulation. Unfortunately, for electrically large domains and for long-term simulation, the method consumes large computational resources owing to the limit of numerical dispersion and stability. Up to now, some more efficient solutions have been presented. For example, in 1989, Fang proposed high-order accurate FDTD method [2], which is fourth-order accurate in both space and time. But the method is hard to treat the varying of permittivity and permeability in the inhomogeneous domain on account of the application of third-order spatial derivatives to substitute for third-order correctional temporal derivatives. Another approach is to use fourth-order accurate Runge-Kutta (R-K) method [3] in the time direction and central-difference with Yee lattice in the space direction, yet, the method is dissipative and requires additional memory.

The symplectic schemes have demonstrated their advantage in energy conservation for the Hamiltonian system over other high-order methods [4]. A symplectic FDTD (SFDTD) scheme [5–7], which is explicit fourth-order accurate in both space and time, was introduced to the computational electromagnetism by Hirono for analyzing waveguide's eigenmode. It has been verified that the SFDTD scheme is nondissipative and saves memory. Moreover, the total field and scattered field technique and the near-to-far-field transformation [8, 9] have been further developed, by which the radar cross section (RCS) of dielectric sphere was successfully computed.

However, considering the high-order difference approximation for the spatial derivatives and the staircase model for the curved surfaces, the advantage of the SFDTD scheme cannot extend to electromagnetic scattering of the curved three-dimensional perfectly conducting objects. Here the diagonal SFDTD scheme is presented to overcome the problem. The Yee cells in the scheme are classified and discriminatingly handled, which not only eliminates the spurious solutions, but also maintains the numerical results accurate.

The paper is organized as follows. The formulation of the SFDTD scheme is given in Section 2, followed by the diagonal split-cell model for treating curved surfaces specified in Section 3, numerical results are presented in Section 4, and summary is concluded in Section 5.

2. GENERAL FORMULATION

The formulation of the x component of the normalized electric field ($\widetilde{E}_x = \sqrt{\varepsilon_0/\mu_0}E_x$) for the SFDTD scheme can be written as [7]

$$\widetilde{E}_x^{n+l/m}\left(i+\frac{1}{2},\, j,\, k\right) = \widetilde{E}_x^{n+(l-1)/m}\left(i+\frac{1}{2},\, j,\, k\right) + d_l$$
$$\times\left\{\lambda_1 \times\left[CFL_y \times\left(H_z^{n+(l-1)/m}\left(i+\frac{1}{2},\, j+\frac{1}{2},\, k\right) - H_z^{n+(l-1)/m}\left(i+\frac{1}{2},\, j-\frac{1}{2},\, k\right)\right)\right.\right.$$

$$-CFL_z \times \left(H_y^{n+(l-1)/m}\left(i+\frac{1}{2}, j, k+\frac{1}{2}\right) - H_y^{n+(l-1)/m}\left(i+\frac{1}{2}, j, k-\frac{1}{2}\right)\right)\Bigg]$$
$$+\left(\frac{\lambda_2}{3}\right)\times\Bigg[CFL_y \times \left(H_z^{n+(l-1)/m}\left(i+\frac{1}{2}, j+\frac{3}{2}, k\right) - H_z^{n+(l-1)/m}\left(i+\frac{1}{2}, j-\frac{3}{2}, k\right)\right)$$
$$-CFL_z \times \left(H_y^{n+(l-1)/m}\left(i+\frac{1}{2}, j, k+\frac{3}{2}\right) - H_y^{n+(l-1)/m}\left(i+\frac{1}{2}, j, k-\frac{3}{2}\right)\right)\Bigg]\Bigg\} \quad (1)$$

$$CFL_y = \frac{1}{\sqrt{\mu_0\varepsilon_0}}\frac{\Delta_t}{\Delta_y} \qquad CFL_z = \frac{1}{\sqrt{\mu_0\varepsilon_0}}\frac{\Delta_t}{\Delta_z} \quad (2)$$

where Δ_y, Δ_z are, respectively, the lattice space increments in the y and z coordinate directions, Δ_t is the time increment, d_l are the constant coefficients of the symplectic integrator, i, j, k, n, l and m are integers, $n + l/m$ denotes the l-th stage after the n-th time step, and m is the total stage number. Here we use $m = 5$, a five-stage fourth-order symplectic integrator is constructed. In addition, when $\lambda_1 = 1$ and $\lambda_2 = 0$, the expression is second-order accurate in space, and when $\lambda_1 = 9/8$ and $\lambda_2 = -1/8$, that is fourth-order accurate in space.

3. DIAGONAL SPLIT-CELL MODEL

For high-order spatial difference, every H-component is surrounded by eight E-components, and vice versa. In Fig. 1(a), the near four circulating E-components equidistant to the H-component can link one closed loop referred to as L_1, and those far can link another loop L_2. The closed PEC surface is notated by S. The L_1, L_2 and S can be viewed as the point set. We treat all the cells drawn in Fig. 1(b)∼(d) according to the following strategies.

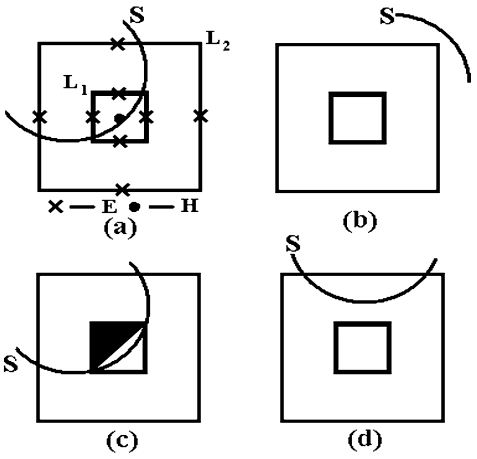


Figure 1: (a) The locations of the H-component and the E-components, (b) Undistorted cell, (c) Completely distorted cell. The black area denotes the diagonal split-cell, (d) Partially distorted cell.

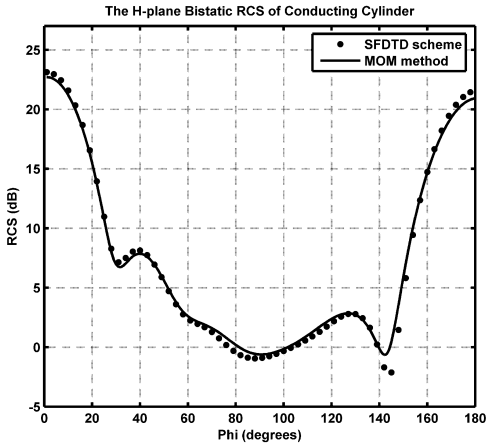


Figure 2: The bistatic RCS of conducting cylinder.

(1) Undistorted cells ($(L_1 \bigcup L_2) \bigcap S = \varnothing$). Here the fourth-order spatial difference is employed. The parameter λ_1 and λ_2 in (1) are taken to be $9/8$ and $-1/8$.

(2) Completely distorted cells ($L_1 \bigcap S \neq \varnothing$). The E-fields are treated in the traditional staircased FDTD way by setting $\lambda_1 = 1$ and $\lambda_2 = 0$. The H-fields are handled according to the diagonal split-cell model [11]. If and only if the circulating E-fields are designated as split along the cell diagonal, the H-fields are to be updated by setting $\lambda_1 = 2$ and $\lambda_2 = 0$. Besides, $\lambda_1 = 1$ and $\lambda_2 = 0$ should be set.

(3) Partially distorted cells ($L_1 \bigcap S = \varnothing$ and $L_2 \bigcap S \neq \varnothing$). Both E-fields and H-fields are initially computed by using the traditional staircased FDTD approach. Then, in order to eliminate the reflection from the undistorted cells to the completely distorted cells, the partially distorted cells are specially treated to keep the field components continuous. The interpolation equations for

the E-fields and H-fields are proposed as

$$F_\delta(h) = W_1 F_\delta(h) + W_2 \times \left[\frac{1}{2}\left(F_\delta(h+1) + F_\delta(h-1)\right)\right] \tag{3}$$

$$F = \widetilde{E}, \quad H, \quad \delta = x,\, y,\, z, \quad h = i,\, j,\, k \tag{4}$$

where W_1 and W_2 are weighting coefficients satisfying $W_1 + W_2 = 1$. Generally, for $W_1 \in [0.87,\, 0.97]$, the stable and accurate numerical results can be obtained. What's more, if the incident plane wave is a unit Gaussian pulse, the largest magnitude of the field values at the partially distorted cells is below the order of 10^{-3}.

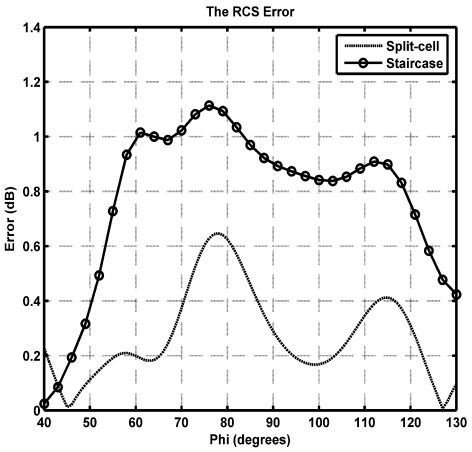


Figure 3: The absolute RCS Error.

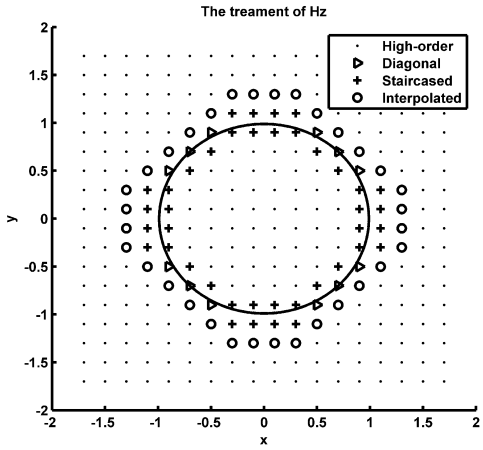


Figure 4: The location of spatial cells for the H_z component. The points denote the undistorted cells, the circles denote the partially distorted cells, and the pluses and triangles denote the completely distorted cells using the staircased and the diagonal approximation.

4. NUMERICAL RESULTS

Without loss of generality, it can be assumed that the incident plane wave with frequency of 300 MHz propagates along the z direction, and the electric field is polarized along the x direction, and the cubic grid is adopted, i.e., $\Delta_x = \Delta_y = \Delta_z = \Delta_\delta$ and $CFL_x = CFL_y = CFL_z = CFL_\delta$.

1. The scattering of conducting cylinder with the radius of 1 m and the height of 1 m is considered. The space increment and the Courant-Friedrichs-Levy (CFL) number are set to $\Delta_\delta = 0.1$ m and $CFL_\delta = 0.5$, respectively. The H-plane bistatic RCS plotted in Fig. 2 is computed by the SFDTD scheme using the diagonal split-cell model. The result calculated by the diagonal SFDTD scheme is in a good agreement with the reference solution calculated by the moment methods. The absolute RCS error curves are illustrated in Fig. 3. It can be seen that the diagonal SFDTD scheme can obtain more accurate results than the staircased SFDTD scheme. Similarly, the relative 2-norm error for the diagonal SFDTD scheme is 0.0556 compared with 0.0816 for the traditional diagonal FDTD method.

2. The monostatic RCS of conducting sphere with the radius of 1 m is calculated. The space increment is unchanged and the CFL number is retaken to $CFL_\delta = 0.70$. The Fig. 4 shows the location of spatial cells for the H_z component. Because of the utilization of low-order spatial difference near the curved boundaries and the symplectic structure in the time direction, the SFDTD solution shown in Fig. 5 still keeps accurate and stable after 5000 time steps.

3. The proposed scheme is employed to analyze the far response of conducting prism displayed in Fig. 6. The side length of the prism are, respectively, 1 m, 1 m, and $\sqrt{2}$ m, and the height is chosen to be 1 m. Under the same relative error condition, the diagonal SFDTD scheme occupies $71 \times 71 \times 71$ cells with $\Delta_\delta = 1.0/8.0$ and $CFL_\delta = 0.65$, by contrast, the traditional staircased FDTD approach occupies $111 \times 111 \times 111$ cells with $\Delta_\delta = 1.0/15.1$ and $CFL_\delta = 0.50$. About 32.4% memory and 39.5% CPU time are saved by the proposed scheme.

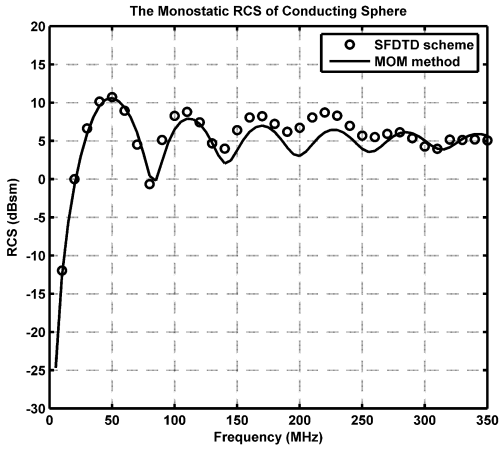


Figure 5: The monostatic RCS of conducting sphere.

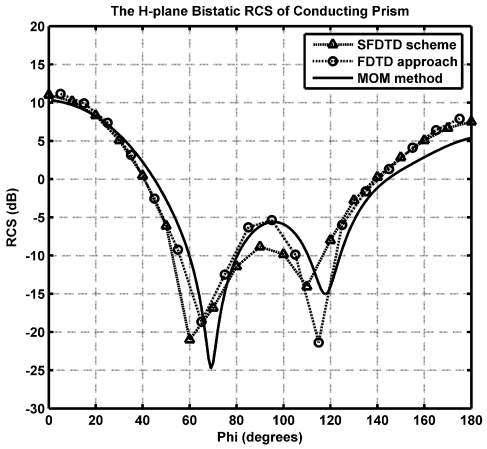


Figure 6: The bistatic RCS of conducting prism.

5. CONCLUSION

The high-order SFDTD scheme using the diagonal split-cell model can accurately and efficiently solve the scattering of three-dimensional perfectly conducting objects with curved metal boundaries. The high numerical stability of the scheme can obviate the instability problem due to the traditional diagonal approximation. Further more, the improved SFDTD scheme is easy to implement and places little additional computations on the original scheme. The future work will focus on the development of the proposed scheme in conjugation with the subgridding method.

ACKNOWLEDGMENT

This work was supported in part by "The national natural science foundation of China (No. 60371041)".

REFERENCES

1. Taflove, A., *Computational Electrodynamics: the Finite-Difference Time-Domain Method*, Artech House, Norwood, MA, 1995.
2. Fang, J., "Time domain finite difference computation for Maxwell's equations," *Dept. Elect. Eng.*, Ph.D. dissertation, Univ. California, Berkeley, CA, 1989.
3. Taflove, A., etc., *Advances in Computational Electrodynamics: The Finite-Difference Time-Domain Method*, Artech House, Norwood, MA, 1998.
4. Sanz-Serna, J. M. and M. P. Calvo, *Numerical Hamiltonian Problems*, Chapman & Hall, London, U.K., 1994.
5. Hirono, T., W. W. Lui, and S. Seki, "Successful applications of PML-ABC to the symplectic FDTD scheme with 4th-order accuracy in time and space," *IEEE MTT-S International Microwave Symposium Digest*, Vol. 3, 1293–1296, 1999.
6. Saitoh, I., Y. Suzuki, and N. Takahashi, "The symplectic finite difference time domain method," *IEEE Transactions on Magnetics*, Vol. 37, 3251–3254, 2001.
7. Hirono, T., W. Lui, S. Seki, and Y. Yoshikuni, "A three-dimensional fourth-order finite-difference time-domain scheme using a symplectic integrator propagator," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 49, 1640–1648, 2001.
8. Zhai, P. W., G. W. Kattawar, P. Yang, and C. H. Li, "Application of the symplectic finite-difference time-domain method to light scattering by small particles," *Applied Optics*, Vol. 44, 1650–1656, 2005.
9. Sha, W., Z. X. Huang, X. L. Wu, and M. S. Chen, "Total field and scattered field technique for fourth-order symplectic finite difference time domain method," *Chinese Physics Letters*, Vol. 23, 103–105, 2006.
10. Yoshida, H., "Construction of higher order symplectic integrators," *Physica D: Nonlinear Phenomena*, Vol. 46, 262, 1990.

11. Mezzanotte, P., L. Roselli, and R. Sorrentino, "A simple way to model curved metal boundaries in fdtd algorithm avoiding staircase approximation," *IEEE Microwave and Guided Wave Letters*, Vol. 5, 267–269, 1995.

Application of Rational Function Approximation Technique to Hybrid FE/BI/MLFMA for 3D Scattering

Zhen Peng[1,2,3] **and Xin-Qing Sheng**[1]

[1]Department of Electronic Engineering, Beijing Institute of Technology
Beijing, China
[2]Institute of Electronics, Chinese Academy of Sciences, China
[3]Graduated University of Chinese Academy of Sciences, China

Abstract— The rational function approximation technique (RFAT) is applied to the hybrid finite-element/boundary-integral/multilevel fast multipole algorithm (FE/BI/MLFMA) to acquire wide-band and wide-angle backscatter radar-cross-section (RCS) by complex targets in this paper. The two approaches of utilizing the rational function approximation technique, asymptotic waveform evaluation (AWE) and model-based parameter estimation (MBPE), both have been investigated and compared by theoretical analysis and numerical experiments. The numerical results acquired by the developed computing algorithm of integrating the hybrid FE/BI/MLFMA with the RFAT are presented in the paper, demonstrating that the rational function approximation technique can greatly speed up the hybrid FE/BI/MLFMA to acquire wide-band and wide-angle backscatter radar-cross-section (RCS) by complex targets.

1. INTRODUCTION

The electromagnetic scattering by large complex targets is one of important and challenge problems, which attracts a great number of researchers over past years. It has been proven that the hybrid finite-element/boundary-integral/multilevel fast multipole algorithm (FE/BI/MLFMA) is a suitable and powerful method for this problem, which is robust, efficient, and general [1–5]. However, the conventional way to acquire RCS over a frequency band and angle range by this hybrid FE/BI/MLFMA is to compute one frequency by one frequency and one angle by one angle, which is very time-consuming. For many practical applications, our interest often includes a wide band of frequencies and a wide range of angles. Hence, how to efficiently compute RCS over a wide band and wide angle is very important and urgent.

The previous works show that employing the rational function approximation is a kind of efficient way to compute RCS over a wide band and wide angle. There are two approaches to generate the rational approximated function. The first one is an interpolation model known as the model-based parameter estimation (MBPE). Miller [6–8] presents detailed descriptions of MBPE with regards to specific problems in electromagnetic (EM) observables. In [9, 10], MBPE is used to interpolate antenna radiation patterns in both the spatial and frequency domains. In [11], MBPE is employed in the moment of method (MOM) to compute exterior and interior resonances of antenna. Recently, a two-stage approach of MBPE is presented to construct approximated models of frequency-domain responses of grounding systems to accelerate the MOM calculations [12]. The second approach to generate rational approximated function is the asymptotic waveform evaluation (AWE), which was originally developed for high-speed circuit analysis [13]. This AWE has been applied to the finite element and finite difference analysis of electromagnetic problem [14–16]. In [17, 18], AWE has also been applied to MOM for solving scattering problem. Recently, the Multipoint Galerkin AWE (MGAWE) and the well-conditioned AWE (WCAWE) have been developed in [19, 20].

In this work, the rational function approximation technique is applied to the hybrid FE/BI/MLFMA for computing scattering by large, complex targets. The both AWE and MBPE for generating the rational approximated function have been applied to the hybrid FE/BI/MLFMA. The advantages and disadvantages of AWE and MBPE are analyzed and discussed. The numerical results computed by the developed algorithm of integrating hybrid FE/BI/MLFMA with the rational function approximation technique are presented to demonstrate that the rational function approximation technique can greatly speed up the hybrid FE/BI/MLFMA to acquire wide-band and wide-angle backscatter radar-cross-section (RCS) by complex targets.

2. APPLICATION OF RFAT TO HYBRID FE/BI/MLFMA

The complete hybrid FE/BI system is

$$\begin{bmatrix} K_{II} & K_{IS_I} & 0 \\ K_{S_II} & K_{S_IS_I} & B \\ 0 & P & Q \end{bmatrix} \begin{Bmatrix} E_I \\ E_{S_I} \\ H_S \end{Bmatrix} = \begin{Bmatrix} 0 \\ 0 \\ b \end{Bmatrix} \tag{1}$$

where $[K_{II}]$, $[K_{S_IS_I}]$, $[K_{IS_I}]$, $[K_{S_II}]$, $[B]$ are sparse FEM matrixes, $[P]$ and $[Q]$ are dense BI matrixes. To compute large objects, the powerful MLFMA has to employ to speed up the matrix-vector multiplications of $[P]$, $[Q]$ and $\{E_{S_I}\}$, $\{H_{S_I}\}$ in the iterative solvers such as CGNS or GMRES for (1). Rewrite Equation (1) as

$$\boldsymbol{Z}(k)\boldsymbol{I}(k,\,\theta,\,\phi) = \boldsymbol{V}(k,\,\theta,\,\phi) \tag{2}$$

The coefficient matrix $\boldsymbol{Z}(k)$ is a function of frequency, the unknown vector $\boldsymbol{I}(k,\,\theta,\,\phi)$ and the right side of the equation $\boldsymbol{V}(k,\,\theta,\,\phi)$ is a function of both frequency and incident angle. Given the frequency f_0 (corresponding to the wave number k_0) and the incident angle $(\theta_0,\,\phi_0)$, $\boldsymbol{I}(k,\,\theta,\,\phi)$ can be obtained by solving (2). Then, the RCS by the target can be computed using $\boldsymbol{I}(k,\,\theta,\,\phi)$. The goal of this paper is to apply the rational function approximation technique to (9) to efficiently compute $\boldsymbol{I}(k,\,\theta,\,\phi)$ over a wide band of frequency and a wide range of angle, which will be detailed in the following sections.

2.1. Application of MBPE To Hybrid FE/BI/MLFMA

The basic idea of MBPE is to firstly represent the unknown vector $\boldsymbol{I}(k,\,\theta,\,\phi)$ in a form of a fractional polynomial function as

$$I(x) = \frac{N(x)}{D(x)} = \frac{N_0 + N_1x + N_2x^2 + \cdots + N_nx^n}{D_0 + D_1x + D_2x^2 + \cdots + D_dx^d} \tag{3}$$

where x represents the frequency f or the angle $(\theta,\,\varphi)$ and $D_d = 1$. Then, the $n+d+1$ unknown coefficients in (3) is determined by matching $I(x)$ at a total of $m(m \geq n+d+1)$ sampling points, which is called point-matching MBPE, or by matching a total number of $m(m \geq n+d+1)$ derivative samples at one point or several points, which is called Frequency-Derivative Interpolation (FDI) [21]. FDI has been proven to be equivalent to AWE in [17]. Let us consider how to determine unknown coefficients in detail in the point-matching MBPE. By sampling $I(x)$ at a total of $m(m \geq n+d+1)$ sampling points, which could be either frequency points or angle points, the following equation can be obtained

$$I_i\underline{D}_i = \underline{N}_i, \quad I = 1, \ldots, m \tag{4}$$

with

$$\begin{aligned} I_i &= I(x_i) \\ \underline{D}_i &= D(x_i) = D_0 + D_1x_i + D_2(x_i)^2 + \cdots + D_d(x_i)^d \\ \underline{N}_i &= N(x_i) = N_0 + N_1x_i + N_2(x_i)^2 + \cdots + D_n(x_i)^n \end{aligned} \tag{5}$$

Then the unknown coefficient vector $\boldsymbol{u} = [N_0\ N_1\ \cdots\ N_n\ D_0\ D_1\ \cdots\ D_{d-1}]^T$ can be determined by

$$\boldsymbol{A}\boldsymbol{u} = \boldsymbol{B} \tag{6}$$

$$\boldsymbol{A} = \begin{bmatrix} 1 & x_1 & \cdots & x_1^n & -I_1 & -I_1x_1 & \cdots & -I_1x_1^{d-1} \\ 1 & x_2 & \cdots & x_2^n & -I_2 & -I_2x_2 & \cdots & -I_2x_2^{d-1} \\ 1 & x_3 & \cdots & x_3^n & -I_3 & -I_3x_3 & \cdots & -I_3x_3^{d-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_m & \cdots & x_m^n & -I_m & -I_mx_m & \cdots & -I_mx_m^{d-1} \end{bmatrix} \tag{7}$$

$$B = \begin{bmatrix} I_1x_1^d & I_2x_2^d & I_3x_3^d & \cdots & I_mx_m^d \end{bmatrix}^T \tag{8}$$

After the unknown coefficient vector $\boldsymbol{u}$ is obtained by solving (6), the rational function (3) is generated, from which $I(k,\,\theta,\,\varphi)$ at any frequency or angle can be computed.

2.2. Application of AWE to Hybrid FE/BI/MLFMA

The basic idea of AWE is to firstly expand the function $\boldsymbol{I}(k,\,\theta,\,\phi)$ into Taylor series, then generate the Pade rational function through the Pade Approximant. Taking $\boldsymbol{I}(k)$ in (2) into consideration, we expand $\boldsymbol{I}(k)$ into a Taylor series:

$$\boldsymbol{I}(k)=\sum_{n=0}^{Q}m_n\left(k-k_0\right)^n=\sum_{n=0}^{Q}\frac{\boldsymbol{I}^{(n)}\left(k_0\right)}{n!}\left(k-k_0\right)^n \tag{9}$$

where $\boldsymbol{I}^{(n)}\left(k_0\right)=\dfrac{d^n}{dk^n}\boldsymbol{I}(k)|_{k=k_0}$, and k_0 is the expansion point. Expanding the coefficient matrix $Z(k)$ and the excitation vector $V(k)$ into the Taylor series, and finally matching the coefficients on both sides yield the recursive relation for the moment vectors:

$$\boldsymbol{m}_0=\boldsymbol{I}\left(k_0\right)=\boldsymbol{Z}^{-1}\left(k_0\right)\boldsymbol{V}\left(k_0\right) \tag{10}$$

$$\boldsymbol{m}_n=\boldsymbol{Z}^{-1}\left(k_0\right)\left[\frac{\boldsymbol{V}^{(n)}}{n!}-\sum_{i=1}^{n}\frac{\boldsymbol{Z}^{(i)}\left(k_0\right)\boldsymbol{m}_{n-i}}{i!}\right] \tag{11}$$

where $\boldsymbol{Z}^{-1}(k_0)$ denotes the inverse of $\boldsymbol{Z}(k_0)$, $\boldsymbol{Z}^{(i)}(k_0)$ denotes the ith derivative of $\boldsymbol{Z}(k_0)$, and likewise, $\boldsymbol{V}^{(n)}$ denotes the nth derivative of $\boldsymbol{V}$. Utilizing the Pade Approximant, we represent $\boldsymbol{I}(k)$ as a Pade rational function

$$\boldsymbol{I}(k)=\sum_{n=0}^{L+M+1}\boldsymbol{m}_n\left(k-k_0\right)^n=\frac{\sum\limits_{i=0}^{L}\boldsymbol{a}_i\left(k-k_0\right)^i}{\sum\limits_{j=1}^{M}\boldsymbol{b}_j\left(k-k_0\right)^j+1} \tag{12}$$

where $L+M+1=Q$.

The key of the implementation of applying AWE to the hybrid FE/BI/MLFMA is the computation of the derivative matrix $\boldsymbol{Z}^{(i)}(k_0)$ and the derivative vector $\boldsymbol{V}^{(n)}(k_0)$. The following will give the details of computing matrix $\boldsymbol{Z}^{(i)}(k_0)$ and vector $\boldsymbol{V}^{(n)}(k_0)$. It is known that the matrix $\boldsymbol{Z}^{(i)}(k_0)$ consists of the full BI matrix and the sparse FEM matrix. The FEM matrix is a simple function of frequency, which can be obtained from the following nth derivative of functional. The BI matrix is more complex, which is generated from CFIE integral equation. The kernel operator of CFIE is $\boldsymbol{L}_p$ and $\boldsymbol{K}_p$, whose nth derivatives are given by

$$\boldsymbol{g}_i\cdot\boldsymbol{L}^n(k)=\frac{jk}{4\pi}\iint\limits_{s\;s'}\boldsymbol{g}_i\cdot\boldsymbol{g}_j(-jR)^n\left(1-\frac{n}{jkR}\right)\frac{\exp(-jkR)}{R}dsds'$$

$$-\frac{jk}{4\pi}\iint\limits_{s\;s'}\boldsymbol{g}_i\cdot\left\{\nabla'\cdot\boldsymbol{g}_j\left[-\frac{j\boldsymbol{R}}{kR^2}(-jR)^n\exp(-jkR)-\frac{\boldsymbol{R}}{R^2}\left(\frac{1}{k^2}(-jR)^n\left(\sum_{p=0}^{n}\frac{\overline{n!}}{\overline{(n-p)!}}\frac{\exp(-jkR)}{(jkR)^p}\right)\right)\right]\right\}dsds' \tag{13}$$

$$\hat{\mathrm{n}}\times\boldsymbol{g}_i\cdot\boldsymbol{K}^{(n)}(k)=-\frac{1}{4\pi}\iint\limits_{s\;s'}\boldsymbol{g}_i\boldsymbol{g}_j\times\boldsymbol{R}\left[j\exp(-jkR)(-jR)^{n-3}(1+jkR-n)\right]dsds' \tag{14}$$

$\boldsymbol{V}^{(n)}(k_0)$ can be obtained by discretizing the following equation

$$\boldsymbol{E}^{(n)}\left(k_0\right)+\hat{\mathrm{n}}\times\bar{\boldsymbol{H}}^{(n)}\left(k_0\right)=\left(-\boldsymbol{j}\hat{\boldsymbol{k}}^i\cdot\boldsymbol{r}\right)^n\left(\boldsymbol{E}^i+\hat{\mathrm{n}}\times\bar{\boldsymbol{H}}^i\right) \tag{15}$$

To sum up, the implementation of applying AWE to the hybrid FE/BI/MLFMA has four steps. First, we compute the derivative matrix $\boldsymbol{Z}^{(i)}(k_0)$ and vector $\boldsymbol{V}^{(n)}(k_0)$ using formulation (16)–(21). Then, we compute $\boldsymbol{m}_n$ using (11)–(12). Afterwards, coefficient $\boldsymbol{a}_i$ and $\boldsymbol{b}_i$ are obtained by solving Equation (14)–(15). Finally, the Pade rational function is generated from (13), from which $\boldsymbol{I}(k)$ at any frequency point near the expansion point f_0 can be computed.

2.3. Comparisons between AWE and MBPE

The formulation of application of AWE and MBPE to hybrid FE/BI/MLFMA has been described in details in the above sections. In this section, the numerical performance of AWE and MBPE will be compared analytically and numerically.

Let us first analyze the advantages and disadvantages of AWE and MBPE. First, it is easy to see that MBPE is an interpolation technique, whereas AWE is an extrapolation technique, which makes AWE have an advantage over MBPE — acquiring RCS at high frequencies using a coarse mesh. Second, MBPE does not increase the memory requirements, whereas AWE leads to the increasing of memory requirements, since both the coefficient matrix and the derivatives of the coefficient matrix in AWE need to be stored. Third, more importantly, MBPE can be easily incorporated into the hybrid FE/BI/MLFMA, whereas AWE is very difficult to be incorporated into the hybrid FE/BI/MLFMA since there is an essential difficulty when MLFMA is employed to accelerate the multiplications of the derivatives of the coefficient matrix in AWE and vectors.

3. NUMERICAL EXPERIMENTS

In this section, we compute the scattering by large targets in a wide band of frequency and a wide range of angle to show the efficiency and generality of the MBPE-incorporated FE/BI/MLFMA.

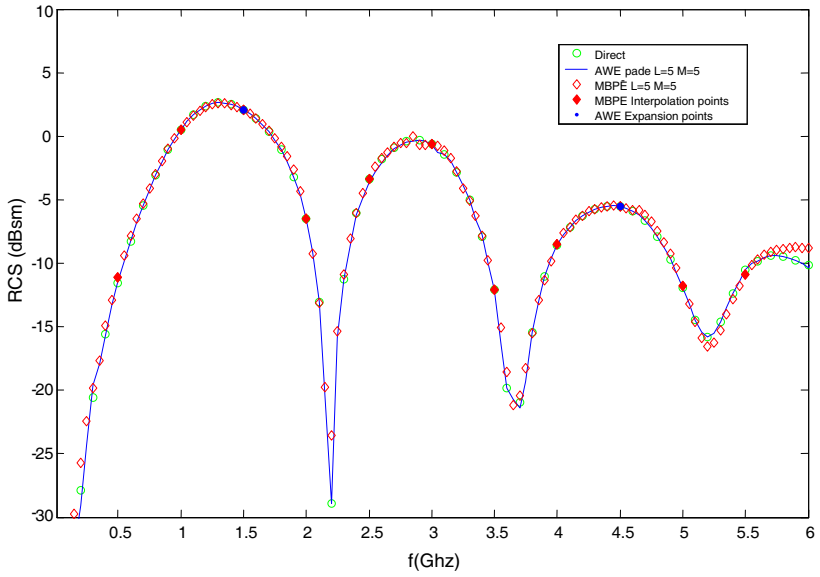


Figure 1: Monostatic RCS of a coated sphere from 0.2 GHz to 6 GHz using the direct, AWE- and MBPE-incorporated FE/BI/MLFMA.

Table 1: The comparisons of CPU time and memory requirement for the wide frequency band computation using the Direct, AWE- and MBPE-Incorporated FE/BI/MLFMA.

Problem	Direct			AWE				MBPE			
	Calculated points	Total CPU times(s)	Memory (MB)	Expansion points	Total CPU times(s)	Memory (MB)	Speed-up factor	Interpolation points	Total CPU times(s)	Memory (MB)	Speed-up factor
Fig. 1 Results	60	17580	11.2	2	2006	104.0	8.76	11	1923	11.0	9.15
Fig. 2 Results	65	2355	11.0	2	301	99.0	7.82	11	292	11.0	8.07

The first example is a coated sphere. The coated sphere has a radius of 4 cm and its conducting core has a radius of 3 cm. The dielectric coating has a relative permittivity $\varepsilon_r = 4$. The direct Hybrid FE/BI/MLFMA takes 17580 seconds to compute the monostatic RCS from 0.2 to 6.0 GHz with a frequency step of 0.1 GHz; the AWE-incorporated FE/BI/MLFMA only takes 2006 seconds

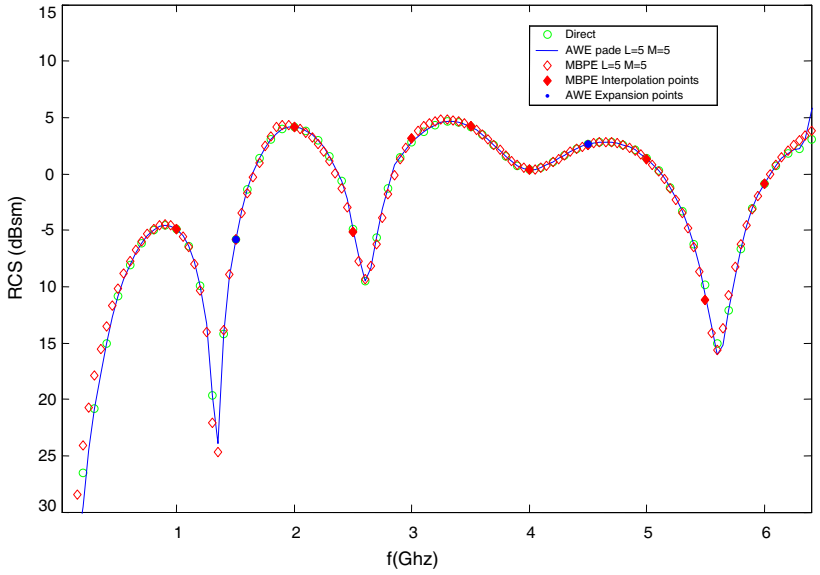


Figure 2: Monostatic RCS of the brick with cavity at $\theta = 0°$,.$\varphi = 0°$ from 0.1 GHz to 6.5 GHz using the direct, AWE- and MBPE-incorporated FE/BI/MLFMA.

with a smaller frequency step of 0.05 GHz and two expansion points located at 1.5 GHz and 4.5 GHz and tenth-order rational function ($L = 5$, $M = 5$); the MBPE-incorporated FE/BI/MLFMA takes 1923 seconds with eleven interpolation points and ten-order ration function ($L = 5$, $M = 5$). The comparison of the calculated RCS results is given in Fig. 3, and the computational information is also given in Table 1.

The second example is a conducting brick with empty cavity illustrated in Fig. 4. The size of the conducting brick is 5 cm×5 cm×10 cm and the cavity has a size of 3 cm×3 cm×3 cm. First, we employ the direct, AWE-incorporated and MBPE-incorporated FE/BI/MLFMA to compute the monostatic RCS from 0.2 GHz to 6.5 GHz illuminated by a plane wave normal incident. It takes 2355, 301, 292 seconds for the direct, AWE-incorporated and MBPE-incorporated FE/BI/MLFMA respectively. The calculated results are given in Fig. 5, from which again we see the agreement is quite good. The computational information is given in Table 1. Next, we compute the monostatic RCS illuminated under the incident angle range of $\phi \in (0° - 180°)$, $\theta = 0°$, and $\theta \in (0° - 180°)$, $\phi = 0°$ at 3 GHz. The calculated results are given in Fig. 6 and Fig. 7. In this calculation, only one expansion point at $\phi = 90°$ is used in AWE and eleven interpolation points in MBPE in Fig. 6, and two expansion points at $\theta = 45°$ and $\theta = 135°$ in AWE and eleven interpolation points in MBPE in Fig. 7. The CPU time and memory requirement are presented in Table 2.

It is well known that larger targets, more complex curves of RCS. Hence, more sampling points are required in MBPE for larger targets. The numerical experiments show that a segment in the RCS curve, from valley point — through apex point — again to valley point, usually can be approximated by only one or two interpolation points in MBPE, whereas ten points are usually required to obtain the segment by the direct computation. Thus, the MBPE-incorporated FE/BI/MLFMA

Table 2: The comparisons of CPU time and memory requirement for the wide angle range computation using the Direct, AWE- and MBPE-Incorporated FE/BI/MLFMA.

Problem	Direct			AWE				MBPE			
	Calculated points	Total CPU times(s)	Memory (MB)	Expansion points	Total CPU times(s)	Memory (MB)	Speed-up factor	Interpolation points	Total CPU times(s)	Memory (MB)	Speed-up factor
Fig. 3 Results	180	2601	11.0	1	107	11.0	24.31	11	159	11.0	16.36
Fig. 4 Results	180	2610	11.0	2	201	11.0	12.99	11	159	11.0	16.36

usually has 5–10 times faster than the direct FE/BI/MLFMA for either small or large targets.

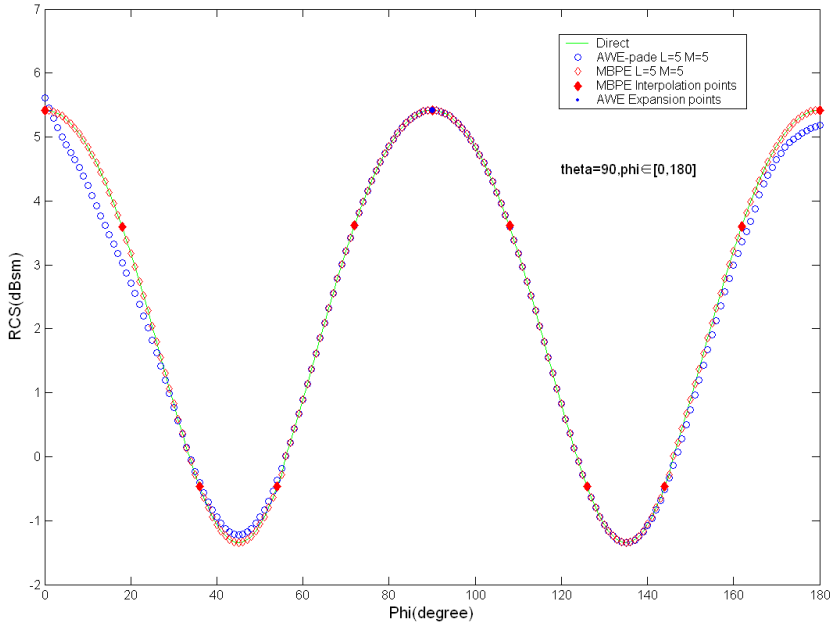


Figure 3: Monostatic RCS of the brick with cavity at 3 GHz and $\theta = 0°$ using the direct, AWE- and MBPE-incorporated FE/BI/MLFMA.

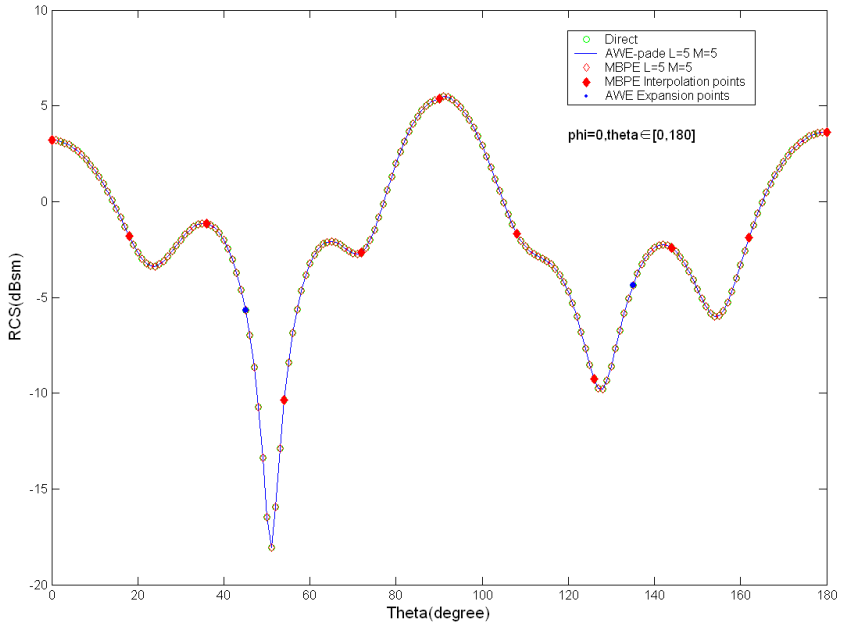


Figure 4: Monostatic RCS of the brick with cavity at 3 GHz and $\varphi = 0°$ using the direct, AWE- and MBPE-incorporated FE/BI/MLFMA.

4. CONCLUSION

The rational function approximation technique is first applied to the hybrid FE/BI/MLFMA for scattering by complex targets. The two approaches of applying RFAT to hybrid FE/BI/MLFMA, AWE and MBPE, are compared analytically and numerically, demonstrating that MBPE is more suitable for the hybrid FE/BI/MLFMA. The MBPE-incorporated FE/BI/MLFMA has proven to be accurate and efficient for computing scattering by large, complex targets in a wide band

of frequency and a wide range of angle. The analysis and numerical experiments show that the MBPE-incorporated FE/BI/MLFMA usually can achieve 5–10 times faster than the direct FE/BI/MLFMA.

REFERENCES

1. Song, J., C. Lu, and W. C. Chew, "Multilevel fast multipole algorithm for electromagnetic scattering by large complex objects," *IEEE Trans. Antennas Propagat.*, Vol. 45, No. 10, 1488–1493, Oct. 1997.
2. Jin, J., *The Finite Element Method in Electromagnetics*, Wiley, New York, 1993.
3. Sheng, X.-Q., J. Song, C. Lu, and W. C. Chew, "On the formulation of hybrid finite-element and boundary-integral method for 3D scattering," *IEEE Trans. Antennas Propagat.*, Vol. 46, No. 3, 303–311, Mar. 1998.
4. Liu, J. and J. Jin, "A novel hybridization of higher order finite element and boundary integral methods for electromagnetics scattering and radiation problems," *IEEE Trans. Antennas Propagat.*, Vol. 49, No. 9, 1794–1806, Dec. 2001.
5. Sheng, X.-Q. and E. K. N. Yung, "Implementation and experiments of a hybrid algorithm of the finite-element boundary-integral method for open-region inhomogeneous electromagnetic problems,"*IEEE Trans. Antennas Propagat.*, Vol. 50, No. 2, 163–167, Feb. 2002.
6. Miller, E. K., "Model-based parameter estimation in electromagnetics-I: Background and theoretical development," *Appl. Comput. Electromagn. Soc. Newslett.*, Vol. 10, No. 3, 40–63, 1995.
7. Miller, E. K., "Model-based parameter estimation in electromagnetics-II: Applications to EM observables," *Appl. Comput. Electromagn. Soc. Newslett.*, Vol. 11, No. 1, 35–56, 1996.
8. Miller, E. K., "Model-based parameter estimation in electromagnetics-III: Applications to EM integral equations," *Appl. Comput. Electromagn. Soc. Newslett.*, Vol. 10, No. 3, 9–29, 1995.
9. Werner, D. H. and R. J. Allard, "The simultaneous interpolation of antenna radiation patterns in both the spatial and frequency domains using model-based parameter estimation," *IEEE trans. Antennas propagat.*, Vol. 48, No. 3, 383–400, Mar. 2000.
10. Allard, R. J. and D. H. Werner, "The model-based parameter estimation of antenna radiation patterns using windowed interpolation and spherical harmonics," *IEEE trans. Antennas propagat.*, Vol. 51, No. 8, 1891–1906, Aug. 2003.
11. Li, L. and C. H. Liang, "Generalized system function analysis of exterior and interior resonances of antenna and scattering problems," *IEEE trans. Antennas propagat.*, Vol. 52, No. 8, 2064–2072, Aug. 2004.
12. Guo, J., J. Zou, B. Zhang, J. L. He, and Z. C. Guan, "An interpolation model to accelerate the frequency-domain response calculation of grounding systems using the method of moments," *IEEE trans. Power Del.*, Vol. 21, No. 1, 121–128, Jan. 2006.
13. Pillage, L. T. and R. A. Rohrer, "Asymptotic waveform evaluation for timing analysis," *IEEE Trans. Computer-Aided Design*, Vol. 9, No. 4, 352–366, Apr. 1990.
14. Zhang, J. P. and J. M. Jin, "Preliminary study of AWE method for fem analysis of scattering problems," *Microwave Opt. Tech. Lett.*, Vol. 17, No. 1, 7–12, Jan. 1999.
15. Kolbehdari, M. A., M. Srinivasan, M. S. Nakhla, Q. J. Zhang, and R. Achar, "Simultaneous time and frequency domain solutions of EM problems using finite element and CFH techniques," *IEEE Trans. Microwave Theory Tech.*, Vol. 44, No. 9, 1526–1934, Sep. 1996.
16. Polstyanko, S. V., R. D. Edlinger, and J. F. Lee, "Fast frequency sweep technique for the efficient analysis of dielectric waveguides," *IEEE Trans. Microwave Theory Tech.*, Vol. 45, No. 7, 1118–1126, July 1997.
17. Reddy, C. J., M. D. Deshpande, C. R. Cockrell, and F. B. Beck, "Fast RCS computation over a frequency band using method of moments in conjunction with asymptotic waveform evaluation technique," *IEEE trans. Antennas propagat.*, Vol. 46, No. 8, 1229–1233, Aug. 1998.
18. Jiao, D., X. Y. Zhu, and J. M. Jin, "Fast and accurate frequency-sweep calculation using asymptotic waveform evaluation and the combined-field intrgral equation," *Radio Science*, Vol. 34, No. 5, 1055–1063, Oct. 1999.
19. Slone, R. D., R. Lee, and J. F. Lee, "Multipoint galerkin asymptotic wave form evaluation for model order reduction of frequency domain FEM electromagnetic radiation problems," *IEEE trans. Antennas propagat.*, Vol. 49, No. 10, 1504–1513, Oct. 2001.
20. Slone, R. D., R. Lee, and J. F. Lee, "Well conditioned asymptotic waveform evaluation for finite elements," *IEEE trans. Antennas Propagat.*, Vol. 51, No. 9, 2442–2447, Sep. 2003.

21. Burke, G. J. and E. K. Miller, "Use of frequency-derivative information to reconstruct an electromagnetic transfer function," *Fourth Annual ACES Review*, 22–34, Naval Postgraduate School, Monterey, CA, March 1988.

A GL Metro Carlo EM Inversion

Ganquan Xie, **Jianhua Li**, **Lee Xie**, **and Feng Xie**
GL Geophysical Laboratory, USA

Abstract— In this paper, we propose a Global and Local (GL) Metro Carlo (GLMC) Electromagnetic (EM) inversion method to recover the EM material parameters. The GLMC inversion is a compromising method between the deterministic and random Monte Carlo simulation methods. The GL modeling strategy simulation process is used to construct the GLMC inversion. Our GL EM modeling method is very fast to calculate data when the EM parameters are changed in each one sub domain successively, and can greatly increase the speed of GLMC inversion. The GLMC EM inversion is fast, stable and has reasonable resolution. The GLMC inversion method can be extended to the seismic, acoustic and their coupled inversion. The GLMC inversion algorithms and software are also developed.

1. INTRODUCTION

We have proposed the global and local electromagnetic field modeling and inversion in PIERS 2005 in Hangzhou of China [1]; The paper "GL method and its advantages for resolving historical difficulties" has been published in PIER 63, and in PIERS 2006 in Cambridge [2]. The paper "SGILD stochastic modeling and inversion" [3] has been published in Physica D. In the AGILD stochastic modeling and inversion [4], we derived the EM statistical field modeling equations and the AGILD pair equations of the parameter moments, then we presented the SAGILD approach to recover the ensemble mean, covariance moments, standard derivation and obtained confidence interval. In this paper, we propose a GLMC EM inversion method [5] to recover the EM parameters. Using the GLMC process and Bayes' formula, the GLMC Markov Chain EM inversion is developed for recovering EM parameter ensemble mean moment.

There are many research papers to study Monte Carlo forward and inverse simulations. The book [6] described the frame of the Markov Chain Monte Carlo (MCMC) practice. The MT inversion using Bayesian statistics was described in [7]. The paper [8] described an electromagnetic scattering model for rice canopy based on Monte Carlo simulation. The book [9] described the Monte Carlo methods in condensed matter physics. In the MCMC inversion, forward modeling need to be solved repeatedly when the parameters are changed in each sub domain. If the finite difference (FD) or finite element method (FEM) forward methods are used in the simulation, the high cost of computational time is a main and unaccepted disadvantage of the MCMC. The inaccurate absorption conditions on the artificial boundary make noises to downgrade the resolution in the MCMC inversion. The deterministic EM inversion may be too determinism and causes the sensitive ill posed. In the other hand, the MCMC inversion may be too random and fuzz up the resolution of the target. It is necessary to introduce a metro path to connect and compromise the deterministic and random MCMC inversion. In this paper we propose the GL Metro Carlo (GLMC) inversion [5] and the GLMC Markov Chain inversion simulation. The GL Metro Carlo EM inversion is a compromising inversion between the deterministic inversion and random Monte Carlo simulation. We use the Metro Carlo instead of the Monte Carlo to distinguish our GLMC inversion in this paper. The GL EM modeling method [1] [2] is very fast to calculate model data when the EM parameters are changed in each sub domain, and greatly increase the speed of GLMC inversions. The GL method has no artificial boundary and no inaccurate absorption boundary condition that reduces the modeling noise for the GLMC inversion. There is no big matrix to be solved in the GLMC that effectively reduces the computational time cost. The GLMC EM inversion method can be useful for material sciences, nanometer materials, surface physics, rough surface, medical sciences, biology, geophysical exploration and finance.

The plan of this paper is as follows. The introduction has been described in the Section 1. In the Section 2, we describe the 3D GL EM modeling. The GLMC EM inversions are presented in the Section 3. In the Section 4, we describe the advantages of the GLMC and GLMC Markov Chain inversions. The simulations and application are described in the Section 5. Finally, we conclude our paper in the Section 6.

2. 3D GL AND AGILD EM MODELING

We have proposed the 3D GL EM modeling in PIERS 2005 in Hangzhou of China [1] and PIERS 2006 in Cambridge [2]. We briefly describe it here and use it to construct the GLMC EM inversion.

2.1. The 3D GL EM Modeling

The 3D GL modeling method is as follows:

(2.1) The domain Ω is divided into a set of the N mesh or meshless sub domains Ω_k, $\Omega = \bigcup_{k=1}^{N} \Omega_k$.

(2.2) In each Ω_k, we solve the EM Green's tensor integral equation system based on the equations (1). By dual operation, the equations are reduced into 6×6 matrix equations. By solving the 6×6 equations, we obtain Green's tensor field $E_k^J, \cdots, H_k^M$.

(2.3) We improved the Global EM field $[E_k(r), H_k(r)]$ by the local scattering filed,

$$\begin{bmatrix} E(r) \\ H(r) \end{bmatrix}_k = \begin{bmatrix} E(r) \\ H(r) \end{bmatrix}_{k-1} + \int_{\Omega_k} G_{E,H,k}^{J,M}(r', r) \left([D] - [D]_b\right) \begin{bmatrix} E(r') \\ H(r') \end{bmatrix}_{k-1} dr', \tag{1}$$

$k = 1, 2, \cdots, N$, successively, the 6×6 EM Green's tensor function

$$\begin{aligned} G_{E,H,k}^{J,M}(r', r) &= \begin{bmatrix} E_k^J(r', r) & H_k^J(r', r) \\ E_k^M(r', r) & H_k^M(r', r) \end{bmatrix}, \\ [D] &= \begin{bmatrix} (\sigma + i\omega\varepsilon) I & 0 \\ 0 & -i\omega\mu I \end{bmatrix}, \\ [D]_b &= \begin{bmatrix} (\sigma + i\omega\varepsilon)_b I & 0 \\ 0 & -i\omega\mu_b I \end{bmatrix}, \end{aligned} \tag{2}$$

where D and D_k are 6×6 matrix, I is 3×3 identity matrix. The $[E_N(r), H_N(r)]$ is the GL solution of the 3D GL EM modeling.

The 3D GL modeling is available for the multiple dimensional EM, elastic, acoustic, mechanical and the stochastic GL modeling.

2.2. The 3D AGILD EM modeling

The EM strip second type integral equation on the boundary strip and the Galerkin equation in the internal domain for the EM field are coupled to construct AGILD modeling. The stochastic SAGILD modeling [4] can calculate the statistical moment field.

3. THE GLMC EM INVERSION

3.1. The GLMC EM inversion

We propose the GLMC inversion as follows.

(1) The domain Ω is divided into set of the N mesh or meshless sub domains Ω_k, $\Omega = \bigcup_{k=1}^{N} \Omega_k$.

(2) Let

$$[D]_{i,k} = \begin{bmatrix} (\sigma_{i,k} + i\omega\varepsilon_{i,k}) I & 0 \\ 0 & -i\omega\mu_{i,k} I \end{bmatrix}. \tag{3}$$

(3) When $l = 0$, $[D]_k^{(0)} = [D]_b$, by the $(l-1)^{th}$ iteration, we have

$$[D]_{i_{(k)},k}^{(l-1)} = \begin{bmatrix} \left(\sigma_{i_{(k)},k}^{(l-1)} + i\omega\varepsilon_{i_{(k)},k}^{(l-1)}\right) I & 0 \\ 0 & -i\omega\mu_{i_{(k)},k}^{(l-1)} I \end{bmatrix}, k = 1, 2, \cdots, N. \tag{4}$$

(4) In each Ω_k, $i = 1, 2, \cdots, M$, we solve the EM Green's tensor integral equation system based on the equations (1). By dual operation, the equations are reduced into 6×6 matrix equations. By solving the 6×6 matrix equations, we obtain Green's tensor field in the frequency set, $E_{i,k}^{J,(l)}(r', r)$, $E_{i,k}^{M,(l)}(r', r)$, $H_{i,k}^{J,(l)}(r', r)$, $H_{i,k}^{M,(l)}(r', r)$, $r' \in \Omega_k$, $r \in \Omega_k \cup D_S$.

(5) For the Ω_k , for $i = 1, 2, \cdots, M$, we calculate the GL modeling data at data site DS and in the frequency set,

$$\left(\left[\begin{array}{c} E(r) \\ H(r) \end{array} \right]_{i,k}^{(l)} \right) |_{r \in DS} = \left(\left[\begin{array}{c} E(r) \\ H(r) \end{array} \right]_{i_{(k-1)},k-1}^{(l)} \right) |_{r \in DS} + \int_{\Omega_k} G_{E,H,i,k}^{J,M,(l)}(r', r) \left([D]_{i,k} - [D]_{i_{(k)},k}^{(l-1)} \right) \left[\begin{array}{c} E(r) \\ H(r) \end{array} \right]_{i_{(k-1)},k-1}^{(l)} dr', \quad (5)$$

where the Green's tensor function

$$G_{E,H,i,k}^{J,M,(l)}(r', r) = \left[\begin{array}{cc} E_{i,k}^{J,(l)}(r', r) & H_{i,k}^{J,(l)}(r', r) \\ E_{i,k}^{M,(l)}(r', r) & H_{i,k}^{M,(l)}(r', r) \end{array} \right].$$

Let

$$EH_{i,k}^{(l)} = \left(\left[\begin{array}{c} E(r) \\ H(r) \end{array} \right]_{i,k}^{(l)} \right) |_{r \in DS}, \quad (6)$$

$$EH_{measured} = \left[\begin{array}{c} E(r) \\ H(r) \end{array} \right] |_{r \in DS, measured}, \quad (7)$$

$$dEH_{i,k}^{(l)} = EH_{i,k}^{(l)} - EH_{measured}. \quad (8)$$

In the i circle, $i = 1, 2, \cdots, M$, there is $i_0 = i_{(k)}$ such that

$$\begin{aligned} e_k^{(l)} &= \left\| W_m \left(dEH_{i_{(k)},k}^{(l)} \right) \right\|_{Lm} + \alpha_m \left\| D_{i_{(k)},k} - \bar{D} \right\|_{L_2} \\ &= \min_{i=1,2,\cdots M} \left(\left\| W_m \left(dEH_{i,k}^{(l)} \right) \right\|_{Lm} + \alpha_m \left\| D_{i,k} - \bar{D} \right\|_{L_2} \right). \end{aligned} \quad (9)$$

In the i circle the M times data modeling (4) and (5) are very fast.

(6) For the fixed $i_0 = i_{(k)}$, to perform the GL modeling (2.2) and (2.3). Using $[D]_{i_{(k)},k} = [D]_{i_{(k)},k}^{(l)}$ to replace $[D]_{i_{(k)},k}^{(l-1)}$, the k circle is going to next $(k+1)^{th}$ step.

(7) After finishing the k circle, $k = 1, 2, \cdots, N$, we get

$$e^{(l)} = \max_{k=1,2,\cdots,N} e_k^{(l)}. \quad (10)$$

(8) If the $e^{(l)}$ is less than ε or $l > l_s$ for the given ε and l_s, the GLMC inversion is stopped, otherwise the l circle is going to (3) to run the $(l+1)^{th}$ step.

In the formula (9), W_m is a weight matrix, L_m is a norm, α_m is a regularizing parameter, and $\bar{D}$ is the average EM parameter by experiment and by historical records. The metro parameters W_m and L_m are chosen by the frequency band, data configuration, and experiment. In this paper, we use $W_m = I$, L_2, $\alpha_m = 1.963$. For the airborne exploration and ground penetration radar, we use the integral geometry and travel time in the wave front surface to determine the W_m and L_m [10]. The α_m is chosen by the time interval and frequency band [11].

3.2. The GLMC Stochastic (GLMCS) EM inversion

Based on the GLMC EM inversion, we describe the GLMCS EM inversion as follows. Suppose that we know the prior probability distribution function on the EM data $p(EH|D)$ and on the parameters $p(D)$.

The steps (1)–(4) of the GLMCS EM inversion are as same as these GLMC's steps. The step (5) of GLMCS EM inversion should be changed as follows,

(5) Using the formulas (6), (7), and (8) in the step (5) of the GLMC inversion iteration and Bayes' formula in [13], we calculate the

$$p(D_{i,k}^{(l)}|dEH_{i,k}^{(l)}) = \frac{p(dEH_{i,k}^{(l)}|D_{i,k}^{(l)})p\left(D_{i,k}^{(l)}\right)}{\sum_{i=1}^{M} p(dEH_{i,k}^{(l)}|D_{i,k}^{(l)})p\left(D_{i,k}^{(l)}\right)}. \tag{11}$$

For each l and k, we finish the i circle in the step (5) of GLMC, $i = 1, 2, \cdots, M$, there is $i_0 = i_{(k)}$ such that

$$p(D_{i_{(k)},k}^{(l)}|dEH_{i_{(k)},k}^{(l)}) = \max_{i=1,2,\cdots M} p(D_{i,k}^{(l)}|dEH_{i,k}^{(l)}). \tag{12}$$

(6) For the fixed $i_0 = i_{(k)}$ and $D_{i_{(k)},k}^{(l)}$, the GL modeling (2.2) and (2.3) is performed quickly. Using $[D]_{i_{(k)},k} = [D]_{i_{(k)},k}^{(l)}$ to replace $[D]_{i_{(k)},k}^{(l-1)}$, the k circle is going to the next $(k+1)^{th}$ step.

(7) After finishing the k circle, $k = 1, 2, \cdots N$, we get

$$p^{(l)} = \min_{k=1,2,\cdots,N} p(D_{i_{(k)},k}^{(l)}|dEH_{i_{(k)},k}^{(l)}). \tag{13}$$

(8) if $p^{(l)}$ is large than η_s or $l > l_s$ for the given η_s and l_s, the GLMCS EM inversion is stopped, otherwise the l *circle* is going to (3) to run the $(l+1)^{th}$ step. After finishing the l circle, we take the probability average of $[D]_{i_{(k)},k}^{(l)}$

$$\begin{aligned} D_k &= A\left([D]_{i_{(k)},k}^{(l)}\right) \\ &= \sum_{l=1}^{l_n} p(D_{i_{(k)},k}^{(l)}|dEH_{i_{(k)},k}^{(l)})D_{i_{(k)},k}^{(l)}. \end{aligned} \tag{14}$$

The D_k, $k = 1, 2, \cdots N$, form the EM parameter image of the GLMCS EM inversion.

3.3. The GLMC Markov Chain EM inversion

Suppose that the prior probability distribution function on the EM data $p(EH|D)$ and parameters $p(D)$ are given. The EM data and parameters are random variable in the probability space $\Re\left(D, EH, P\left(EH|D\right), P\left(D\right)\right)$. The l iteration circle in the GLMC is the Markov chain or the skeleton Markov chain [12]. The conditional probability density function $P\left(D|EH\right)]$ is the same in the space $\Re$ for each l. After finishing the GLMCS EM inversion iteration, we take the ensemble mean of the EM parameters,

$$D_k = \frac{1}{l_n}\sum_{i=1}^{M}\sum_{l=1}^{l_n} p(D_{i,k}^{(l)}|dEH_{i,k}^{(l)})D_{i,k}. \tag{15}$$

The D_k, $k = 1, 2, \cdots N$ form the EM parameter image of the GLMC Markov Chain EM inversion.

3.4. The SAGILD EM inversion

The GLMCS and GLMC Markov Chain EM inversion methods can be used to recover the ensemble mean of the EM parameter. For recovering the other statistical moments, we briefly describe the SAGILD EM inversion as follows [4].

From the maximum probability (12), we have

$$-\log\left(P\left(D\right|\left(EH\right)|_{DS}\right)) = \min, \tag{16}$$

where D is EM parameter, $EH = (E, H)$, and

$$[D] = \begin{bmatrix} (\sigma + i\omega\varepsilon)\, I & 0 \\ 0 & -i\omega\mu I \end{bmatrix}.$$

If the random EM parameter D and data EH have the Gaussian distribution, the equation (16) is equivalent to the regularizing inversion formula [4],

$$\|(EH)\,|_{measured\ \text{on DS}} - (EH)\,|_{model\ \text{on DS}}\|_{Lm} + \alpha \left\|D - \bar{D}\right\|_{L_2} = \min, \tag{17}$$

where α is the regularizing parameter. In the paper [4], we coupled the EM boundary strip first type integral equation in the boundary zone and Galerkin parameter equation in the internal domain to construct SAGILD inversion for the EM parameter statistics moments.

4. ADVANTAGES OF THE STOCHASTIC GLMC INVERSION

The conventional MCMC inversion requests repeatedly to solve many forward modeling when the parameters are changed in each sub domain. The finite difference [7] or finite element forward methods in the MCMC have same heaviness cost for repeatedly sampling parameter in each sub domain. The high cost of computational time is a main and unaccepted disadvantage of the MCMC. The inaccurate absorption conditions on the artificial boundary make noises to downgrade the resolution in the MCMC inversion. The GLMC is a metro path and bridge to connect and compromise the deterministic and stochastic inversions. From (1) – (8) in the Section 3.1, the GL modeling and GLMC inversion are simultaneous processes. There is no big matrix to be solved in the GL modeling and GLMC inversion that effectively reduces the computational time cost. The GLMC EM inversion can be the deterministic inversion, it is adjusted by the suitable weight, norm, and regularizing factors which have likelihood function. Moreover, the GLMC EM inversion is easy to switch to the stochastic GLMCS and GLMC Markov Chain inversion. The GL EM modeling method [1][2] is very fast to calculate model data when the EM parameters are changed in each sub domain, and greatly increase the speed of GLMC inversions. The GL method has no artificial boundary and no inaccurate absorption boundary condition that reduces the modeling noise for the GLMC inversion. If the EM data and the parameters have the Gaussian distribution, the maximum probability condition (12) in the GLMCS and GLMC Markov Chain inversion is equivalent to threshold condition (9) in the GLMC inversion. The GLMC EM inversion method can be useful for material sciences, nanometer materials, surface physics, rough surface, medical sciences, biology, geophysical exploration and finance.

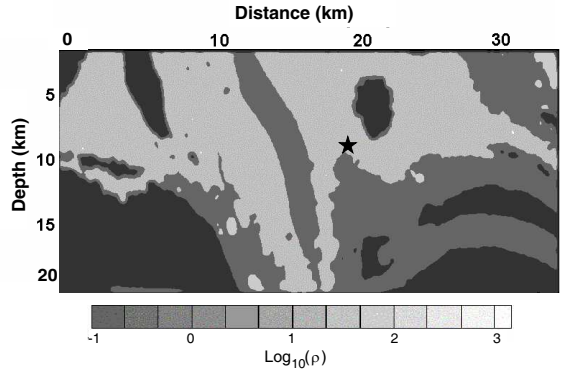


Figure 1: The resistivity imaging of the MT data is reconstructed by AGILD and GLMC MT inversion. Not only the distinct low resistivity anomaly beneath the hypocenter (the solid star) of the Chi-Chi earthquake is shown, but also the CLP fault is clearly displayed in location (at about distance 12 km and dips to the right) and size.

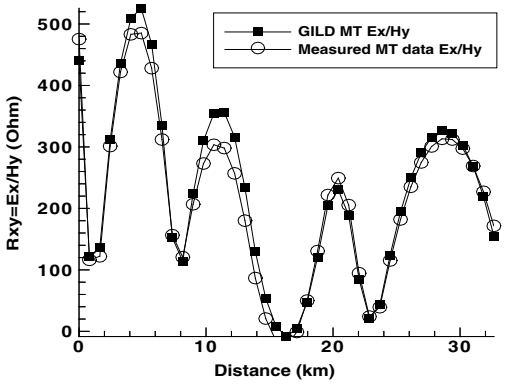


Figure 2: Misfit between measured MT data E_x/H_y and modeling data E_x/H_y by GLMC MT modeling and inversion. The square root residual of the misfit is 0.3×10^{-1}.

5. APPLICATIONS

We use the GLMC EM inversion and SAGILD MT inversion for the MT imaging in geophysical exploration for the Chi Chi Earthquake in Taiwan. First, we use the AGILD MT inversion for the geological geometry and conductivity imaging in the large zone. Second, we use the GLMC CSMT EM inversion and seismic joint inversion for local conductivity imaging in the crack. We obtained reasonable resolution of the resistivity image. The conductivity imaging and data misfit are shown

in the Figures 1 and 2. The patent right of the algorithms and software of the GLMC EM inversion method are reserved by authors in the GLGEO.

6. CONCLUSION

The synthetic and field data simulation tests show that GLMC inversion and its Markov inversion are fast, stable and have reasonable resolution. It will be used for seismic, hydrological, EM coupled inversion. Also it is useful for finance and biological stochastic modeling and inversion.

ACKNOWLEDGMENT

Authors thank Professor Kong and Professor Michael Oristaglio for their encouragements.

REFERENCES

1. Xie, G., J. Li, and F. Xie, "New global and local electromagnetic field (GL) modeling and inversion," *Proc. of PIERS 2005*, 68, Hangzhou, 2005.
2. Xie, G., F. Xie, L. Xie, and J. Li, "GL method and its advantages for resolving historical difficulties," *Progress In Electromagnetics Research*, PIER 63, 141–152, and *Proc. of PIERS 2006*, 306, Cambridge, 2006.
3. Xie, G. and J. Li, "New parallel SGILD modeling and inversion," *Physics D*, Vol. 133, 477–487, 1999.
4. Li, J., G. Xie, and F. Xie, "New stochastic AGLID EM modeling and inversion," *Proc. of PIERS 2006*, 490, Cambridge, 2006.
5. Xie, G., J. Li, L. Xie, "A new GL Metro Carlo inversion for electromagnetic and seismic image.," *GLGEO technology patent report GLREPORT082006*, 2006.
6. Gilks, W. R., S. Richardson, and D. J. Spiegelhalter, *Markov Chain Monte Carlo in Practice*, Chapman and Hall, London, 1996.
7. Spichak, V., M. Menvielle, M. Roussignol, "Three dimensional inversion of MT fields using Rayesian statistics," *Three Dimensional Electromagnetics,* Edited by M. Oristaglio and B. Spies, GD, 406–417, 1999.
8. Wang, L.-F., J. A. Kong, K. H. Ding, T. Le Toan, F. Ribbes, and N. Floury, "Electromagnetic Scattering Model for Rice Canopy Based on Monte Carlo Simulation," *Progress In Electromagnetics Research*, PIER 52, 153–171, 2005.
9. Binder, K., *The Monte Carlo Method in Condensed Matter Physics*, topic in applied physics, Vol. 71, 1992.
10. Xie, G., J. Li, "Nonlinear integral equation of coefficient inversion of acoustic wave equation and iteration," *Science in China*, Vol. 32, No. 5, 513–523, 1989.
11. Xie, G., J. Li, M. Majer, D. Zuo, and M. Oristaglio, "3D electromagnetic modeling and nonlinear inversion", *Geophysics*, Vol. 65, No. 3, 804–822, 2000.
12. Hou, Z., *Markov skeleton processes*, Hunan Science and Technology Press, 2000.
13. Meyer, P. L., *Introductory probability and statistical applications* , Addison–Wesley Publishing Company, 1970.

A 3D GL EM Modeling and Inversion for Forest Exploration and Felling

Jianhua Li, **Ganquan Xie**, **Lee Xie**, **and Feng Xie**
GL Geophysical Laboratory, USA

Abstract— In this paper, we proposed a 3D GL electromagnetic (EM) modeling and inversion for the forest Exploration and Felling (GLEMFEF). We consider the inhomogeneous forest domain which is embedded in the earth and air two layered media. The forest domain includes inhomogeneous EM parameters. The EM field differential integral equation and integral equation are used to formulate the EM field in the forest domain. The inhomogeneous forest domain is divided into the several small sub domains. The forest EM parameter in the sub domain is isotropy or anisotropy variable. The Global field is updated by the Local EM scattering field in the sub domain. When the Global forest field propagates through the all sub domains, the GL forest EM field is obtained. The GL forest modeling is very fast to calculate EM field when the EM parameters are updated in the each sub domain. This merit accelerates the GL Metro Carlo inversion for the forest imaging. The 3D GL forest modeling and inversion can be useful for forest exploration, forest environment, animal, Chinese medical tree and grass, forest fire fighting, felling, and mine exploration in the ground under forest.

1. INTRODUCTION

The electromagnetic modeling and inversion method has important applications in the forest science and engineering. In this paper, we proposed the GL [1] modeling and inversion for the forest exploration, fire fighting, and felling. The EM differential integral equation [2] and EM integral equation [1] are used to formulate the EM field in the forest domain. The inhomogeneous forest domain is divided into the several small sub domains. The forest EM parameter in the sub domain is isotropy or anisotropy variable. The Global field is updated by the Local scattering field in the sub domain. When the Global forest field propagates through the all sub domains, the GL forest EM field is obtained. The GL forest modeling is very fast to calculate EM field when the EM parameters are updated in the each sub domain. This merit accelerates the GL Metro Carlo inversion [3] for the forest imaging. The 3D GL forest modeling and inversion can be useful for forest exploration, forest environment, forest fire fighting, felling, and mine exploration in the ground under forest.

There are research works on the EM modeling for forest. The paper [4] described analysis of electromagnetic wave propagation in forest environment along multiple paths. The paper [5] described a full wave three dimensional analysis of forest remote sensing using the electric integral equation in the VHF frequency band. In this paper, the GLEMFEF modeling and inversion method for the forest exploration and felling is different from the above papers.

The plan of this paper is as follows. The introduction has been described in Section 1. The electromagnetic field differential integral equation and EM integral equation are described in Section 2. In Section 3, we propose the GL modeling for EM field in the forest. The GL inversion for the forest imaging is presented in Section 4. In Section 5, we describe forest exploration, fire fighting, and environment. The forest felling is described in Section 6. The simulations are described in Section 7. We conclude this paper in Section 8.

2. THE ELECTROMAGNETIC DIFFERENTIAL INTEGRAL EQUATION AND INTEGRAL EQUATION

In the is section, we describe the electromagnetic field differential integral equation and integral equation which are proposed in the paper [1, 2]. We use these equations to formulate the EM wave propagation in the forest, air, and ground inhomogeneous domain.

2.1. The Magnetic Field Differential Integral Equation

The magnetic field differential integral equation is as follows,

$$\mathbf{H}(r) = \mathbf{H}_b(r) + \int_{\Omega} D\left(\sigma, \varepsilon; \sigma_b, \varepsilon_b\right) \mathbf{E}_b^M\left(r', r\right) \cdot \nabla \times \mathbf{H}(r') dr', \tag{1}$$

where

$$D\left(\sigma,\varepsilon;\sigma_b,\varepsilon_b\right)=\frac{\left(\sigma(r)+i\omega\varepsilon(r)\right)-\left(\sigma_b(r)+i\omega\varepsilon_b(r)\right)}{\sigma(r)+i\omega\varepsilon(r)},$$

$\mathbf{H}(r)$ is the magnetic field, $\mathbf{H}_b(r)$ is the background magnetic field with air and ground two layered background EM media, σ is the electric conductivity, ε is the dielectric parameter, σ_b is the background electric conductivity, ε_b is the background dielectric parameter, r is the space variable vector, $r=(x,y,z)$, r' is space integrand variable, $r'=(x',y',z')$,

$$\mathbf{E}_b^M(r',r)=\frac{1}{\sigma_b(r)+i\omega\varepsilon_b(r)}\nabla\times\mathbf{H}_b^M(r',r),$$

$\mathbf{E}_b^M(r',r)$ is the electric Green's field tensor excited by the magnetic dipole source,
$\mathbf{H}_b^M(r',r)$ is the magnetic Green's field tensor excited by the magnetic dipole source,
$\mathbf{E}_b^M(r',r)$ is the weaker singular integrative kernel and easy to be discretized [2].
If the μ is the continuous function or constant, the $\mathbf{H}_b^M(r',r)$ and $\mathbf{H}(r)$ are continuous.

2.2. The Electric Field Differential Integral Equation

The electric field differential integral equation is as follows [2],

$$\mathbf{E}=\mathbf{E}_b+\int\limits_{\Omega}\left(\frac{\mu-\mu_b}{\mu}\right)\mathbf{H}_b^J(r',r)_b\nabla\times\mathbf{E}(r')dr', \tag{2}$$

where

$$\mathbf{H}_b^J(r',r)=\frac{1}{\mu_b}\nabla\times\mathbf{E}_b^J(r',r),$$

μ is the magnetic permeability, μ_b is the background magnetic permeability.
$\mathbf{E}_b^J(r',r)$ is the electric Green's field tensor excited by the electric dipole source,
$\mathbf{H}_b^J(r',r)$ is the magnetic Green's field tensor excited by the electric dipole source,
$\mathbf{H}_b^J(r',r)$ is the weaker singular integrative kernel and easy to be discretized [2].
If the σ and ε are the continuous function or constant, the $\mathbf{E}_b^J(r',r)$ and $\mathbf{E}(r)$ are continuous.

2.3. The Electromagnetic Field Integral Equation

The electromagnetic field integral equation is as follows [1]

$$\begin{bmatrix}\mathbf{E}(r)\\ \mathbf{H}(r)\end{bmatrix}=\begin{bmatrix}\mathbf{E}_b\left(r_s\right)\\ \mathbf{H}_b(r)\end{bmatrix}+\int\limits_{\Omega}\begin{bmatrix}\mathbf{E}_b^J(r',r) & \mathbf{H}_b^J(r',r)\\ \mathbf{E}_b^M(r',r) & \mathbf{H}_b^M(r',r)\end{bmatrix}\left[D\left(\sigma,\varepsilon;\sigma_b,\varepsilon_b\right)\right]\begin{bmatrix}\mathbf{E}(r')\\ \mathbf{H}(r')\end{bmatrix}dr', \tag{3}$$

$$\begin{bmatrix}\mathbf{E}(r)\\ \mathbf{H}(r)\end{bmatrix}=\begin{bmatrix}\mathbf{E}_b(r)\\ \mathbf{H}_b(r)\end{bmatrix}+\int\limits_{\Omega}\begin{bmatrix}\mathbf{E}^J(r',r) & \mathbf{H}^J(r',r)\\ \mathbf{E}^M(r',r) & \mathbf{H}^M(r',r)\end{bmatrix}\left[D\left(\sigma,\varepsilon;\sigma_b,\varepsilon_b\right)\right]\begin{bmatrix}\mathbf{E}_b(r')\\ \mathbf{H}_b(r')\end{bmatrix}dr'. \tag{4}$$

where $D\left(\sigma,\varepsilon;\sigma_b,\varepsilon_b\right)$,

$$\left[D\left(\sigma,\varepsilon;\sigma_b,\varepsilon_b\right)\right]=\begin{bmatrix}\left(\sigma(r)+i\omega\varepsilon(r)\right)-\left(\sigma_b(r)+i\omega\varepsilon_b(r)\right)I & 0\\ 0 & i\omega\left(\mu_b(r)-\mu(r)\right)I\end{bmatrix},$$

is the 6×6 EM material parameter variation matrix for the isotropy materials, for anisotropy materials the $[D]$ will be 6×6 full matrix, $\mathbf{E}(r)$ is the electric field, $\mathbf{H}(r)$ is the magnetic field, $\mathbf{E}_b(r)$ is incident electric field in the background medium, $\mathbf{H}_b(r)$ is incident magnetic field in the background medium, $\mathbf{E}_b^M(r',r),\ldots,\mathbf{H}_b^M(r',r)$ is electric or magnetic background field Green tensor exciting by the electric or magnetic dipole source respectively. The integral equations (3) and (4) are dual system each other.

3. GL EM MODELING FOR EM FIELD IN THE FOREST

In this section, we propose the Global and Local EM field modeling for the forest science and engineering. In short, it is called the GL forest EM modeling.

3.1. The GL EM Modeling I Based on the Magnetic Field Differential Integral Equation

The GL EM modeling based on the magnetic field differential integral equation is suitable for forest domain with inhomogeneous conductivity and dielectric parameters.

3.1.1.

The forest domain Ω is divided into the sub domains Ω_k which are included in the trunk, skin, branch, and foliage domain respectively, such that $\Omega = \bigcup_{k=1}^{N} \Omega_k$.

3.1.2.

In each Ω_k, we solve the magnetic field differential integral equation (1). By dual curl operation, the equation systems are reduced into a 3×3 matrix equations. By solving the 3×3 equations, we obtain Green tensor field $\mathbf{H}_k^M$ and $\mathbf{E}_k^M$.

3.1.3.

We improved the Global Magnetic field $\mathbf{H}_k(r)$ by the Local scattering filed,

$$\mathbf{H}_k(r) = \mathbf{H}_{k-1}(r) + \int_{\Omega_k} D\left(\sigma_{k-1}, \varepsilon_{k-1}; \sigma_k, \varepsilon_k\right) \mathbf{E}_k^M\left(r', r\right) \cdot \nabla \times \mathbf{H}_{k-1}(r') dr', \tag{5}$$

where we consider isotropic EM model,

$$D\left(\sigma_{k-1}, \varepsilon_{k-1}; \sigma_k, \varepsilon_k\right) = \frac{\left(\sigma_k(r) + i\omega\varepsilon_k(r)\right) - \left(\sigma_{k-1}(r) + i\omega\varepsilon_{k-1}(r)\right)}{\sigma_{k-1}(r) + i\omega\varepsilon_{k-1}(r)}$$

$k = 1, 2, \ldots, N$, successively. The $\mathbf{H}_N(r)$ and $\mathbf{E}_N(r)$, $\mathbf{E}_N(r) = (\nabla \times \mathbf{H}_N(r))/(\sigma + i\omega\varepsilon)$, are GL solution of the 3D GL EM modeling.

3.2. The GL EM Modeling II Based on the Electric Field Differential Integral Equation

The GL EM modeling based on the electric field differential integral equation [1] is suitable for forest domain with inhomogeneous magnetic permeability parameters.

3.2.1.

The step is the same as 3.1.1.

3.2.2.

In each Ω_k, $k = 1, 2, \ldots, n$, we solve the electric field differential integral equation (2) to find $\mathbf{H}_k^J(r', r)$ successively. By the dual curl operation, only 3×3 matrix equations is solved.

3.2.3.

We improved the Global EM field $\mathbf{E}_k(r)$ by the Local scattering filed

$$\mathbf{E}_k(r) = \mathbf{E}_{k-1}(r) + \int_{\Omega_k} \frac{\mu_k - \mu_b}{\mu_{k-1}} \mathbf{H}_k^J\left(r', r\right) \cdot \nabla \times \mathbf{E}_{k-1}\left(r'\right) dr', \tag{6}$$

$k = 1, 2, \ldots, N$, successively. $\mathbf{E}_N(r)$ is GL electric field solution of (2).

3.3. The GL EM Modeling III Based on the Electromagnetic Field Integral Equation

The GL EM modeling based on the electromagnetic field integral equation [1] is suitable for forest domain with inhomogeneous electric conductivity, dielectric, and magnetic permeability parameters. The paper [1] detailed described the GL EM modeling.

3.3.1.

The step is same as the 3.1.1.

3.3.2.

In each Ω_k, we solve the EM Green tensor integral equation system based on the equations (3) and (4). By dual curl operation, the equation system is reduced into a 6×6 matrix equations. By solving the 6×6 equations, we obtain Green tensor field $\mathbf{E}_k^J$ and $\mathbf{H}_k^M$.

3.3.3.

We improved the Global EM field $[\mathbf{E}_k(r),\ \mathbf{H}_k(r)]$ by the Local scattering filed

$$\begin{bmatrix}\mathbf{E}(r)\\ \mathbf{H}(r)\end{bmatrix}_k = \begin{bmatrix}\mathbf{E}(r)\\ \mathbf{H}(r)\end{bmatrix}_{k-1} + \int\limits_{\Omega_k} \begin{bmatrix}\mathbf{E}_k^J(r',r) & \mathbf{H}_k^J(r',r)\\ \mathbf{E}_k^M(r',r) & \mathbf{H}_k^M(r',r)\end{bmatrix} [D] \begin{bmatrix}\mathbf{E}(r')\\ \mathbf{H}(r')\end{bmatrix}_{k-1} dr', \tag{7}$$

$k = 1, 2, \ldots, N$, successively. The $[\mathbf{E}_N(r),\ \mathbf{H}_N(r)]$ is GL solution of the EM integral equation (3) and (4).

The 3D GL modeling is also available for EM wave field in the anisotropic forest domain.

4. THE GL METRO CARLO INVERSION FOR FOREST IMAGING

We propose the GL Metro Carlo EM inversion [3] for forest imaging.

4.1. The GL Metro Carlo Inversion I for Forest Imaging

The GL Metro Carlo Inversion I for Forest Imaging with respect to GL EM modeling I is as follows.

4.1.1.

The forest domain Ω is divided into the sub domains Ω_k which are included in the trunk, branch, and foliage domain respectively, such that $\Omega = \bigcup\limits_{k=1}^{N} \Omega_k$. In each Ω_k, we define electric conductivity and dielectric parameter set, $\{\sigma\} = \{\sigma_{i,k}\},\ \{\varepsilon\} = \{\varepsilon_{i.k}\},\ k = 1, 2, \ldots, N,\ i = 1, 2, \ldots, M$. Suppose that for each Ω_k , $\{\{\sigma_{i,k}\},\ i = 1, \ldots, M_1\}$ are conductivities in the ground, $\{\{\sigma_{i,k}\},\ i = M_1+1, \ldots, M_2\}$ are conductivities in the tree trunk, $\{\{\sigma_{i,k}\},\ i = M_2+1, \ldots, M_3\}$ are conductivities in the tree branch, $\{\{\sigma_{i,k}\},\ i = M_3+1, \ldots, M_4\}$ are conductivities in the tree foliage, $\{\{\sigma_{i,k}\},\ i = M_4+1, \ldots, M_5\}$ are conductivities in the dry air and moisture air. Suppose that the probability $p\{\sigma_{i,k}\}$ and $p\{\varepsilon_{i,k}\}$ are given.

4.1.2.

Suppose for lth iteration, $\{\{\sigma_{i(k),k}^{(l)}\},\ \{\varepsilon_{i(k),k}^{(l)}\},\ k = 1, \ldots, N\}$ is obtained by the previous lth iteration. For each k, $k = 1, 2, \ldots, N$, and we do i-circle, $i = 1, 2, \ldots, M$, we calculate $[\mathbf{H}(r)]_{i,k}^{(l+1)}|_{r\in DS}$ by the following differential integral equation,

$$[\mathbf{H}(r)]_{i,k}^{(l+1)}|_{r\in DS} = [\mathbf{H}(r)]_{i(k-1),k-1}^{(l+1)}|_{r\in DS} + \int\limits_{\Omega_k} D_{i,k,k-1}^{l+1} \mathbf{E}_{i,k}^{M,(l+1)}\left(r', r\right) \cdot \nabla \times [\mathbf{H}(r')]_{i(k-1),k-1}^{(l+1)} dr', \tag{8}$$

where

$$D_{i,k,k-1}^{l+1} = \frac{(\sigma_{i,k}(r) + i\omega\varepsilon_{i,k}(r)) - \left(\sigma_{i(k-1),k-1}^{(l+1)}(r) + i\omega\varepsilon_{i(k-1),k-1}^{(l+1)}(r)\right)}{\sigma_{i(k-1),k-1}^{(l+1)}(r) + i\omega\varepsilon_{i(k-1),k-1}^{(l+1)}(r)},$$

$\mathbf{E}_{i,k}^{M,(l+1)}(r', r)$ is obtained by using the dual curl operation and solving the 3×3 discrete magnetic field differential integral equation (1).

4.1.3.

Define

$$d\mathbf{H}_{i,k}^{(l+1)} = [\mathbf{H}(r)]_{i,k}^{(l+1)}|_{r\in DS} - [\mathbf{H}(r)]|_{r\in DS, measured},$$

in the i-circle, $i = 1, 2, \ldots, M$, there is $i_0 = i_{(k)}$, such that

$$\begin{aligned} e_k^{(l+1)} &= \left\|\mathbf{W}_F d\mathbf{H}_{i(k),k}^{(l+1)}\right\| + \alpha_F \left\|\left(\sigma_{i(k),k}(r) + i\omega\varepsilon_{i(k),k}(r)\right) - (\bar{\sigma} + i\omega\bar{\varepsilon})\right\| \\ &= \min_{i=1,2,\ldots,M} \left\|\mathbf{W}_F d\mathbf{H}_{i,k}^{(l+1)}\right\| + \alpha_F \left\|(\sigma_{i,k}(r) + i\omega\varepsilon_{i,k}(r)) - (\bar{\sigma} + i\omega\bar{\varepsilon})\right\|, \end{aligned} \tag{9}$$

where weight matrix $\mathbf{W}_F$ and regularizing parameter α_F are depended on the local forest EM property and data configuration. In the i-circle, the GL EM modeling is very fast. After finishing the i-circle, we obtain

$$\sigma_{i(k),k}^{(l+1)}(r) + i\omega\varepsilon_{i(k),k}^{(l+1)}(r) = \sigma_{i_0,k}(r) + i\omega\varepsilon_{i_0,k}(r). \tag{10}$$

4.1.4.

After finishing k-circle, we obtain $\{\{\sigma^{(l+1)}{}_{i(k),k}\},\ \{\varepsilon^{(l+1)}{}_{i(k),k}\},\ k = 1, \ldots, N\}$, also we calculate

$$e^{(l+1)} = \max_{i=1,2,\ldots,M} e_k^{(l+1)}. \tag{11}$$

4.1.5.

If the $e^{(l+1)}$ less than a given small number η or l larger than a given large number L, then the GL Metro Carlo EM inversion is stopped. Otherwise the *l-circle* iteration is continuous [1].

4.2. The GL Metro Carlo Inversion II and III for Forest Imaging

The GL Metro Carlo Inversion II and III for forest imaging are similar with the GL EM Inversion I which is described in the papers [1, 3].

5. THE FOREST EXPLORATION

In geophysical exploration, there are seismic, acoustic and electromagnetic methods. The electromagnetic method plays very important main role in the forest exploration. It can be used to explore forest location, size, type, rain forest, forest weather environment, and forest fire etc.

5.1. Data Configuration

The one source to multiple receivers, multiple source to multiple receivers airborne VHF EM data configuration are employed for the forest exploration. The GL EM modeling and inversion software can be used to interpret these data to detect forest location, size, type, rain forest, forest weather environment, and forest fire etc.

5.2. Forest Fire Exploration

GL EM and thermal coupled modeling and inversion [6] can be used for rain forest, forest weather, and forest fire exploration.

6. THE FOREST FELLING

The felling is very important forest engineering. EM, acoustic, controlling seismic wave can be used to make the high resolution imaging for local forest and trees group imaging that is very important for felling safe and speed. It is necessary for automatic remote robot forest felling.

6.1. High Resolution Forest Imaging

The microwave radar (MRD) or infrared radar (IRD) multiple sources to multiple receivers data is suitable for the high resolution forest imaging. The GL EM modeling and inversion can be used for deleting the branch, skin and foliage of the individual tree or tree group. The GL microwave near field imaging can be used for felling, and to explore animals, Chinese medical tree and grass in the forest, and to find mine in the groud under forest.

6.2. To Detect the Tree Type Using the Tree Skin Imaging

The microwave radar high frequency near field data and GL EM inversion can make high resolution tree skin imaging. The tree skin rough surface EM imaging play important role to detect the tree type.

6.3. To Control Fallen Direction

In the felling, the controlling fallen direction is very important. The vertical felling working performance is prohibited. The horizontal felling performance is working. The transverse hill fallen is very dangerous and to block the wood transportion along the hill. The transverse hill fallen must be avoided. The up hill fallen and down hill fallen are two important felling fallen directions. The GL EM inversion and MRD data imaging can be useful to control trees fallen in that directions.

6.4. GL Dynamic Elastic Modeling and Inversion

After the local forest high resolution imaging is obtained by using the GL EM inversion and MRD data, we can use our GL dynamic elastic modeling and inversion [6] to simulate the fallen tree's slip motion along down hill in the following cases. The case I is to do simulation for traditional cutting in the bottom of the tree. In the case II, the bottom and some branch are cutting to control the tree to be fallen down hill and to be controlled slipping speed.

7. SIMULATIONS

7.1. The Simulation of the GL EM Modeling for Forest Area

We design a 3D forest domain which is embedded into the air and ground two layered infinite domain. The height of the forest is 18 m in the z direction. The horizontal distance in x and y direction is 100 m respectively. The electric conductivity in the ground, $\sigma_g = 0.1\,\mathrm{S/M}$, the relative dielectric $\varepsilon_g = 45$. The electric conductivity in the trunk domain, $\sigma_t = 0.2 \times 10^{-3}\,\mathrm{S/M}$, the relative dielectric $\varepsilon_t = 60$. The electric conductivity in the foliage domain, $\sigma_f = 0.15 \times 10^{-3}\,\mathrm{S/M}$, the relative dielectric $\varepsilon_f = 75$. The vertical magnetic dipole source is located in the section $y = 0$, $x = -32\,\mathrm{m}$, $z = 58\,\mathrm{m}$. The Figure 1 shows that the transverse magnetic wave H_y propagates at 0.012 µs which shows the source location. The Figure 2 shows that magnetic wave H_y propagates

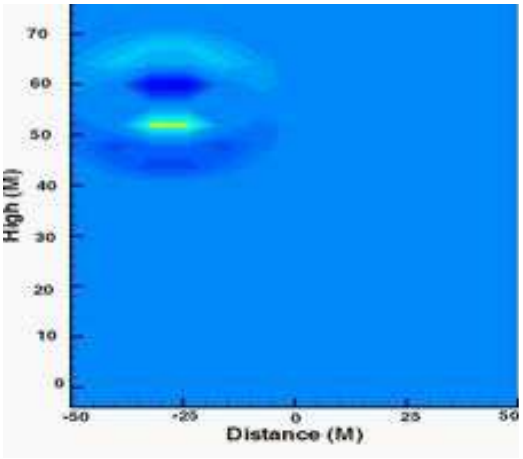


Figure 1: The magnetic wave H_y at 0.012 µs.

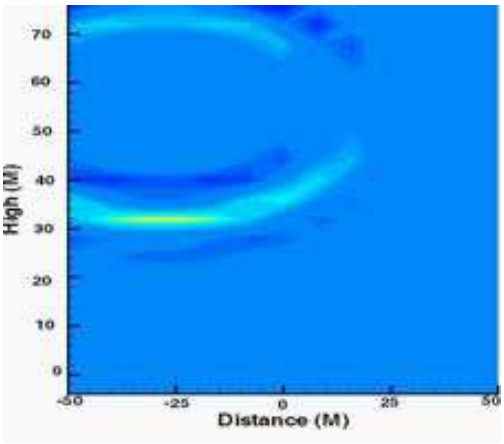


Figure 2: The magnetic wave H_y at 0.072 µs.

at 0.072 µs. The Figure 3 shows that magnetic wave H_y propagates at 0.108 µs. The Figure 4 shows that magnetic wave H_y propagates at 0.144 µs. The H_y interacts with foliage and results the scattering wave from the foliage. The Figure 5 shows that magnetic wave H_y propagates at 0.192 µs. The H_y interacts with the trunk and foliage domains and results the scattering from the trunk and foliage domains. The Figure 6 shows that magnetic wave H_y propagates at 0.252 µs which interacts with ground surface, the trunk and foliage domains, and results the scattering

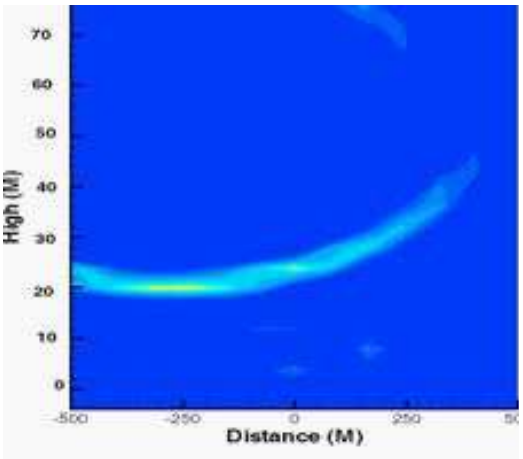


Figure 3: The magnetic wave H_y at 0.108 µs.

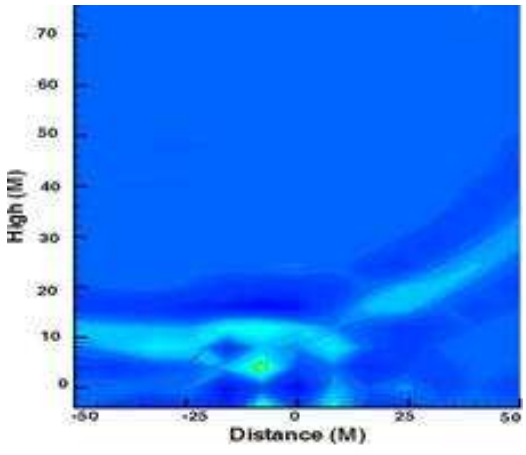


Figure 4: The magnetic wave H_y at 0.144 µs.

from the earth surface, trunk and foliage domains. The Figure 7 shows that magnetic wave H_y propagates at 0.312 µs which interacts with under ground, surface, trunk and foliage domains and results the scattering from the under ground, surface, trunk and foliage domains. The scattering wave has been propagates back to the source high level which is shown in Figure 8. By observing the Figure 1 to Figure 8, the magnetic wave H_y accurately propagates in the air, foliage, branch, trunk,

surface, and under ground. There is no any error boundary refection in the wave propagation. The scattering wave is also accurate and has clearly propagation.

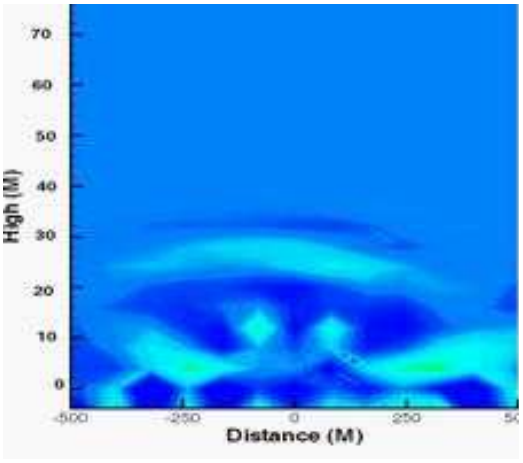


Figure 5: The magnetic wave H_y at 0.192 µs.

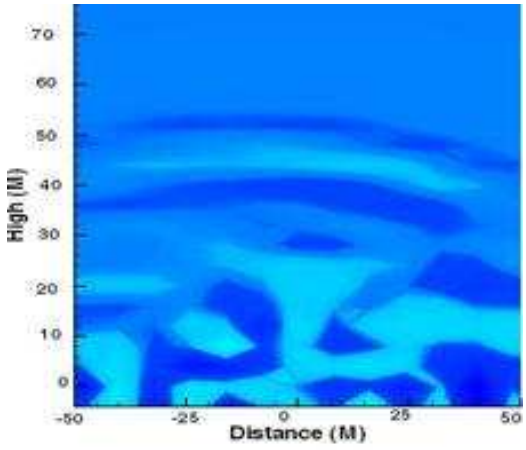


Figure 6: The magnetic wave H_y at 0.252 µs.

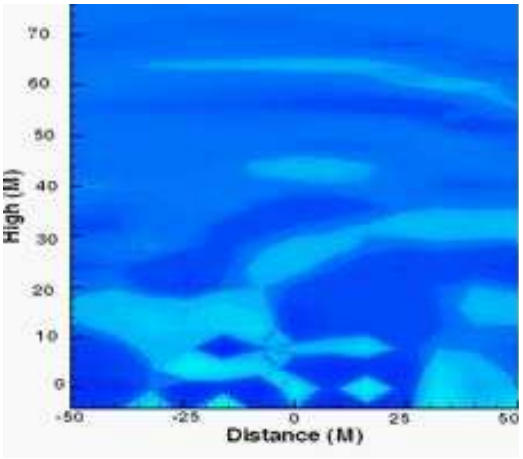


Figure 7: The magnetic wave H_y at 0.312 µs.

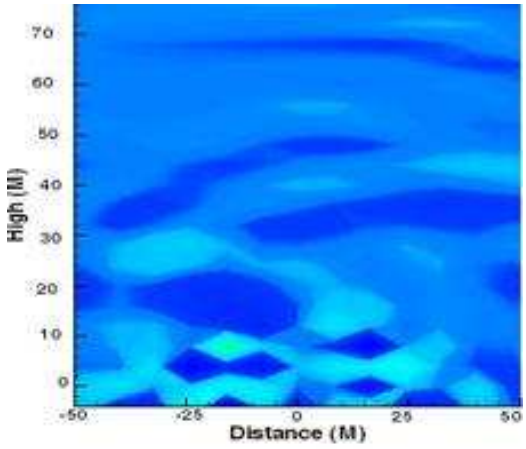


Figure 8: The magnetic wave H_y at 0.348 µs.

7.2. The Simulation of GL EM Inversion for Forest Area

An individual tree and group tree have been used for GL EM inversion simulation. We present the imaging of GL EM inversion for individual tree. The tree height is 7.3 m. Its diameter is 0.6 m. The vertical magnetic dipole source is located in the center and has height 15 m , the 4 receivers are located in 4 corners with height 1 m. We run the GLEM modeling to get data, then remove the inhomogeneous media. We use GLEMFEF inversion to recover the tree imaging. The tree model 1 is shown in Figure 9. Its imaging by GLEMFEF inversion is shown in Figure 10. The tree model 2 is shown in Figure 11. Its imaging by GLEMFEF inversion is shown in Figure 12. We can see that the GLEMFEF inversion can make high resolution for trunk and branch. We also can see the foliage imaging, although its resolution is lower than the resolution of the trunk and branch domains.

7.3. Simulation for GLEMFEF Imaging and GL Elastic Dynamic Deformation Modeling

The tree imaging is made by GLEMFEF inversion. We use the image to be input data for GL elastic dynamic deformation modeling and inversion [6]. By the simulation result, the some branch has been cut which is shown in Figure 13. Then we cut down the tree in the button. The tree has been fallen down hill in controlling direction. In particular, after slipping 0.5 m, the tree is stopped completely in the prediction position, which is shown in the Figure 14.

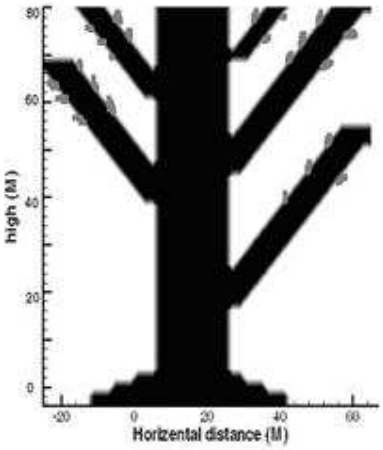


Figure 9: The tree model 1.

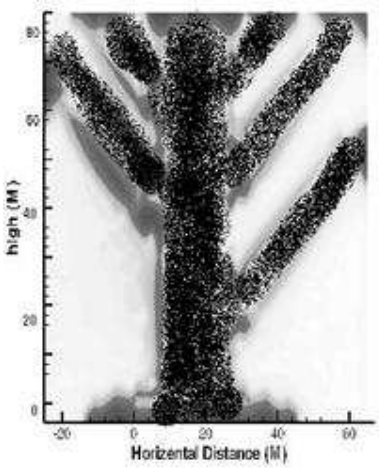


Figure 10: The tree imaging of model 1 by GLEMFEF.

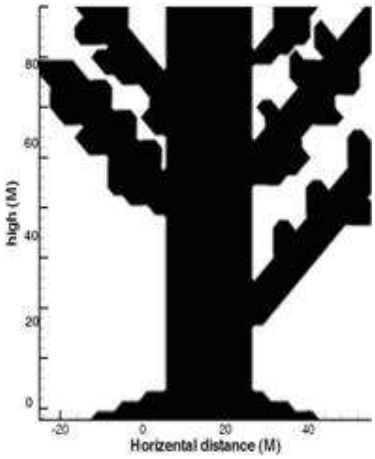


Figure 11: The tree model 2.

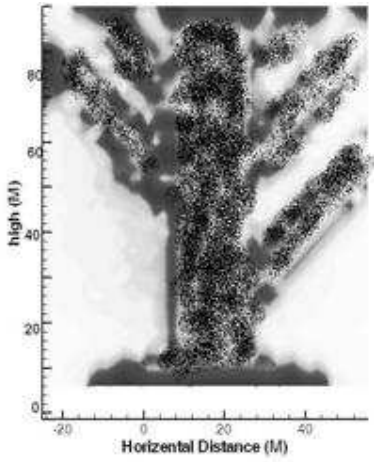


Figure 12: The tree imaging of model 1 by GLEMFEF.

Figure 13: According GLEMFEF inversion imaging and GLMD larger deformation analysis, the branches are suitable cut to control stable collapse and control down hill fallen direction.

Figure 14: The tree is stable collapse and fallen in prediction direction and stopped in given position. The error between the practice fallen position and GLEMFEF controlled direction and estimated position is only 13 cm.

8. CONCLUSION

The GL EM modeling and inversion for the forest exploration and felling are developed. The simulations show that the GL forest modeling and inversion is accurate, stable, fast and high resolution. The GLEMFEF algorithm and software can be used to make high resolution forest imaging for forest exploration and felling. The 3D GL forest modeling and inversion GLEMFEF can be useful for forest exploration, forest environment, animal, Chinese medical tree and grass detection, forest fire fighting, felling, and mine exploration in the ground under forest. The patent of the GLEMFEF algorithm and software belongs to authors of this paper.

REFERENCES

1. Xie, G., F. Xie, L. Xie, and J. Li, "GL method and its advantages for resolving historical difficulties," *Progress in Electromagnetics Research*, Vol. PIER 63, 141–152.
2. Xie, G., J. Li, E. Majer, D. Zuo, and M. Oristaglio, "3-D electromagnetic modeling and nonlinear inversion," *Geophysics*, Vol. 65, No. 3, 804–822, 2000.
3. Xie, G., J. Li, L. Xie, and F. Xie, "A GL Metro Carlo EM inversion," *Journal of Electromagnetic Waves and Applications*, Vol. 20, No. 14, 1991–2000, 2006.
4. Li, L. W., T. S. Yeo, P. S. Kooi, M. S. Leong, and J. H. Koh, "Analysis of electromagnetic wave propagation in forest environment along multiple paths," *Progress in Electromagnetics Research*, Vol. PIER 23, 137–164, 1999.
5. Angot, L., H. Roussel, and W. Tabbara, "A full wave three dimensional forest remote sensing using electromagnetic wave," *Progress in Electromagnetics Research*, Vol. PIER 38, 311–331, 2002.
6. Li, J., G. Xie, L. Xie, and F. Xie, "A GL elastic dynamical deformation analysis," GL Geophysical Laboratory Report, GLGEOP0015, 2006.

Multi-grid Method of Quasi Elastic Fluid Flow

Xuequan Li[1], **Haitao Cai**[2], **Hongwei Zhang**[2, 3], **and Yinghao Cai**[4]

[1]Hunan Normal College, Changsha, Hunan, China
[2]College of Mathmatics, Central South University, Changsha 410083, China
[3]College of Math & Computing Science, University of Science and Technology, Changsha, China
[4]Institute of Automation, Chinese Academy of Sciences, 100080, China

Abstract— For the steady state and time-dependent viscoelastic fluid flow in three dimensional domains, we put forward a V-cycle multi-grid method. At the same time, we discuss existence and error estimates of the approximate solution. The approximate stress, velocity and pressure are respectively p_k discontinuous, p_{k+1} continuous, p_k continuous.

1. INTRODUCTION

Lately, more and more attentions have been paid to multi-grid method. The multi-grid method provides an optimal algorithm for solving many problems. The multi-grid method has two main features [1]: smoothing step on the current grid and error correction on a coarser grid. The smoothing step has the effect of damping out the oscillatory part of the error. The smooth part of the error can then be accurately corrected on the coarser grid. In this paper, we study the numerical analysis of a mixed finite element approximation of a viscoelastic fluid flow obeying an Oldroyd B constitutive equation. We put forward a V-cycle multi-grid method For the steady state and time-dependent viscoelastic fluid flow in three dimensional domains. At the same time, we discuss existence and error estimates of the approximate solution. The approximate stress, velocity and pressure are p_k discontinuous, p_{k+1} continuous, p_k continuous, respectively. Using V-cycle multi-grid method, we show that the approximate problem solution's existence, uniqueness and convergence are established.

We begin by describing a model of viscoelastic fluid flow obeying an Oldroyd B constitutive equation in Section 1. In Section 2, we provide the finite discretization of this model. In Section 3, we use V-cycle multi-grid method to approximate this equation. Finally, we give the convergence analysis in Section 4.

Consider the model of Oldroyd for viscoelastic fluid flow (Problem Oldroyd)

$$\begin{aligned} \lambda\sigma_t + \sigma + \lambda(u\cdot\nabla)\sigma + \lambda g_\alpha(\sigma, \nabla u) - 2\alpha d(u) &= 0, \quad \text{in } \Omega\times[0,T],\\ u_t - \nabla\cdot\sigma - 2(1-\alpha)\nabla\cdot d(u) + \nabla p &= f, \quad \text{in } \Omega\times[0,T],\\ \nabla\cdot u = 0, \quad \text{in } \Omega\times[0,T], \qquad u &= 0, \quad \text{on } \Gamma\times[0,T]. \end{aligned}$$

If $t=0$, then $u=u_0$, $\sigma=\sigma_0$. Here $\Omega\subset R^2$ is an open domain with boundary Γ, σ is the symmetrical stress tensor, u is velocity vector, p is the pressure (scalar), $d(u)=\frac{1}{2}(\nabla u+\nabla u^T)$ is the rate of strain tensor, $\lambda\geq 0$ is Weissenberg constant, α is a third dimensionless, which may be considered as the fraction of viscoelastic viscosity ($\alpha=1$ for Max Well's model). For simplify, we only consider the case of $0<\alpha<1$. $g_a: R^4\times R^4\to R^4$ is a bilinear form:

$$g_a(\sigma,\nabla u)=\sigma\omega(u)-\omega(u)\sigma-a(d(u)\sigma+\sigma d(u))+\frac{1-a}{2}(\sigma\nabla u+\nabla u^T\sigma)-\frac{1+a}{2}(\nabla u\sigma+\sigma\nabla u^T),$$

Here $-1<a<1$ and $\omega(u)=\frac{1}{2}(\nabla u-\nabla u^T)$ is the vorticity tensor.

For analysis the convergence property of approximation solution, we only consider simply form of viscoelastic fluid flow

$$\sigma - 2\alpha d(u) = 0, \quad \text{on } \Omega\times[0,T] \tag{1}$$

$$u_t - \nabla\cdot\sigma - 2(1-\alpha)\nabla\cdot d(u) + \nabla p = f, \quad \text{on } \Omega\times[0,T] \tag{2}$$

$$\nabla\cdot u = 0, \quad \text{on } \Omega\times[0,T], \qquad u = 0, \quad \text{on } \Gamma\times[0,T] \tag{3}$$

We define the energy space as follow

$$\begin{aligned} T &= \left\{\tau=(\tau_{ij});\ \tau_{ij}\in L^2(\Omega), 1\leq i,j\leq 2\right\},\\ X &= (H^1(\Omega))^2,\\ Q &= \left\{p\in L^2(\Omega);\ \int_\Omega p dx = 0\right\}. \end{aligned}$$

We use the standard Sobolev space $W^{m,p}(\Omega)$ with a norm $\|\cdot\|_{m,p}$ given by $\|\phi\|^p_{m,p} = \sum_{|\alpha|\le m} \|D^\alpha \phi\|^p_{L^p(\Omega)}$. For $n = 2$, $H^2(\Omega) = W^{m,2}(\Omega)$ and $\|\cdot\|_m = \|\cdot\|_{m,2}$, $\|\cdot\| = \|\cdot\|_{0,2}$ [2].

Definition 1 $\{\Gamma_h\}_{h>0}$ be a quasi-uniform family of finite element triangular partitions of Ω and any two neighbor units constitutes h^2-approximation parallelogram, there exists an independent of h constant C, such that $|P_1P_2 - P_3P_4| \le Ch^2$, then the partitions family be denoted strong uniform triangular partitions [3].

Definition 2 If the operator $Q_h : L^2(\Omega) \to W_h$ such that for any $g \in L^2(\Omega)$, we have $(Q_h g,\, w_h) = (g,\, w_h)$, $\forall w_h \in W_h$, then Q_h be denoted L^2 projection operator in W_h [2].

Definition 3 If the operator $Q_{l-1} : V_l \to V_{l-1}$ such that for any $w_{l-1} \in V_{l-1}$ we have $(Q_{l-1}v_l, w_{l-1}) = (v_l, w_{l-1})$, then Q_{l-1} be denoted L^2 projection operator in V_{l-1} [1].

For simplify, we introduce some properties in Sobolev space.

Approximation property:

$$\|u - Q_h u\|_{1,2} \le Ch^{k+1}\|u\|_{k+2,2} \tag{4}$$

$$|p - Q_h p|_{0,2} + h|p - Q_h p|_{1,2} \le Ch^{k+1}\|p\|_{k+1,2} \tag{5}$$

$$|\sigma - Q_h \sigma|_{0,2} + h|\sigma - Q_h \sigma|_{1,2} \le Ch^{k+1}\|\sigma\|_{k+1,2} \tag{6}$$

where $(\sigma,\, u,\, p) \in H^{k+1}(\Omega)^4 \times H^{k+2}(\Omega)^2 \times H^{k+1}(\Omega)$.

Inverse estimates: Let Γ_h be a uniform triangulation of Ω, $K \in \Gamma_h$, and P be a finite dimensional subspace of $H^l(K) \cap H^m(K)(0 \le m \le l)$, then there exist constant C such that $\|v\|_{l,2} \le Ch^{m-l}\|v\|_{m,2}, \forall v \in P$.

2. FINITE ELEMENT DISCRETIZATION

Let Γ_h be a uniform triangulation family of Ω, $\bar{\Omega} = \{\cup K, K \in \Gamma_h\}$. For integer $k > 0$, we construct the finite element subspace $T_h \in T$, $X_h \in X$, $Q_h \in Q$

$$\begin{aligned} T_h &= \left\{\tau \in T;\, \tau|_K \in P_k(K)^4,\, \forall K \in \Gamma_h\right\}, \\ X_h &= \left\{v \in X; v|_K \in P_{k+1}(K)^2, \forall K \in \Gamma_h\right\}, \\ Q_h &= \{q \in Q; q|_K \in P_k(K), \forall K \in \Gamma_h\} \end{aligned}$$

where $P_k(K)$ denotes the space of polynomials of degree no more than k on $K \in \Gamma_h$.

The corrpsponding finite element approximation of problem (1–3) is

$$(\sigma_h, \tau_h) - 2\alpha(d(u_h), \tau_h) = 0,\ \forall \tau_h \in T_h \tag{7}$$

$$(u_{h_t}, v_h) + (\sigma_h, d(v_h)) + 2(1-\alpha)(d(u_h), d(v_h)) - (p_h, \nabla \cdot v_h) = (f, v_h),\ \forall v_h \in X_h \tag{8}$$

$$(\nabla \cdot u_h,\, q_h) = 0,\ \forall q_h \in Q_h \tag{9}$$

Let $\Delta t > 0$, $N > 0$, $N = T/\Delta t$ and $t^n = n\Delta t$, $u^n = u(\cdot, t^n)$, $\partial_t u^n = \frac{u^n - u^{n-1}}{\Delta t}$.

Then we get a partitions family of discrete time of mixed finite element aproximation [4], assume that $(\sigma_h^0,\, u_h^0,\, p_h^0) \in T_h \times X_h \times Q_h$, $n = 1, 2, \ldots, N$, we seek $(\sigma_h^n,\, u_h^n,\, p_h^n) \in T_h \times X_h \times Q_h$, such that:

$$(\sigma_h^n,\, \tau_h) - 2\alpha(d(u_h^n),\, \tau_h) = 0,\ \forall \tau_h \in T_h \tag{10}$$

$$(\partial_t u_h^n,\, v_h) + (\sigma_h^n,\, d(v_h)) + 2(1-\alpha)(d(u_h^n),\, d(v_h)) - (p_h^n,\, \nabla \cdot v_h) = (f,\, v_h),\ \forall v_h \in X_h \tag{11}$$

$$(\nabla \cdot u_h^n,\, q_h) = 0,\ \forall q_h \in Q_h \tag{12}$$

Furthermore, we consider two uniform triangulation Γ_{h_1} and Γ_{h_2} of Ω determined as follows. Suppose Γ_{h_1} is given and then Γ_{h_2} be obtained from Γ_{h_1} via regular subdivision: edge midpoints in Γ_{h_1} are connected by new edges to form Γ_{h_2}. Note that for any $T \in \Gamma_{h_2}$, the four subtriangles in Γ_{h_1} are all similar to T and have half the side of T. Thus, $h_2 = h_1/2$ [5, 6].

3. V-CYCLE MULTI-GRID METHOD OF VISCOELASTIC FLUID FLOW

We use V-cycle multi-grid method to solve viscoelastic fluid flow:

Step 1: (presmoothing) In Γ_{h_2}, $\forall \tau_{h_2} \in T_{h_2}$, $v_{h_2} \in X_{h_2}$, $q_{h_2} \in Q_{h_2}$,

Assume that $(\sigma_{h_2}^0,\, u_{h_2}^0,\, p_{h_2}^0) \in T_{h_2} \times X_{h_2} \times Q_{h_2}$, find $(\sigma_{h_2}^n,\, u_{h_2}^n,\, p_{h_2}^n) \in T_{h_2} \times X_{h_2} \times Q_{h_2}$ such that

$$(\sigma_{h_2}^n,\, \tau_{h_2}) - 2\alpha(d(u_{h_2}^n),\, \tau_{h_2}) = 0,\ \forall \tau_{h_2} \in T_{h_2} \tag{13}$$
$$(\partial_t u_{h_2}^n,\, v_{h_2}) + (\sigma_{h_2}^n,\, d(v_{h_2})) + 2(1-\alpha)(d(u_{h_2}^n),\, d(v_{h_2})) - (p_{h_2}^n, \nabla \cdot v_{h_2}) = (f,\, v_{h_2}),\ \forall v_{h_2} \in X_{h_2} \tag{14}$$
$$(\nabla \cdot u_{h_2}^n,\, q_{h_2}) = 0,\ \forall q_h \in Q_{h_2} \tag{15}$$

Step 2: (Intergrid transfer) By connection (2.7–2.9) and (1–3), then

$$(\sigma^n - \sigma_{h_2}^n,\, \tau_{h_2}) - 2\alpha(d(u^n - u_{h_2}^n),\, \tau_{h_2}) = 0,\ \forall \tau_{h_2} \in T_{h_2} \tag{16}$$
$$(\partial_t(u^n - u_{h_2}^n),\, v_{h_2}^n) + (\sigma^n - \sigma_{h_2}^n,\, d(v_{h_2})) + 2(1-\alpha)(d(u^n - u_{h_2}^n),\, d(v_{h_2})) - (p^n - p_{h_2}^n, \nabla \cdot v_{h_2}) = 0,\ \forall v_{h_2} \in X_{h_2} \tag{17}$$
$$(\nabla \cdot (u^n - u_{h_2}^n),\, q_{h_2}) = 0,\ \forall q_{h_2} \in Q_{h_2} \tag{18}$$

Then we transfer remainder to Γ_{h_2} by L^2 projection operator Q_{l-1}, denote $\hat{\sigma}_{h_1}^0 = Q_{l-1}(\sigma^n - \sigma_{h_2}^n)$, $\hat{u}_{h_1}^0 = Q_{l-1}(u^n - u_{h_2}^n)$, $\hat{p}_{h_1}^0 = Q_{l-1}(p^n - p_{h_2}^n)$.

Step 3: (Error correction) Fixed $(\hat{\sigma}_{h_1}^0, \hat{u}_{h_1}^0, \hat{p}_{h_1}^0)$, then (13–15) become: Assume that $(\hat{\sigma}_{h_1}^0, \hat{u}_{h_1}^0, \hat{p}_{h_1}^0) \in T_{h_1} \times X_{h_1} \times Q_{h_1}$, find $(\hat{\sigma}_{h_1}^n,\, \hat{u}_{h_1}^n,\, \hat{p}_{h_1}^n) \in T_{h_1} \times X_{h_1} \times Q_{h_1}$ such that

$$(\hat{\sigma}_{h_1}^n, \tau_{h_1}) - 2\alpha(d(\hat{u}_{h_1}^n), \tau_{h_1}) = 0,\ \forall \tau_{h_1} \in T_{h_1} \tag{19}$$
$$(\partial_t \hat{u}_{h_1}^n,\, v_{h_1}) + (\hat{\sigma}_{h_1}^n,\, d(v_{h_1})) + 2(1-\alpha)(d(\hat{u}_{h_1}^n),\, d(v_{h_1})) - (\hat{p}_{h_1}^n,\, \nabla \cdot v_{h_1}) = 0,\ \forall v_{h_1} \in X_{h_1} \tag{20}$$
$$(\nabla \cdot (\hat{u}_{h_1}^n),\, q_{h_1}) = 0,\ \forall q_{h_1} \in Q_{h_1} \tag{21}$$

Step 4: (Iteration solution) By the Error correction, we obtains a correction solution $(\hat{\sigma}_{h_1}^n, \hat{u}_{h_1}^n, \hat{p}_{h_1}^n)$.

Let $\tilde{\sigma}_h^n = \sigma_{h_2}^n + \hat{\sigma}_{h_1}^n$, $\tilde{u}_h^n = u_{h_2}^n + \hat{u}_{h_1}^n$, $\tilde{p}_h^n = p_{h_2}^n + \hat{p}_{h_1}^n$, we obtains solution $(\tilde{\sigma}_h^n,\, \tilde{u}_h^n,\, \tilde{p}_h^n)$.

4. CONVERGENCE ANALYSIS

Theorem 1 (existence and unique) Solution of V-cycle multi-grid method of viscoelastic fluid flow's simply form (1–3) be existence and unique.

Proof Obviously, the existence and unique of V-cycle multi-grid method's solution only decided by step 1 and step 3 in all four steps.

At first, we consider solution's existence and unique of step 1. We define bilinear form in $T_{h_2} \times X_{h_2} \times Q_{h_2}$

$$\begin{aligned} &A((\sigma_{h_2}^n,\, u_{h_2}^n,\, p_{h_2}^n),\, (\tau_{h_2},\, v_{h_2},\, q_{h_2})) = (\sigma_{h_2}^n, \tau_{h_2}) - 2\alpha(d(u_{h_2}^n),\, \tau_{h_2}) + 2\alpha(d(v_{h_2}),\, \sigma_{h_2}^n) \\ &+2\alpha(\partial_t u_{h_2}^n,\, v_{h_2}) + 4\alpha(1-\alpha)(d(u_{h_2}^n), d(v_{h_2})) - 2\alpha(p_{h_2}^n, \nabla \cdot v_{h_2}) + 2\alpha(q_{h_2}, \nabla \cdot u_{h_2}^n) \\ &+ \sum_K h_2^2(-\nabla \cdot \sigma_{h_2}^n - 2(1-\alpha)\nabla \cdot d(u_{h_2}^n) + \nabla p_{h_2}^n - \nabla \cdot \tau_{h_2} - 2(1-\alpha)\nabla \cdot d(v_{h_2}) + \nabla q_{h_2})_K \end{aligned}$$

At the same time, we define [8]

$$|||(\sigma_{h_2}^n,\, u_{h_2}^n,\, p_{h_2}^n)|||^2 \equiv \sum_K h_2^2 |\nabla \cdot \sigma_{h_2}^n + 2(1-\alpha)\nabla \cdot d(u_{h_2}^n) - \nabla p_{h_2}^n|_K^2$$

Linear form

$$F(\tau_{h_2},\, v_{h_2},\, q_{h_2}) = 2\alpha < f,\, v_{h_2} > + \sum_K h_2^2 < f, -\nabla \cdot \tau_{h_2} - 2(1-\alpha)\nabla \cdot d(v_{h_2}) + \nabla q_{h_2} >_K,$$

then approximation form of problem (2.7–2.9): Find $(\sigma_{h_2}^n,\, u_{h_2}^n,\, p_{h_2}^n) \in T_{h_2} \times X_{h_2} \times Q_{h_2}$ such that

$$A((\sigma_{h_2}^n,\, u_{h_2}^n,\, p_{h_2}^n),\, (\tau_{h_2},\, v_{h_2},\, q_{h_2})) = F(\tau_{h_2},\, v_{h_2},\, q_{h_2}) \tag{22}$$

Now we prove the bilinear form $A((\sigma_{h_2}^n,\, u_{h_2}^n,\, p_{h_2}^n),\, (\tau_{h_2},\, v_{h_2},\, q_{h_2}))$ be coercive. By the define of the bilinear form, we have

$$\begin{aligned} &A((\sigma_{h_2}^n,\, u_{h_2}^n,\, p_{h_2}^n),\, (\sigma_{h_2}^n,\, u_{h_2}^n,\, p_{h_2}^n)) = \|\sigma_{h_2}^n\|^2 - 2\alpha(d(u_{h_2}^n), \sigma_{h_2}^n) + 2\alpha(d(u_{h_2}^n), \sigma_{h_2}^n) \\ &+2\alpha(\partial_t u_{h_2}^n, u_{h_2}^n) + 4\alpha(1-\alpha)\|d(u_{h_2}^n)\|^2 - 2\alpha(p_{h_2}^n, \nabla \cdot u_{h_2}^n) + 2\alpha(p_{h_2}^n, \nabla \cdot u_{h_2}^n) \\ &+ \sum_K h_2^2 \left|\nabla \cdot \sigma_{h_2}^n + 2(1-\alpha)\nabla \cdot d(u_{h_2}^n) - \nabla p_{h_2}^n\right|_K^2. \end{aligned}$$

Then

$$\begin{aligned} & A((\sigma_{h_2}^n,\, u_{h_2}^n,\, p_{h_2}^n),\, (\sigma_{h_2}^n,\, u_{h_2}^n,\, p_{h_2}^n)) \\ = \; & \|\sigma_{h_2}^n\|^2 + 2\alpha(\partial_t u_{h_2}^n, u_{h_2}^n) + 4\alpha(1-\alpha)\|d(u_{h_2}^n)\|^2 + \sum_K h_2^2 \left|\nabla\cdot\sigma_{h_2}^n + 2(1-\alpha)\nabla\cdot d(u_{h_2}^n) - \nabla p_{h_2}^n\right|_K^2. \end{aligned}$$

Choose a proper $0<\alpha<1$, we have

$$A((\sigma_{h_2}^n,\, u_{h_2}^n,\, p_{h_2}^n),\, (\sigma_{h_2}^n,\, u_{h_2}^n,\, p_{h_2}^n)) \geq \alpha |||(\sigma_{h_2}^n,\, u_{h_2}^n,\, p_{h_2}^n)|||^2.$$

Then, bilinear form $A((\sigma_{h_2}^n,\, u_{h_2}^n,\, p_{h_2}^n),\, (\tau_{h_2},\, v_{h_2},\, q_{h_2}))$ be coercivity. Hence, by Lax-Milgram theorem, we show the solution's existence and unique of step 1.

In step 3, we define the bilinear form in $T_{h_1}\times X_{h_1}\times Q_{h_1}$ [9]:

$$\begin{aligned} & B((\hat\sigma_{h_1}^n,\, \hat u_{h_1}^n,\, \hat p_{h_1}^n),\, (\tau_{h_1},\, v_{h_1},\, q_{h_1})) \\ = \; & (\hat\sigma_{h_1}^n,\, \tau_{h_1}) - 2\alpha(d(\hat u_{h_1}^n), \tau_{h_1}) + 2\alpha(d(v_{h_1}),\, \hat\sigma_{h_1}^n) + 2\alpha(\partial_t \hat u_{h_1}^n,\, v_{h_1}) \\ & + 4\alpha(1-\alpha)(d(\hat u_{h_1}^n),\, d(v_{h_1})) - 2\alpha(\hat p_{h_1}^n, \nabla\cdot v_{h_1}) + 2\alpha(q_{h_1},\, \nabla\cdot\hat u_{h_1}^n) + \sum_K h_1^2(\nabla\hat p_{h_1}^n,\, \nabla q_{h_1}) \end{aligned}$$

At the same time, we define [10]

$$|||(\hat\sigma_{h_1}^n,\, \hat u_{h_1}^n,\, \hat p_{h_1}^n)|||_B^2 = \|\hat\sigma_{h_1}^n\|^2 + \|d(\hat u_{h_1}^n)\|^2 + \sum_K h_1^2 \left|\hat p_{h_1}^n\right|_K^2$$

Linear form $F\equiv 0$, then approximation form of problem (19–21):

Find $(\sigma_{h_2}^n,\, u_{h_2}^n,\, p_{h_2}^n)\in T_{h_2}\times X_{h_2}\times Q_{h_2}$ such that $B((\hat\sigma_{h_1}^n,\, \hat u_{h_1}^n,\, \hat p_{h_1}^n),\, (\tau_{h_1},\, v_{h_1},\, q_{h_1})) = 0$

Now we show bilinear form $B((\hat\sigma_{h_1}^n,\, \hat u_{h_1}^n,\, \hat p_{h_1}^n),\, (\tau_{h_1},\, v_{h_1},\, q_{h_1}))$ be coercivity.

At first, by the define of the bilinear form, we have

$$\begin{aligned} & B((\hat\sigma_{h_1}^n,\, \hat u_{h_1}^n,\, \hat p_{h_1}^n),\, (\hat\sigma_{h_1}^n,\, \hat u_{h_1}^n,\, \hat p_{h_1}^n)) \\ = \; & \|\hat\sigma_{h_1}^n\|^2 - 2\alpha(d(\hat u_{h_1}^n), \hat\sigma_{h_1}^n) + 2\alpha(d(\hat u_{h_1}^n), \hat\sigma_{h_1}^n) + 2\alpha(\partial_t \hat u_{h_1}^n, \hat u_{h_1}^n) \\ & + 4\alpha(1-\alpha)\|d(\hat u_{h_1}^n)\|^2 - 2\alpha(\hat p_{h_1}^n,\, \nabla\cdot\hat u_{h_1}^n) + 2\alpha(\hat p_{h_1}^n,\, \nabla\cdot\hat u_{h_1}^n) + \sum_K h_1^2 \left|\nabla\hat p_{h_1}^n\right|_K^2. \end{aligned}$$

Then

$$B((\hat\sigma_{h_1}^n, \hat u_{h_1}^n, \hat p_{h_1}^n), (\hat\sigma_{h_1}^n, \hat u_{h_1}^n, \hat p_{h_1}^n)) = \|\hat\sigma_{h_1}^n\|^2 + 2\alpha(\partial_t \hat u_{h_1}^n, \hat u_{h_1}^n) + 4\alpha(1-\alpha)\|d(\hat u_{h_1}^n)\|^2 + \sum_K h_1^2 \left|\nabla\hat p_{h_1}^n\right|_K^2.$$

Choose a proper $0<\alpha<1$, we have

$$B((\hat\sigma_{h_1}^n,\, \hat u_{h_1}^n,\, \hat p_{h_1}^n),\, (\hat\sigma_{h_1}^n,\, \hat u_{h_1}^n,\, \hat p_{h_1}^n)) \geq \alpha \| \left|(\hat\sigma_{h_1}^n,\, \hat u_{h_1}^n,\, \hat p_{h_1}^n)\right| \|_B^2.$$

Hence, bilinear form $B((\hat\sigma_{h_1}^n,\, \hat u_{h_1}^n,\, \hat p_{h_1}^n),\, (\tau_{h_1},\, v_{h_1},\, q_{h_1}))$ be coercivity. So solution's existence and unique of step 1 have proved by Lax-Milgram theorem.

The desired result then follows.

Theorem 2 Assume the solution $(\sigma^n, u^n, p^n)\in H^{k+1}(\Omega)^4\times H^{k+2}(\Omega)^2\times H^{k+1}(\Omega)$ of problem (1–3) and the irrelative solution $(\sigma_{h_2}^n,\, u_{h_2}^n,\, p_{h_2}^n)\in T_{h_2}\times X_{h_2}\times Q_{h_2}$ in step 1, then there exist a constant C independent of h_2 such that

$$|||(\sigma^n - \sigma_{h_2}^n,\, u^n - u_{h_2}^n,\, p^n - p_{h_2}^n)|||_A^2 \leq C h_2^{2k} \tag{23}$$

Proof Let $e_\sigma = \sigma_{h_2}^n - \sigma^n$, $e_u = u_{h_2}^n - u^n$, $e_p = p_{h_2}^n - p^n$. By the definition of projection operator, let

$$\begin{aligned} e_{\sigma_{h2}^n} &= Q_h\sigma_{h2}^n - \sigma^n, \quad e_{u_{h2}^n} = Q_h u_{h2}^n - u^n, \\ e_{p_{h2}^n} &= Q_h p_{h2}^n - p^n, \quad \xi_{\sigma_{h2}^n} = Q_h\sigma_{h2}^n - \sigma_{h2}^n, \\ \xi_{u_{h2}^n} &= Q_h u_{h2}^n - u_{h2}^n, \quad \xi_{p_{h2}^n} = Q_h p_{h2}^n - p_{h2}^n. \end{aligned}$$

By triangle inequality, we have $|||e_\sigma, e_u, e_p|||^2 \le |||e_{\sigma_{h2}^n}, e_{u_{h2}^n}, e_{p_{h2}^n}|||^2 + |||\xi_{\sigma_{h2}^n}, \xi_{u_{h2}^n}, \xi_{p_{h2}^n}|||^2$. At the same time, by approximation property and inverse estimates, we obtains

$$\begin{aligned}|||(\xi_{\sigma_{h_2}^n}, \xi_{u_{h_2}^n}, \xi_{p_{h_2}^n})|||^2 &\le Ch_2^2(\|\xi_{\sigma_{h_2}^n}\|_{1,2}^2 + \|d(\xi_{u_{h_2}^n})\|_{1,2}^2 + \|\xi_{p_{h_2}^n}\|_{1,2}^2)\\ &\le Ch_2^2(h_2^{2(k+1)} + \left|\xi_{u_{h2}^n}\right| t_{1,2}^2 h_2^{-2} + h_2^{2(k+1)})\\ &\le Ch_2^2(h_2^{2(k+1)} + h_2^{2k} + h_2^{2(k+1)}) \le Ch_2^{2k+2}.\end{aligned}$$

So we only need to prove $|||e_{\sigma_{h_2}^n}, e_{u_{h_2}^n}, e_{p_{h_2}^n}|||^2 \le Ch_2^{2k}$. Now we show that:

$$\begin{aligned}&C|||e_{\sigma_{h_2}^n}, e_{u_{h_2}^n}, e_{p_{h_2}^n}|||^2\\ \le\ & A((e_{\sigma_{h_2}^n}, e_{u_{h_2}^n}, e_{p_{h_2}^n}), (e_{\sigma_{h_2}^n}, e_{u_{h_2}^n}, e_{p_{h_2}^n}))\\ =\ & A((\xi_{\sigma_{h_2}^n} - e_\sigma, \xi_{u_{h_2}^n} - e_u, \xi_{p_{h_2}^n} - e_p), (e_{\sigma_{h_2}^n}, e_{u_{h_2}^n}, e_{p_{h_2}^n}))\\ =\ & A((\xi_{\sigma_{h_2}^n}, \xi_{u_{h_2}^n}, \xi_{p_{h_2}^n}), (e_{\sigma_{h_2}^n}, e_{u_{h_2}^n}, e_{p_{h_2}^n}))\\ =\ & (\xi_{\sigma_{h_2}^n}, e_{\sigma_{h_2}^n}) - 2\alpha(d(\xi_{u_{h_2}^n}), e_{\sigma_{h_2}^n}) + 2\alpha(d(e_{u_{h_2}^n}), \xi_{\sigma_{h_2}^n}) + 2\alpha\left(\frac{\xi_{u_{h_2}^n} - \xi_{u_{h_2}^{n-1}}}{\Delta t}, e_{u_{h_2}^n}\right)\\ &+ 4\alpha(1-\alpha)(d(\xi_{u_{h_2}^n}), d(e_{\sigma_{h_2}^n})) - 2\alpha(\xi_{p_{h_2}^n}, \nabla\cdot e_{u_{h_2}^n}) + 2\alpha(e_{p_{h_2}^n}, \nabla\cdot\xi_{u_{h_2}^n})\\ &+ \sum_K h_2^2(-\nabla\cdot\xi_{\sigma_{h_2}^n} - 2(1-\alpha)\nabla\cdot d(\xi_{u_{h_2}^n}) + \nabla\xi_{p_{h_2}^n}, -\nabla\cdot e_{\sigma_{h_2}^n} - 2(1-\alpha)\nabla\cdot d(e_{u_{h_2}^n}) + \nabla e_{p_{h_2}^n})_K\\ \le\ & \|\xi_{\sigma_{h_2}^n}\|\|e_{\sigma_{h_2}^n}\| + 2\alpha\|d(\xi_{u_{h_2}^n})\|\|e_{\sigma_{h_2}^n}\| + 2\alpha\|d(e_{u_{h_2}^n})\|\|\xi_{\sigma_{h_2}^n}\|\\ &+ \frac{2\alpha}{\Delta t}\|\xi_{u_{h_2}^n} - \xi_{u_{h_2}^{n-1}}\|\|e_{u_{h_2}^n}\| + 4\alpha(1-\alpha)\|d(\xi_{u_{h_2}^n})\|\|d(e_{u_{h_2}^n})\| + 2\alpha\|\xi_{p_{h_2}^n}\|\|\nabla\cdot e_{u_{h_2}^n}\|\\ &+ 2\alpha\|e_{p_{h_2}^n}\|\|\nabla\cdot\xi_{u_{h_2}^n}\| + h_2^2(\|\nabla\cdot\xi_{\sigma_{h_2}^n}\| + 2(1-\alpha)\|\nabla\cdot d(\xi_{u_{h_2}^n})\| + \|\nabla\xi_{p_{h_2}^n}\|)\\ &\cdot(\|\nabla\cdot e_{\sigma_{h_2}^n}\| + 2(1-\alpha)\|\nabla\cdot d(e_{u_{h_2}^n})\| + \|\nabla e_{p_{h_2}^n}\|)\\ \equiv\ & I_1 + I_2 + I_3 + I_4 \end{aligned} \tag{24}$$

By Poincare equality in (23), we have

$$\begin{aligned}I_1 \le\ & \frac{1}{K_1}\|\xi_{\sigma_{h_2}^n}\|^2 + K_1\|e_{\sigma_{h_2}^n}\|^2 + \frac{4\alpha^2}{K_1}\|d(\xi_{u_{h_2}^n})\|^2 + K_1\|e_{\sigma_{h_2}^n}\|^2 + K_1 4\alpha^2\|d(e_{u_{h_2}^n})\|^2 + \frac{1}{K_1}\|\xi_{\sigma_{h_2}^n}\|^2,\\ I_2 =\ & \frac{1}{\Delta t}\|\xi_{u_{h_2}^n} - \xi_{u_{h_2}^{n-1}}\|\|e_{u_{h_2}^n}\| \le \frac{1}{K_2(\Delta t)^2}\|d(\xi_{u_{h_2}^n})\|^2 + K_2\|e_{u_{h_2}^n}\|^2,\\ I_3 \le\ & 4\alpha(1-\alpha)(\frac{1}{K_3}\|d(\xi_{u_{h_2}^n})\|^2 + K_3\|d(e_{u_{h_2}^n})\|^2) + \frac{1}{K_3}\|\xi_{p_{h_2}^n}\|^2 + 4\alpha^2 K_3\|d(e_{u_{h_2}^n})\|^2\\ &+ \frac{4\alpha^2}{K_4}\|\nabla\cdot\xi_{u_{h_2}^n}\|^2 + K_4\|e_{p_{h_2}^n}\|^2,\\ I_4 \le\ & h_2^2(1/K_4\|\nabla\cdot\xi_{\sigma_{h_2}^n}\|^2 + K_4\|\nabla\cdot e_{\sigma_{h_2}^n}\|^2 + 4(1-\alpha)^2/K_4\|\nabla\cdot d(\xi_{u_{h_2}^n})\|^2 + K_4\|\nabla\cdot e_{\sigma_{h_2}^n}\|^2\\ &+ 1/K_4\|\nabla\xi_{p_{h_2}^n}\|^2 + K_4\|e_{\sigma_{h_2}^n}\|^2 + 1/K_4\|\nabla\cdot\xi_{\sigma_{h_2}^n}\|^2 + K_4\|\nabla\cdot e_{u_{h_2}^n}\|^2 + 4(1-\alpha)^2/K_4\|\nabla\cdot d(\xi_{u_{h_2}^n})\|^2\\ &+ K_4\|\nabla\cdot e_{u_{h_2}^n}\|^2 + 1/K_4\|\nabla\xi_{p_{h_2}^n}\|^2 + K_4\|e_{u_{h_2}^n}\|^2 + 1/K_4\|\nabla\cdot\xi_{\sigma_{h_2}^n}\|^2 + K_4\|\nabla\cdot e_{p_{h_2}^n}\|^2\\ &+ 4(1-\alpha)^2/K_4\|\nabla\cdot d(\xi_{u_{h_2}^n})\|^2 + K_4\|\nabla\cdot e_{p_{h_2}^n}\|^2 + 1/K_4\|\nabla\xi_{p_{h_2}^n}\|^2 + K_4\|e_{p_{h_2}^n}\|^2.\end{aligned}$$

Then, we sum all the terms of I_1, I_2, I_3, I_4 and turn the terms including $e_{\sigma_{h_2}^n}$, $e_{u_{h_2}^n}$, $e_{p_{h_2}^n}$ to left. Choose the proper K_i $(i = 1, 2, 3, 4)$, we have $|||e_{\sigma_{h_2}^n}, e_{u_{h_2}^n}, e_{p_{h_2}^n}|||^2 \le Ch_2^{2k}$. The desired result then follows.

Theorem 3 Assume the solution$(\hat{\sigma}_{h_1}^0, \hat{u}_{h_1}^0, \hat{p}_{h_1}^0) \in T_{h_1} \times X_{h_1} \times Q_{h_1}$ of problem (19–21), and the irrelative solution $(\hat{\sigma}_{h_1}^n, \hat{u}_{h_1}^n, \hat{p}_{h_1}^n) \in T_{h_1} \times X_{h_1} \times Q_{h_1}$ in step 3, then there exist a constant C independent of h_1 such that

$$|||(\hat{\sigma}_{h_1}^n - \hat{\sigma}_{h_1}^0, \hat{u}_{h_1}^n - u_{h_1}^0, \hat{p}_{h_1}^n - \hat{p}_{h_1}^0)|||_B^2 \le Ch_1^{2k} \tag{25}$$

Proof Let $\varphi_\sigma = \hat{\sigma}^n_{h_1} - \hat{\sigma}^0_{h_1}$, $\varphi_u = \hat{u}^n_{h_1} - \hat{u}^0_{h_1}$, $\varphi_p = \hat{p}^n_{h_1} - \hat{p}^0_{h_1}$, $\varphi_{\sigma^n_{h_1}} = Q_h\hat{\sigma}^n_{h_1} - \hat{\sigma}^0_{h_1}$, $\varphi_{u^n_{h_1}} = Q_h\hat{u}^n_{h_1} - \hat{u}^0_{h_1}$, $\eta_{\hat{\sigma}^n_{h_1}} = Q_h\hat{u}^n_{h_1} - \hat{u}^n_{h_1}$, $\eta_{\hat{p}^n_{h_1}} = Q_h\hat{p}^n_{h_1} - \hat{p}^n_{h_1}$.

Now, we show that $|||\varphi_{\hat{\sigma}^n_{h_1}}, \varphi_{\hat{u}^n_{h_1}}, \varphi_{\hat{p}^n_{h_1}}|||^2 \le C|||\eta_{\hat{\sigma}^n_{h_1}}, \eta_{\hat{u}^n_{h_1}}, \eta_{\hat{p}^n_{h_1}}|||^2$:

$$
\begin{aligned}
& C|||\varphi_{\hat{\sigma}^n_{h_1}}, \varphi_{\hat{u}^n_{h_1}}, \varphi_{\hat{p}^n_{h_1}}|||^2 \\
\le\ & B((\varphi_{\hat{\sigma}^n_{h_1}}, \varphi_{\hat{u}^n_{h_1}}, \varphi_{\hat{p}^n_{h_1}}), (\varphi_{\hat{\sigma}^n_{h_1}}, \varphi_{\hat{u}^n_{h_1}}, \varphi_{\hat{p}^n_{h_1}})) \\
=\ & B((\eta_{\hat{\sigma}^n_{h_1}} - \varphi_{\hat{\sigma}}, \eta_{\hat{u}^n_{h_1}} - \varphi_{\hat{u}}, \eta_{\hat{p}^n_{h_1}} - \varphi_{\hat{p}}), (\varphi_{\hat{\sigma}^n_{h_1}}, \varphi_{\hat{u}^n_{h_1}}, \varphi_{\hat{p}^n_{h1}})) \\
=\ & B((\eta_{\hat{\sigma}^n_{h_1}}, \eta_{\hat{u}^n_{h_1}}, \eta_{\hat{p}^n_{h_1}}), (\varphi_{\hat{\sigma}^n_{h_1}}, \varphi_{\hat{u}^n_{h_1}}, \varphi_{\hat{p}^n_{h1}})) \\
=\ & (\eta_{\hat{\sigma}^n_{h_1}}, \varphi_{\hat{\sigma}^n_{h_1}}) - 2\alpha(d(\eta_{\hat{u}^n_{h_1}}), \varphi_{\hat{\sigma}^n_{h_1}}) + 2\alpha(d(\varphi_{\hat{u}^n_{h_1}}), \eta_{\hat{\sigma}^n_{h1}}) \\
& + \sum_K h_1^2(\nabla\eta_{\hat{p}^n_{h_1}}, \nabla\varphi_{\hat{p}^n_{h_1}})_K + 2\alpha\left(\frac{\eta_{\hat{u}^n_{h_1}} - \eta_{\hat{u}^{n-1}_{h_1}}}{\Delta t}, \varphi_{\hat{u}^n_{h_1}}\right) \\
& + 4\alpha(1-\alpha)(d(\eta_{\hat{u}^n_h}), d(\varphi_{\hat{\sigma}^n_{h1}})) - 2\alpha(\eta_{\hat{p}^n_{h_1}}, \nabla\cdot\varphi_{\hat{u}^n_{h_1}}) + 2\alpha(\varphi_{\hat{p}^n_{h_1}}, \nabla\cdot\eta_{u^n_{h_1}}) \\
\le\ & \|\eta_{\hat{\sigma}^n_{h_1}}\|\|\varphi_{\hat{\sigma}^n_{h_1}}\| + 2\alpha\|d(\eta_{\hat{u}^n_{h_1}})\|\|\varphi_{\hat{\sigma}^n_{h_1}}\| + 2\alpha\|d(\varphi_{\hat{u}^n_{h1}})\|\|\eta_{\hat{\sigma}^n_{h_1}}\| + h_1^2\|\nabla\eta_{\hat{p}^n_{h_1}}\|\|\nabla\varphi_{\hat{p}^n_{h_1}}\| \\
& + \frac{2\alpha}{\Delta t}\|\eta_{\hat{u}^n_{h_1}} - \eta_{\hat{u}^{n-1}_{h_1}}\|\|\varphi_{\hat{u}^n_{h_1}}\| + 4\alpha(1-\alpha)\|d(\eta_{\hat{u}^n_{h_1}})\|\|d(\varphi_{\hat{u}^n_{h_1}})\| \\
& + 2\alpha\|\eta_{\hat{p}^n_{h_1}}\|\|\nabla\cdot\varphi_{\hat{u}^n_{h_1}}\| + 2\alpha\|\varphi_{\hat{p}^n_{h_1}}\|\|\nabla\cdot\eta_{\hat{u}^n_{h_1}}\| \\
\equiv\ & J_1 + J_2 + J_3
\end{aligned}
\tag{26}
$$

By Poincare inequality in (26), we have

$$
\begin{aligned}
J_1 \ &\le\ \frac{1}{L_1}\|\eta_{\hat{\sigma}^n_{h_1}}\|^2 + L_1\|\varphi_{\hat{\sigma}^n_{h_1}}\|^2 + \frac{4\alpha^2}{L_1}\|d(\eta_{\hat{u}^n_{h_1}})\|^2 + L_1\|\varphi_{\hat{\sigma}^n_{h_1}}\|^2 + L_1 4\alpha^2\|d(\varphi_{\hat{u}^n_{h1}})\|^2 + \frac{1}{L_1}\|\eta_{\hat{\sigma}^n_{h_1}}\|^2, \\
J_2 \ &\le\ \frac{1}{L_2(\Delta t)^2}\|d(\xi_{u^n_{h_2}})\|^2 + L_2\|e_{u^n_{h_2}}\|^2 + \frac{h_1^2}{L_2}\|\nabla\eta_{\hat{p}^n_{h_1}}\|^2 + L_2\|\nabla\varphi_{\hat{p}^n_{h_1}}\|^2, \\
J_3 \ &\le\ 4\alpha(1-\alpha)\left(\frac{1}{L_3}\|d(\eta_{\hat{u}^n_{h_1}})\|^2 + L_3\|d(\varphi_{\hat{u}^n_{h_1}})\|^2\right) + \frac{1}{L_3}\|\eta_{\hat{p}^n_{h_1}}\|^2 \\
& \quad + 4\alpha^2 L_3\|d(\varphi_{\hat{u}^n_{h_1}})\|^2 + \frac{4\alpha^2}{L_3}\|\nabla\cdot\eta_{\hat{u}^n_{h_1}}\|^2 + L_3\|\varphi_{\hat{p}^n_{h1}}\|^2,
\end{aligned}
$$

Then, we sum all the terms of J_1, J_2, J_3 and turn the terms including $\varphi_{\hat{\sigma}^n_{h_1}}, \varphi_{\hat{u}^n_{h_1}}, \varphi_{\hat{p}^n_{h_1}}$ to left. Choose the proper L_j, $(j = 1, 2, 3)$, we have

$$
|||\varphi_{\hat{\sigma}^n_{h_1}}, \varphi_{\hat{u}^n_{h_1}}, \varphi_{\hat{p}^n_{h_1}}|||^2 \le C|||\eta_{\hat{\sigma}^n_{h_1}}, \eta_{\hat{u}^n_{h_1}}, \eta_{\hat{p}^n_{h_1}}|||^2.
$$

By approximation property and inverse estimates, we obtains

$$
\begin{aligned}
|||(\eta_{\hat{\sigma}^n_{h_1}}, \eta_{\hat{u}^n_{h_1}}, \eta_{\hat{p}^n_{h_1}})|||^2 &\le Ch_1^2(\|\eta_{\hat{\sigma}^n_{h_1}}\|^2_{1,2} + \|d(\eta_{\hat{u}^n_{h_1}})\|^2_{1,2} + \|\eta_{\hat{p}^n_{h_1}}\|^2_{1,2}) \\
&\le Ch_1^2(h_1^{2(k+1)} + |\eta_{\hat{u}^n_{h_1}}|^2_{1,2}\, h_1^{-2} + h_1^{2(k+1)}) \\
&\le Ch_1^2(h_1^{2(k+1)} + h_1^{2k} + h_1^{2(k+1)}) \\
&\le Ch_1^{2k+2}.
\end{aligned}
$$

By triangle inequality, we have $|||\varphi_{\hat{\sigma}}, \varphi_{\hat{u}}, \varphi_{\hat{p}}|||^2 \le |||\varphi_{\hat{\sigma}^n_{h_1}}, \varphi_{\hat{u}^n_{h_1}}, \varphi_{\hat{p}^n_{h_1}}|||^2 + |||\eta_{\hat{\sigma}^n_{h_1}}, \eta_{\hat{u}^n_{h_1}}, \eta_{\hat{p}^n_{h_1}}|||^2$. The desired result then follows.

Theorem 4 Assume the solution $(\sigma^n, u^n, p^n) \in H^{k+1}(\Omega)^4 \times H^{k+2}(\Omega)^2 \times H^{k+1}(\Omega)$ of problem (1–3), and the irrelative solution $(\sigma^n_{h_2}, u^n_{h_2}, p^n_{h_2}) \in H^{k+1}(\Omega)^4 \times H^{k+2}(\Omega)^2 \times H^{k+1}(\Omega)$ in step 1, then there exist a constant C independent of h_2 such that the irrelative solution $(\tilde{\sigma}^n_h, \tilde{u}^n_h, \tilde{p}^n_h) \in T_h \times X_h \times Q_h$ of V cycle multi-grid method:

$$
\|\sigma^n - \tilde{\sigma}^n_h\|_{0,2} \le Ch_2^k \tag{27}
$$

$$
\|u^n - \tilde{u}^n_h\|_{1,2} + \|p^n - \tilde{p}^n_h\|_{1,2} \le Ch_2^{k-1} \tag{28}
$$

Proof By (4–6) and lemma 3, we obtain

$$\begin{aligned}
\|u^n - \tilde{u}_h^n\|_{1,2} &\le \|\hat{u}_{h_1}^n - \hat{u}_{h_1}^0\|_{1,2} + \|(u^n - u_{h_2}^n) - Q_{l-1}(u^n - u_{h_2}^n)\|_{1,2} \\
&\le Ch_2^{k+1}\|u^n - u_{h_2}^n\|_{k+2,2} + Ch_1^k \\
&\le Ch_2^{k+1} \cdot h_2^{-k}\|u^n - u_{h_2}^n\|_{2,2} + Ch_1^k \\
&\le Ch_2^k, \\
\|\sigma^n - \tilde{\sigma}_h^n\|_{0,2} &\le \|\hat{\sigma}_{h_1}^n - \hat{\sigma}_{h_1}^n\|_{0,2} + \|(\sigma^n - \sigma_{h_2}^n) - Q_{l-1}(\sigma^n - \sigma_{h_2}^n)\|_{0,2} \\
&\le Ch_1^k + Ch_2^{k+1}\|\sigma^n - \sigma_{h_2}^n\|_{k+1,2} \\
&\le Ch_2^{k+1} \cdot h_2^{-k}\|\sigma^n - \sigma_{h_2}^n\|_{1,2} + Ch_1^k \\
&\le Ch_2^{k-1}, \\
\|p^n - \tilde{p}_h^n\|_{1,2} &\le \|(p^n - p_{h_2}^n) - Q_{l-1}(p^n - p_{h_2}^n)\|_{1,2} + \|\hat{p}_{h_1}^n - \hat{p}_{h_1}^0\|_{1,2} \\
&\le Ch_2^{k+1}\|p^n - p_{h_2}^n\|_{k+1,2} + Ch_1^k \\
&\le Ch_2^{k+1} \cdot h_2^{-k}\|p^n - p_{h_2}^n\|_{1,2} + Ch_1^k \\
&\le Ch_2^{k-1} + Ch_1^k \\
&\le Ch_2^{k-1},
\end{aligned}$$

The desired result then follows.

REFERENCES

1. Brenner, S. C. and L. R. Scott, *The Mathematical Theory of Finite Element Methods*, Springer-verlag, 1994.
2. Adams, R. A., *Sobolev Spaces*, Academic Press, New York-San Francisco-London, 1975.
3. Zhang, H. and Z. Lu, "Mixed finite element method of viscoelastic fluid flow," *Journal of Changsha University of Electric Power*, Vol. 4, 2006, to appear.
4. Chen, Y., Y. Q. Huang, and D. Yu, "A Two-grid method for expanded mixed finite-element solution of semilinear reaction-diffusion equations," *Natural Science Journal of Xiangtan University*, Vol. 24, 101–124, 2002.
5. Xu, J., "Two-grid discretization techniques for linear and non-linear PDEs," *SIAM Journal on Nuerical Analysis*, Vol. 33, 1759–1777, 1996.
6. Xu, J., "A novel two-grid method for semilinear equations," *SIAM Journal on Scientific computing*, Vol. 15, 231–237, 1994.
7. Layton, W. J., "A two-level discretization method for the Navier-stokes equations," *Coput. Math. Appl.*, Vol. 26, No. 2, 31–38, 1993.
8. Dawson, C. N. and M. F. Wheeler, "Two-grid methods for mixed finite element approximations of non-linear parabolic equations," *Mathematics*, Vol. 180, 191–203, 1994.
9. Baranger, J. and D. Sandri, "Finite element approximation of viscoelastic fluid flow: Existence of approximate solutions and error bounds," *Numerische Mathematik*, Springer-verlag, Vol. 63, 13–27, 1992.
10. Sandri, D., "Finite element approximation of viscoelastic fluid flow: Existence of approximate solutions and error bounds.II.continuous approximation of the stress," *SIAM Journal on Nuerical Analysis*, Vol. 31, 362–377, 1994.

Result Concerning Delay-Dependent Exponential Stability of Delayed Neural Networks

Xin-Ge Liu, **Hai-Tao Cai**, **and Mei-Lan Tang**

School of Mathematical Science and Computing Technology
Central South University
Changsha, Hunan 410083, China

Abstract— Without assumption of the boundedness of the activation functions, using Lyapunov-Krasovskii functional, relationships among the state vectors of the neural networks together with homeomorphism map, a linear matrix inequality (LMI) approach is developed to derive a new delay-dependent sufficient condition for (i) existence (ii) uniqueness and (iii) global exponential stability of equilibrium point, of a class of time-varying delay neural networks. Example is provided to demonstrate the effectiveness of the proposed result.

1. INTRODUCTION

Neural networks have aroused a tremendous surge of investigation in these years[1–5]. A class of neural networks with time-varying delay (DNN) can be described by

$$\begin{aligned} \dot{x}(t) &= -Ax(t) + Wg(x(t)) + W_1 g(x(t-\tau(t))) + u, \\ x(t) &= \phi(t), \quad t \in [-2h, 0], \end{aligned} \tag{1}$$

where $x(t) = [x_1(t), \ldots, x_n(t)]^T$ is the neuron state vector, $g(x(t)) = [g_1(x_1(t)), \ldots, g_n(x_n(t))]^T$ is the activation function vector, and $u = [u_1, \ldots, u_n]^T$ is a constant external input vector. In the neural networks (1), matrices $A = \mathrm{diag}(a_1, a_2, \ldots, a_n)$, where the scalar $a_i > 0$ is the rate with which the ith unit will reset its potential to the resting state in isolation when disconnected from the network and external inputs at time t. Matrices W and W_1 are the connection weighting matrix and the delayed weighting matrix, respectively. $\tau(t)$ is a continuous function describing the time-varying transmission delays in the neural networks and satisfies $0 \le \tau(t) \le h$, $\dot{\tau}(t) \le \mu$ for all $t \ge 0$, with h and μ being two nonnegative constants. The initial condition function $\phi(t) : [-2h, 0] \to R^n$ is assumed to be a continuous function. Throughout this paper, $\lambda_M(Q)$ denotes the maximum eigenvalue of the symmetric matrix Q and $\lambda_m(Q)$ the minimum; it is assumed that each neuron activation function g_j satisfies the following assumption:

Assumption 1.

$$|g_j(s) - g_j(t)| \le k_j |s - t|, \qquad \forall s, t \in R, \quad j = 1, \ldots, n. \tag{2}$$

Denote $K = \mathrm{diag}(k_1, \ldots, k_n)$ and $k = \max\limits_{i=1,\ldots,n} k_i$. Obviously, $K \ge 0$.

Definition 1. *The vector $x^* = [x_1^*, \ldots, x_n^*]^T$ is said to be an equilibrium point of the DNN in Eq. (1) if $-Ax^* + Wg(x^*) + W_1 g(x^*) + u = 0$.*

Definition 2. *DNN in Eq. (1) is said to be exponentially stable if there is scalars $\alpha > 0$ and $\gamma > 0$ such that*

$$|x(t)| \le \gamma e^{-\alpha t} \sup_{-2h \le \theta \le 0} |x(\theta)|.$$

Definition 3. *A map $H : R^n \to R^n$ is a homeomorphism of R^n onto itself, if $H \in C^0$, H is one-to-one, H is onto and the inverse map $H^{-1} \in C^0$, where C^0 represents the set of all continuous functions from R^n to R^n.*

Lemma 1. *If $H(x) \in C^0$ and satisfies conditions: (1) $H(x)$ is injective on R^n, (2) $\lim\limits_{\|x\| \to +\infty} \|H(x)\| = +\infty$. Then $H(x)$ is a homeomorphism of R^n [5].*

Some existing results on global exponential stability of the equilibrium of DNN concern the case where the activation functions are all bounded and strictly increasing. These assumptions make the results inapplicable to some important engineering problems. In this paper, we develop

a new delay-dependent exponential stability condition for a class of neural networks with time-varying delay by utilizing Lyapunov functional and relationships among the state vectors of the neural networks . We only assume that the activation functions are globally Lipschitz continuous. Under this assumption, both the existence of a unique equilibrium point and the global exponential stability of a given delayed neural networks are proved. The derived condition is less conservative than the sufficient conditions in [3, 4] and can be expressed in term of a linear matrix inequality (LMI). Furthermore, our delay-dependent stability criterion removes the unreasonable restriction that the derivative of the delay function $\tau(t)$ is less than 1. Example is provided to demonstrate the effectiveness of the proposed condition.

2. MAIN RESULT

Theorem 1. *For given scalars $h > 0$ and $\mu \geq 0$, the delayed neural networks in (1) satisfying Assumption 1 has the unique equilibrium point which is globally exponentially stable for any delay $0 < \tau(t) \leq h$ if there exist positive definite matrices $D > 0, P > 0, Q > 0, R > 0, P_k, k = 1, 2, \ldots, 6, T_i, i = 1, 2, \ldots, 5$ and diagonal matrices $S_j > 0$ and $Y_j > 0, j = 1, 2$ such that the following LMI holds:*

$$\Phi = \begin{bmatrix} \Phi_{11} & \Phi_{12} & \Phi_{13} & P_1W_1 - A^TP_4^T + T_4^T & P - P_1 - A^TP_5^T + T_5^T & -A^TP_6^T - T_1 \\ * & \Phi_{22} & \Phi_{23} & S_2 + P_2W_1 - T_4^T & -P_2 - T_5^T & -T_2 \\ * & * & \Phi_{33} & P_3W_1 + W^TP_4^T & -P_3 + W^TP_5^T & W^TP_6^T - T_3 \\ * & * & * & -(1-\mu)Q - Y_2 + 2P_4W_1 & -P_4 + W_1^TP_5^T & W_1^TP_6^T - T_4 \\ * & * & * & * & hD - 2P_5 & -P_6^T - T_5 \\ * & * & * & * & * & -\frac{D}{h} \end{bmatrix} < 0 \quad (3)$$

where $\Phi_{11} = R + KY_1K + 2KS_1 - 2P_1A + 2T_1$, $\Phi_{12} = -A^TP_2^T - T_1 + T_2^T$, $\Phi_{13} = S_1 + P_1W - A^TP_3^T + T_3^T$, $\Phi_{22} = -(1-\mu)R + KY_2K + 2KS_2 - 2T_2$, $\Phi_{23} = P_2W - T_3^T$, $\Phi_{33} = Q - Y_1 + 2P_3W$.

Proof. Define map $H : R^n \to R^n$

$$H(x) = -Ax + Wg(x) + W_1g(x)) + u. \quad (4)$$

Denote $\zeta(x, y) = [(x-y)^T, (x-y)^T, (g(x)-g(y))^T, (g(x)-g(y))^T, (H(x)-H(y))^T, 0]^T$. Then

$$\begin{aligned} & \zeta(x,y)^T\Phi\zeta(x,y) \\ = \; & (x-y)^T[\mu R + K(Y_1+Y_2)K + 2K(S_1+S_2)](x-y) \\ & + 2(x-y)^T[S_1+S_2](g(x)-g(y)) + 2(x-y)^TP(H(x)-H(y)) \\ & + (g(x)-g(y))^T[\mu Q - Y_1 - Y_2](g(x)-g(y)) + h(H(x)-H(y))^TD(H(x)-H(y)). \end{aligned} \quad (5)$$

In the rest of this Section, we prove this new sufficient condition in Theorem 1 by three steps.

Step 1. We prove that $H(x)$ is an injective map on R^n by showing that assuming otherwise leads to a contradiction.

We suppose that vectors x, y exist in R^n such that $x \neq y$ while $H(x) = H(y)$, then $x - y \neq 0$, $A(x-y) = (W + W_1)(g(x) - g(y))$. From Inequality (3), we have

$$\zeta(x,y)^T\Phi\zeta(x,y) < 0. \quad (6)$$

On the other hand, Noting that $\mu \geq 0, R > 0$, $Q > 0$ $S_i \geq 0$ and $Y_i \geq 0, i = 1, 2$, from Eq. (5) and Inequality (2), we have

$$\begin{aligned} & \zeta(x,y)^T\Phi\zeta(x,y) \\ = \; & (x-y)^T[\mu R + K(Y_1+Y_2)K + 2K(S_1+S_2)](x-y) \\ & + 2(x-y)^T[S_1+S_2](g(x)-g(y)) + (g(x)-g(y))^T[\mu Q - Y_1 - Y_2](g(x)-g(y)) \\ \geq \; & 0. \end{aligned} \quad (7)$$

which contradicts Inequality (6), and hence implies that $H(x)$ is an injective map on R^n.

Step 2. We prove that $\lim_{\|x\| \to +\infty} \|H(x)\| = +\infty$.

Since $H(0)$ is a constant vector, it suffices to show that $\lim_{\|x\|\to+\infty} \|\tilde{H}(x)\| = +\infty$ where $\tilde{H}(x) = H(x) - H(0)$.

Since $\zeta(x,0) = [x^T, x^T, (g(x)-g(0))^T, (g(x)-g(0))^T, \tilde{H}(x)^T, 0]^T$, from Inequality (3), there exist a small positive real number $a > 0$ such that

$$\zeta(x,0)^T \Phi \zeta(x,0) \leq -ax^T x. \tag{8}$$

Noting $\tilde{H}(x) = H(x) - H(0)$, $\mu \geq 0, R > 0$, $Q > 0$ $S_i \geq 0$ and $Y_i \geq 0, i = 1, 2$, from Eq. (5) and Inequality (2), we get

$$\zeta(x,0)^T \Phi \zeta(x,0) \geq 2x^T P \tilde{H}(x). \tag{9}$$

So

$$2x^T P \tilde{H}(x) \leq -ax^T x. \tag{10}$$

Since P is positive definite matrix, we have

$$a\|x\|^2 = ax^T x \leq |2\tilde{H}(x)^T P x| \leq 2\lambda_{\max}(P)\|\tilde{H}(x)\|\|x\|, \tag{11}$$

Therefore,

$$\lim_{\|x\|\to+\infty} \|\tilde{H}(x)\| = +\infty.$$

From Steps 1 and 2, the map $H(x)$ is a homeomorphism of R^n. Thus the DNN given in Eq. (1) has a unique equilibrium point. We denote the unique equilibrium point by x^*.

Using the transformation $z(t) = x(t) - x^*$, we can now put the DNN given in Eq. (1) into the following form:

$$\begin{aligned} \dot{z}(t) &= -Az(t) + Wf(z(t)) + W_1 f(z(t-\tau(t))) \\ z(t) &= \phi(t) - x^*, \qquad t \in [-2h, 0] \end{aligned} \tag{12}$$

where $z(t) = [z_1(t), \ldots, z_n(t)]^T$, $f(z(t)) = [f_1(z_1(t)), \ldots, f_n(z_n(t))]^T$, $f_j(z_j(t)) = g_j(z_j(t) + x_j^*) - g_j(x_j^*)$, $j = 1, \ldots, n$. Note that the functions $f_j(z_j(t))$ satisfy the following conditions:

$$|f_j(s)| \leq k_j|s|, \quad f_j(0) = 0, \qquad \forall s \in R, \quad j = 1, \ldots, n. \tag{13}$$

By construction, the origin is the unique equilibrium points of the DNN given in Eq. (12). Next, we will prove the origin is exponential stable.

Step 3. We choose a Lyapunov-Krasovskii functional candidate for the DNN given in (12) as

$$\begin{aligned} V(z_t) &= z^T(t)Pz(t) + \int_{t-\tau(t)}^{t} [z^T(s)Rz(s) + f(z(s))^T Qf(z(s))]ds \\ &\quad + \int_{-h}^{0}\int_{t+r}^{t} y(s)^T Dy(s)dsdr \end{aligned} \tag{14}$$

where $y(s) = \dot{z}(s), D > 0, P > 0, Q > 0$ and $R > 0$.

Calculating the derivative of $V(z_t)$ along the solution of Eq. (12) yields

$$\begin{aligned} \dot{V}(z_t) \leq\ & 2z^T(t)Py(t) + z^T(t)Rz(t) - (1-\mu)z^T(t-\tau(t))Rz(t-\tau(t)) \\ & + f(z(t))^T Qf(z(t)) - (1-\mu)f(z(t-\tau(t)))^T Qf(z(t-\tau(t))) \\ & + hy(t)^T Dy(t) - \int_{t-\tau(t)}^{t} y(s)^T Dy(s)ds \\ & + 2z^T(t)S_1 f(z(t)) + z(t)^T KY_1 Kz(t) + 2z^T(t)KS_1 z(t) - f(z(t))^T Y_1 f(z(t)) \\ & + 2z^T(t-\tau(t))S_2 f(z(t-\tau(t))) + z(t-\tau(t))^T KY_2 Kz(t-\tau(t)) \\ & + 2z^T(t-\tau(t))KS_2 z(t-\tau(t)) - f(z(t-\tau(t)))^T Y_2 f(z(t-\tau(t))) \\ \leq\ & \xi(t)^T \begin{bmatrix} R + KY_1K + 2KS_1 & 0 & S_1 & 0 & P \\ * & -(1-\mu)R + KY_2K + 2KS_2 & 0 & S_2 & 0 \\ * & * & Q - Y_1 & 0 & 0 \\ * & * & * & -(1-\mu)Q - Y_2 & 0 \\ * & * & * & * & hD \end{bmatrix} \xi(t) \\ & - \frac{1}{\tau(t)} \int_{t-\tau(t)}^{t} \tau(t) y(s)^T \frac{D}{h} \tau(t) y(s) ds \end{aligned} \tag{15}$$

where $\xi(t) = [z^T(t), z^T(t-\tau(t)), f(z(t))^T, f(z(t-\tau(t)))^T, y^T(t)]^T$.

It is easy to prove for appropriate dimensional free weight matrices $T_i, i = 1, 2, \ldots, 5$

$$\xi(t)^T \begin{bmatrix} -2P_1A + 2T_1 & -A^TP_2^T - T_1 + T_2^T & P_1W - A^TP_3^T + T_3^T & P_1W_1 - A^TP_4^T + T_4^T & -P_1 - A^TP_5^T + T_5^T \\ * & -2T_2 & P_2W - T_3^T & P_2W_1 - T_4^T & -P_2 - T_5^T \\ * & * & 2P_3W & P_3W_1 + W^TP_4^T & -P_3 + W^TP_5^T \\ * & * & * & 2P_4W_1 & -P_4 + W_1^TP_5^T \\ * & * & * & * & -2P_5 \end{bmatrix} \xi(t)$$
$$+ 2\xi(t)^T \begin{bmatrix} -A^TP_6^T - T_1 \\ -T_2 \\ W^TP_6^T - T_3 \\ W_1^TP_6^T - T_4 \\ -P_6^T - T_5 \end{bmatrix} \int_{t-\tau(t)}^{t} y(s)ds \equiv 0. \tag{16}$$

Adding Eq. (16) on both sides of Inequality (15), we have

$$\dot{V}(z_t) \leq \frac{1}{\tau(t)} \int_{t-\tau(t)}^{t} \eta(t,s)^T \Phi \eta(t,s) ds \tag{17}$$

where $\eta(t,s) = [z^T(t), z^T(t-\tau(t)), f(z(t))^T, f(z(t-\tau(t)))^T, y^T(t), \tau(t)y(s)^T]^T$.

By Inequalities (3) and (17), it is easy to see that there exist scalars $a > 0$, $c > 0$ such that

$$\dot{V}(z_t) \leq -a|z(t)|^2 - cy(t)^T Dy(t). \tag{18}$$

Now, we can choose a scalar $b > 0$ satisfying

$$b\lambda_{\max}(P) + bhe^{bh}(\lambda_{\max}(R) + k^2\lambda_{\max}(Q)) - a \leq 0, \tag{19}$$
$$bh^2e^{bh} - c \leq 0. \tag{20}$$

Then, for the above scalar $b > 0$, from Eq. (14) and Inequality (17) we have

$$\begin{aligned} \frac{d}{dt}(e^{bt}V(z_t)) \quad \leq \quad & e^{bt}\{[b\lambda_{\max}(P) - a]|z(t)|^2 + b[\lambda_{\max}(R) + k^2\lambda_{\max}(Q)] \int_{t-\tau(t)}^{t} z^T(s)z(s)ds \\ & + bh \int_{t-h}^{t} y(s)^T Dy(s)ds - cy(t)^T Dy(t)\}. \end{aligned} \tag{21}$$

Now, integrating both sides of Inequality (21) from 0 to $t > 0$ and using Inequalities (19) and (20), we obtain

$$e^{bt}V(z_t) \leq q(\sup_{-2h \leq \theta \leq 0} |z(\theta)|)^2 \tag{22}$$

where $q = (bhe^{bh} + 1)h[\lambda_{\max}(R) + k^2\lambda_{\max}(Q)] + \lambda_{\max}(P) + 3(bh^2e^{bh} + h)h\lambda_{\max}(D)[\lambda_{\max}^2(A) + k^2\lambda_{\max}^2(W) + k^2\lambda_{\max}^2(W_1)]$.

Hence

$$|z(t)| < e^{-bt/2}\sqrt{\frac{q}{\lambda_{\max}(P)}} \sup_{-2h \leq \theta \leq 0} |z(\theta)|. \tag{23}$$

It follows from Inequalities (23) that the DNN in Eq. (12) is exponentially stable for any delay $0 < \tau(t) \leq h$. This completes the proof.

3. EXAMPLES

Example 1. *We consider the DNN given in Eq. (1) with parameters as*

$$A = \begin{bmatrix} 1.0674 & 0 & 0 \\ 0 & 2.2094 & 0 \\ 0 & 0 & 0.8352 \end{bmatrix}, \quad W = \begin{bmatrix} 0.4094 & 0.5719 & 0.2503 \\ -1.0645 & 0.0410 & 0.9923 \\ -0.7439 & 0.63443 & 0.1066 \end{bmatrix},$$
$$W_1 = \begin{bmatrix} 0.3008 & 0 & 0 \\ 0 & 0.3070 & 0 \\ 0 & 0 & 0.3068 \end{bmatrix}, \quad K = \begin{bmatrix} 0.4911 & 0 & 0 \\ 0 & 0.9218 & 0 \\ 0 & 0 & 0.6938 \end{bmatrix}.$$

For this example, when $\mu = 0.1$ as in [3], if the activation functions are unbounded, Theorem 2 in [3] fails, while, our exponential stability criterion in Theorem 1 remains applicable.

If the time-varying delay function $\tau(t)$ is differentiable, but the upper bound of the derivative of the delay function $\tau(t)$ is greater than 1, Theorem 2 in [3] fails, while, our exponential stability criterion in Theorem 1 remains applicable. For example, if $\mu = 1.2$, we find the allowable maximum delay $h = 3.5336$, under which the DNN in this example is globally exponentially stable.

4. CONCLUSION

Without assuming that activation functions are bounded, a new condition for the global exponential stability of DNN has been obtained. This stability criterion is expressed in term of LMI, which make it computationally efficient and flexible. Numerical example is also given to show the reduced conservatism and effectiveness of the proposed result in this paper.

REFERENCES

1. Guo, S. J. and L. H. Huang, "Stability analysis of Cohen-Grossberg neural networks," *IEEE Trans. Neural Netw.*, Vol. 17, No. 1, 106–117, 2006.
2. Rong, L. B., "LMI-base criteria for robust stability of Cohen-Grossberg neural networks with delay," *Phys. Lett. A*, Vol. 339, 63–72, 2005.
3. Xu, S. Y. and J. Lam, "A new approach to exponential stability analysis of neural networks with time-varying delays," *Neural Netw.*, Vol. 19, 76–83, 2006.
4. Xu, S. Y., J. Lam, D. W. C. Ho, and Y. Zou, "Novel global asymptotic stability criteria for delayed cellilar neural networks," *IEEE Trans.Circuits Syst. II*, Vol. 52, No. 6, 349–353, 2005.
5. Zhang, Q., X. Wei, and J. Xu, "Global asymptotic stability of Hopfield neural networks with transmission delays," *Phys. Lett. A*, Vol. 318, 2003.

Thermal Stress Analysis of Functionally Graded Cemented Carbides

Ziqian Huang[1], **Haitao Cai**[2], **and Baiyun Huang**[1]

[1]State Key Laboratory of Powder Metallurgy, Central South University
Changsha 410083, China
[2]School of Mathmatical Sciences and Computing Technology, Central South University
Changsha 410083, China

Abstract— A new numerical model for cemented carbides with cobalt gradient has been developed. In this study, carburization is used to form cobalt gradient structure with optimized surface characteristics. Finite-element formulations incorporating continuous and smooth spatial variation in the composition and properties of the cemented carbides are used to simulate the evolution of elastoplastic response of functionally graded cemented carbides (FGCC) due to thermal loading. The geometry of specimens is axisymmetric solid cylinder with two-dimension gradient. The constitutive elastoplastic response is developed by constraint factors. The plasticity constraint factor is used to determine the stress-strain relation of plastic deformation. Different shapes of cobalt rich zone are considered to analyze the evolution of thermal stresses. The calculated results show arc transition of gradient in the corner has a good agreement with XRD measurement.

1. INTRODUCTION

One of the latest developments in the field of hardmetals and cermets used for mineral cutting and rock drilling operations is the establishment of a functional gradient in near-surface areas of cemented carbide tools leading to functionally graded cemented carbides (FGCC) [1–4]. Such a functional gradient is a specific smoothly varying distribution of phases and/or element composition to provide a highly resistant surface region which withstands the interaction of the cutting tool with the workpiece at high temperatures and which smoothly attains the hardmetal microstructure. In fact, carburizing and nitriding of metals are examples of attempts to produce functionally graded materials with optimized surface characteristics, where compositional gradients arising from the diffusion of C or N from the surface to the bulk provides a gradual change in properties. Advances in materials processing have now provided a variety of technologically viable methods to produce functionally graded materials (FGMs). Such techniques include chemical vapor deposition, physical vapor deposition, powder metallurgy, thermal spray, and self-propagating high-temperature combustion synthesis [5–7]. The evolution of thermal stresses and the associated mechanical response are influenced by both the geometry of the component and the gradients in composition and properties.

In the 1980s and 1990s, Sandvik introduced new materials that were described as "the most significant innovation in the history of cemented carbide" since the early 1950s [8–10]. The patented dual phase (DP) technology permits a simultaneous increase in both the wear resistance and toughness of cemented carbide, or either property independently. The DP materials were achieved by changing the distribution of the cobalt within the insert, i. e., the formation of a cobalt-gradient structure. For WC-6Co alloy, there was a low cobalt content of around 3% at the surface to provide the high wear resistance required to penetrate the rock; a middle layer with a cobalt content of around 10% to give the insert high toughness and a core with around 6% cobalt to give a balance of properties. This gives a hardness distribution across the button which is high at the outside and the core, falling off in the middle, cobalt rich zone. The DP grade cemented carbides showed an increase of 40% in service lifetime compared with conventional materials.

Thermally induced elastoplastic analyses of graded composites have also been examined in a few earlier studies dealing with finite element simulations (e. g., [11–13]) where the graded region was discretized into strips of homogeneous compositions of the composite. Since a large number of elements are needed to simulate the continuous change in mechanical properties of materials, especially for modeling three dimensional compositional gradients, this approach is time-consuming and less effective. A.E. Glannakopoulos used finite-element formulations for continuous and smooth variation in the composition to simulate the elastic response by the ratio of stress to strain transfer [14]. However the ratio of stress to strain transfer was simply extrapolated to simulate the plastic deformation. Actually the constraint factor of plasticity deformation is varied with the volume fraction of the ductile phase. J.Rohde et al used "functionally graded elements" to determine

the effects of gradient zone on crack propagation [15]. But they didn't consider the effects of plastic deformation. When the cobalt concentration is equal to or greater than 0.25 (vol%) or higher, there should be plastic flow in the hard alloys in temperature cooling from fabrication temperature. Plastic flow has an important effect on thermal stress redistribution and relaxation.

The purpose of this paper is to develop the constitutive response of two-dimension gradient inserts and obtain the distribution of residual macrostresses and stress gradient. Although it is difficult to obtain the stress distribution with experimental method, the compressive stresses in the surface zone can be achieved with X-ray diffraction measurement. The numerical results are used to compare with the XRD measurement. In our work, carburization was used to form cobalt gradient structure. The constitutive response for continuous and smooth variation of material properties is developed by constraint factors. The ratio of strain transfer was used to simulate the elastic deformation and plasticity constraint factor varied with the volume fraction of the cobalt phase was used to simulate the plastic deformation. Material properties are smooth functions of volume fraction as a smooth function of position. The field quantities of calculated results change with the function of cobalt gradient and temperature difference. Good agreement is obtained between calculated results and XRD method.

2. NUMERICAL FORMULATION AND MATERIAL MODEL

As already mentioned, a model for continuous and smooth spatial variations in the composition and properties of materials has been used. The model is available in the MSC. Marc2005 software.

We consider small strain problem of axisymmetric solid cylinders. In the solid cylinder, there is two-dimension gradient of composition in axial and radial directions. Uniform temperature cooling and heating were assumed. The geometry represents the most realistic geometrical conditions in hardmetal tools such as anvil and gradient buttons. The geometry is given as

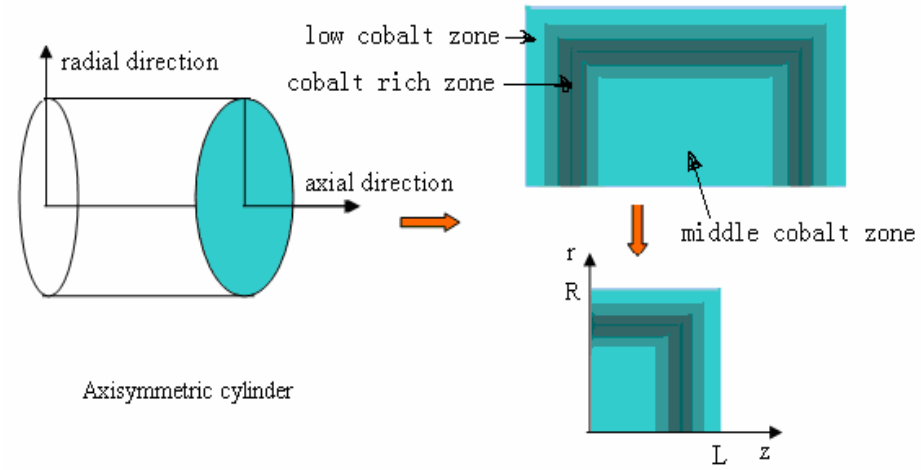


Figure 1: Axisymmetric solid cylinder model.

In the axisymmetric calculations, a cylindrical coordinate system (r, θ, z) is introduced and all field quantities are taken to be independent of θ.The cylinder analyzed numerically is $0 \leq r \leq R$ and $0 \leq z \leq L$, for which the L to R ratio is 1. Free boundary conditions were assumed. On the external boundary of the cylinder,

$$T_r = T_z = 0 \quad \text{on} \quad z = L \quad \text{and} \quad r = R$$
$$u_z = 0 \quad \text{on} \quad z = 0 \quad \text{and} \quad u_r = 0 \quad \text{on} \quad r = 0.$$

The effective values of the coefficient of thermal expansion, α, and Poisson's ratio, ν, of the compositionally graded solid cylinder were computed from a volume fraction-based interpolation. In this way

$$\alpha = f_1\alpha_1 + f_2\alpha_2, \quad \nu = f_1\nu_1 + f_2\nu_2, \quad f_1 + f_2 = 1 \tag{1}$$

where f_i is the volume fraction of material $(i)(0 \leq f_1, f_2 \leq 1)$, subscript 1 and 2 refer to cobalt and tungsten carbide, respectively. Here the isotropic properties of two phases are assumed. The

approximate character of the serial rule of mixture in equations is obvious. The effects of internal constraints at the level of the microstructure of the composite are not taken into account. Here we consider the two-phase formulation for tungsten carbide and cobalt. Actually the specimens have three phases: WC, Co and η phase. We regarded the η phase as WC phase so as to simplify the formulas of constitutive response. The subsequent calculated results show this hypothesis is reasonable. There is a good agreement between the computational results and experimental results.

The elastoplastic response of the mixed WC-Co FGCCs were assumed to be amenable to a rate-independent formulation within the context of the incremental theory of plasticity. An associative Mises yield condition and isotropic hardening were assumed. Isotropic linear elastic response was assumed for tungsten carbide. The composite stress-strain curves were modeled by using the intermediate law of mixture, as originally shown for cemented carbides by Tamura et al., [16] and later adapted for FGMs by Williamson et al., [11]. The FGCCs is treated as isotropic composites for which the uniaxial stress, σ, and strain, ε are related to the corresponding average uniaxial stress and strain of the two constituent phases as

$$\sigma = f_1\sigma_1 + f_2\sigma_2, \quad \varepsilon = f_1\varepsilon_1 + f_2\varepsilon_2 \tag{2}$$

and the ratio of strain transfer, p, is taken to be

$$p = \frac{\varepsilon_1}{\varepsilon_2} \tag{3}$$

If $E_1 < E_2$ is assumed, we have

$$1 \le P \le \frac{E_2}{E_1} \tag{4}$$

Obviously, equal strain indicates that p equals to 1 and equal stress denotes that p equals to E_2/E_1.

From the Equations (2) and (3), effective E is derived explicitly as

$$E = \frac{f_1E_1p + f_2E_2}{f_1p + f_2} \tag{5}$$

where E_i is the Young's modulus of material (i). The value chosen for p is 3, which is close to experimental value of WC-Co alloy. In general, the value of p depends on the internal constraints arising from specific microstructural dispersion of the two phases, the thermomechanical properties, and the residual stresses from processing. Numerous formulae, in closed forms and with bounds, exist for the thermoelastic response of composites which can be used with appropriate modifications in the context of compositionally graded zones. Experimental results show the value of p has a little change with the volume fraction of cobalt phase. Nevertheless, the cobalt concentration we considered is around between 0.03 and 0.3, the difference of p variation can be ignored and has little effects of the elastic deformation.

In 1965, Drucker and Bulten studied the constraint behaviors for tungsten carbide to cobalt phase and microstructure problems [17]. They considered deformation behaviors of alloys were mainly depended on the interactions of components and proposed a concept of plasticity constraint factor. Constraint factor means yield stress ratio for alloy to ductile matrix. The values of constraint factor depend on thickness of cobalt layer and particles of tungsten carbide. According to their investigations, constraint factors of WC-Co alloy show as follows (initial yield stress unit is pound/in^2, here we changed to MPa).

Table 1: Constraint factor and yield stress in different cobalt content.

Cobalt content	0.10	0.19	0.25	0.31	0.37	0.50	0.70
Constraint factor	5.92	3.21	2.50	2.07	1.79	1.46	1.05
Yield stress (MPa)	1656	897	690	572	496	400	289

We obtain the function of constraint factor versus cobalt concentration by means of least square fitting (shown in Figure 2). According to the definition of plasticity constraint factor and the function by least square fitting, the initial yield strength of WC-Co alloy, σ_0, is given by

$$\sigma_0(r,T) = \left(6.25e^{-2.99f_1}\right)\sigma_{01}(T) \tag{6}$$

Here f_1 is cobalt volume fraction as a function of radial position, σ_{01} is the cobalt yield strength varied with the temperature. The parameters 6.25 and -2.99 are obtained from least square fitting. The stress-strain relation of plastic deformation is given by

$$\sigma = \sigma_0(1 + \varepsilon/\varepsilon_0)^N \tag{7}$$

N is the strain hardening exponent. ε_0 is yield strain of alloy. The values chosen for N, ε_0 is 0.2, 0.0001, respectively. The uniaxial plastic behavior of the compositionally graded composite is derived as shown in Figure 2.

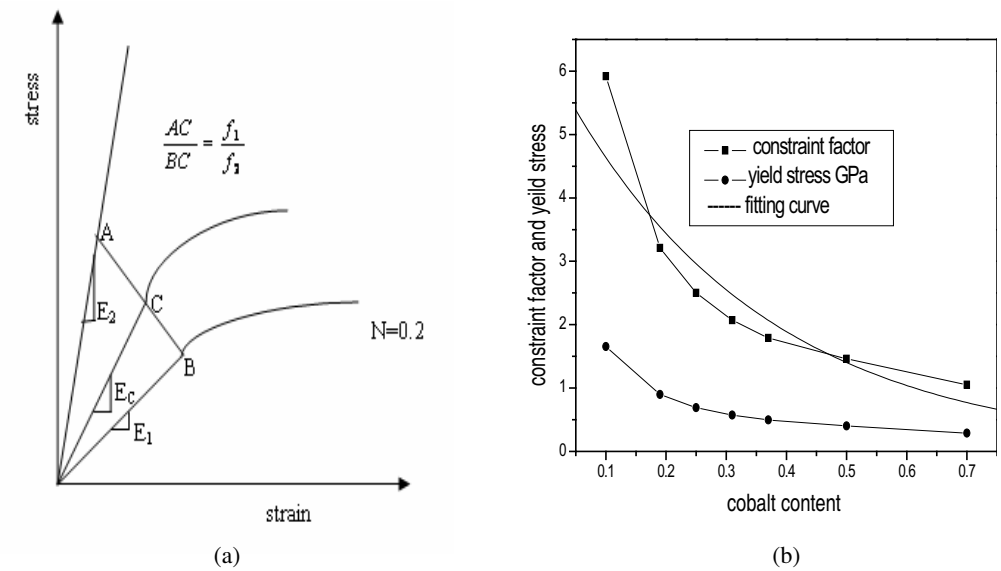


Figure 2: (a) Schematic representation of the stress-strain curve representing the constitutive response of the WC-Co composite as a function of the metal phase, f_1, (b) constraint factor for tungsten carbide to cobalt phase and yield stress vs cobalt content.

Table 2: Temperature-dependence of the thermomechanical properties of α_1, α_2-WC and γ-Co.

(a) WC (α_1, α_2)

E (GPa)	ν	$\alpha(^\circ C^{-1})$
696	0.22	$3.2\times10^{-6}(0^\circ C)$
		$4\times10^{-6}(400^\circ C)$
		$6\times10^{-6}(800^\circ C)$

(b) γ-Co

E (GPa)	ν	$\alpha(^\circ C^{-1})$
207	0.32	$10\times10^{-6}(0^\circ C)$
		$14\times10^{-6}(400^\circ C)$
		$16\times10^{-6}(800^\circ C)$

(c) γ-Co

Temperature	σ_y(MPa)	σ_u(MPa)	$\varepsilon_u(\%)$
0°	200	700	5
400°	100	300	10
800°	72	170	13

From the processing point of view, the above constitutive relations are approximations with respect to the constitutive volume fraction of the metal, f_1, which in fact can be controlled from processing. The formulation is complete, once f_1 is known as a function of the position. Due to

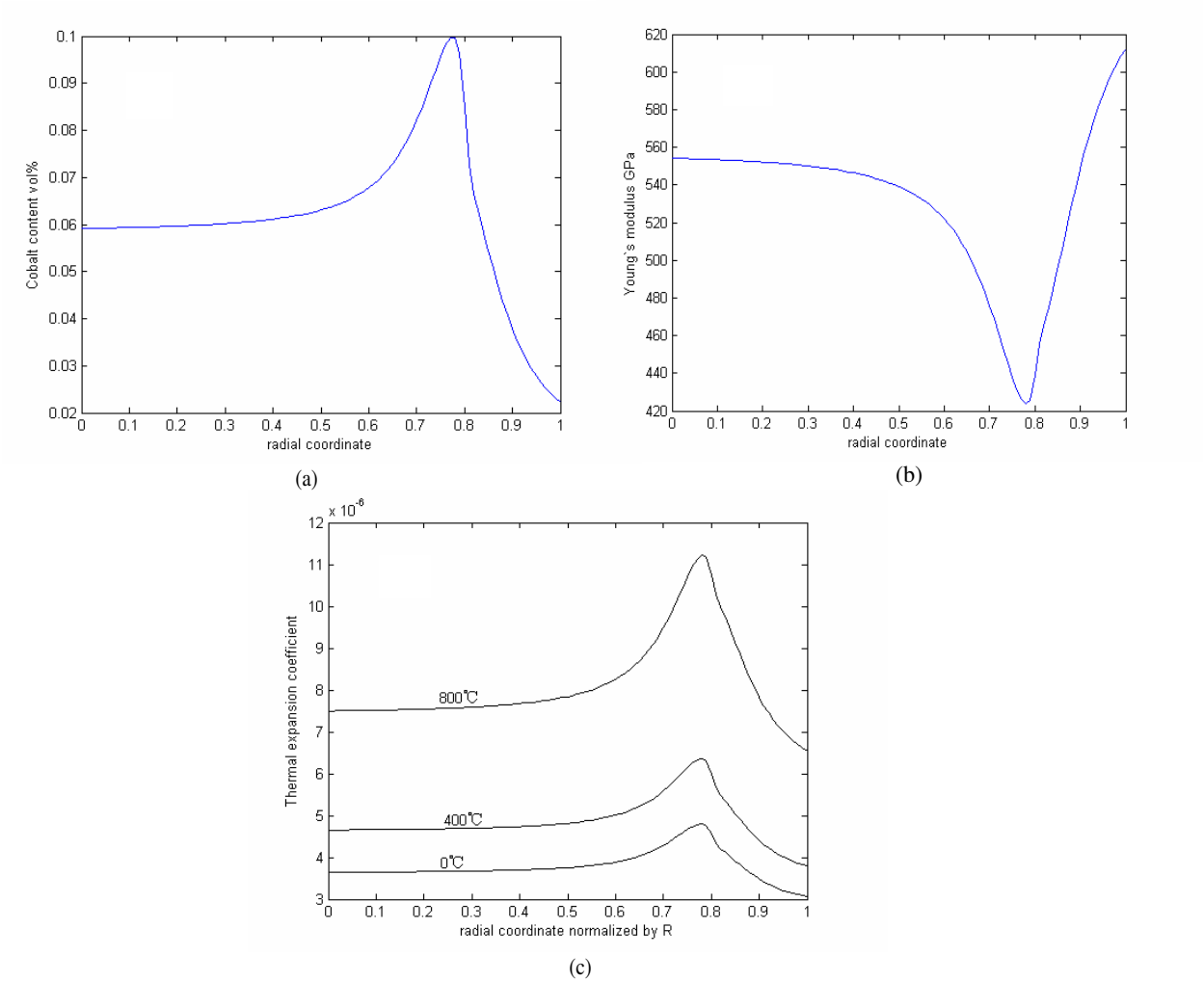


Figure 3: Schematics of materials properties: (a) cobalt content vs radial position, (b) Young's modulus vs radial position, (c) coefficient of thermal expansion vs radial position and temperature.

the irreversible nature of plastic strains, there is no unique theory to characterize the elastoplastic response of compositionally graded WC-Co composites with gradual transitions in phase mixtures.

In our work, carburization used to form the gradient structure in near-surface areas. Firstly, powder mixture with carbon deficient WC-Co alloy was prepared and pressed. Secondly, WC-Co alloy with η phase was prepared by presintering the powder mixture compact. These specimens were carburized with graphite powders to form the gradient structure. The basic mechanism is decomposition of η phase and in-diffusion of cobalt phase under capillary force. The resulting microstructure indicated that all the η phase in the alloy can be fully decomposed by carburization and formed WC(α)+Co(γ): $W_3Co_3C+2C \rightarrow 3WC+3Co$.

According to the experimental results for the distribution of cobalt concentration, we constructed a function to describe the distribution of cobalt content (i. e., the volume fraction f_1 as a function of distance r)

$$f_1 = \frac{m}{100 + (r-n)^2} - k \cdot a\tan(r-n) + l, \quad 0 \le r \le R. \tag{8}$$

The constants, m, n, k, l, represent the modified coefficients of Cobalt concentration. By varying the values of m, n, k, l, we could obtain the different distribution of cobalt content. The factor, m, was used to adjust the nominal cobalt content, n to adjust the position of cobalt rich zone, k used to adjust the slope of cobalt gradients and l to modify the least cobalt content to avoid the appearance of negative value of cobalt content. When m, n, k, l equal to 5, 0.8, 0.015, 0.035, respectively, we obtain the curve of cobalt content distribution (shown in Figure 3(a)). The corresponding distribution of Young's modulus and coefficient of thermal expansion are also shown in Figures 3(b), (c). The majority of the simulation for FGCCs model system incorporated the variation of mechanical

properties with temperature as show in Table 2.

3. NUMERICAL RESULTS

3.1. Thermal Stress Distribution for Orthogonal Gradient in the Corner

The geometry of specimens is solid cylinders with two-dimension gradient. Orthogonal intersection for axial and radial gradient was considered in the corner. The radius and height of the cylinder are 5 mm and 10 mm, respectively. The carburizing time is 1.5–2h. The thickness of cobalt gradient zone is around 1 mm. Uniform temperature cooling and free boundary condition are assumed. Figure 4 shows the distribution of thermal stresses. From Figure 4(a), it is found that equivalent Von Mises stress mainly concentrates in the cobalt gradient zone and it is quite small in the core. There is a big stress gradient from the cobalt rich zone to the core. The maximum value of equivalent Von Mises stress is 241 MPa and the position is in the surface zone. There are two peaks of equivalent Von Mises stress which lies in the cobalt rich zone and the surface zone. The picture also shows the equivalent Von Mises stress in the corner is smaller than the counterparts of the same thickness. From the Figure 4(b), it can be seen that there forms compressive stresses in the surface zone and tensile stresses in the cobalt rich zone after a temperature cooling from initially stress-free temperature of 800°C to 0°C. The maximum value of surface compressive stress is −254 MPa. The compressive stresses are beneficial to inserts lifetime. When the temperature increase from initial stress-free temperature of 0°C to 800°C, there forms the opposite distribution of thermal stresses (shown in Figure 4(c)). Since there is elastic deformation only, it is a reversible process. Actually

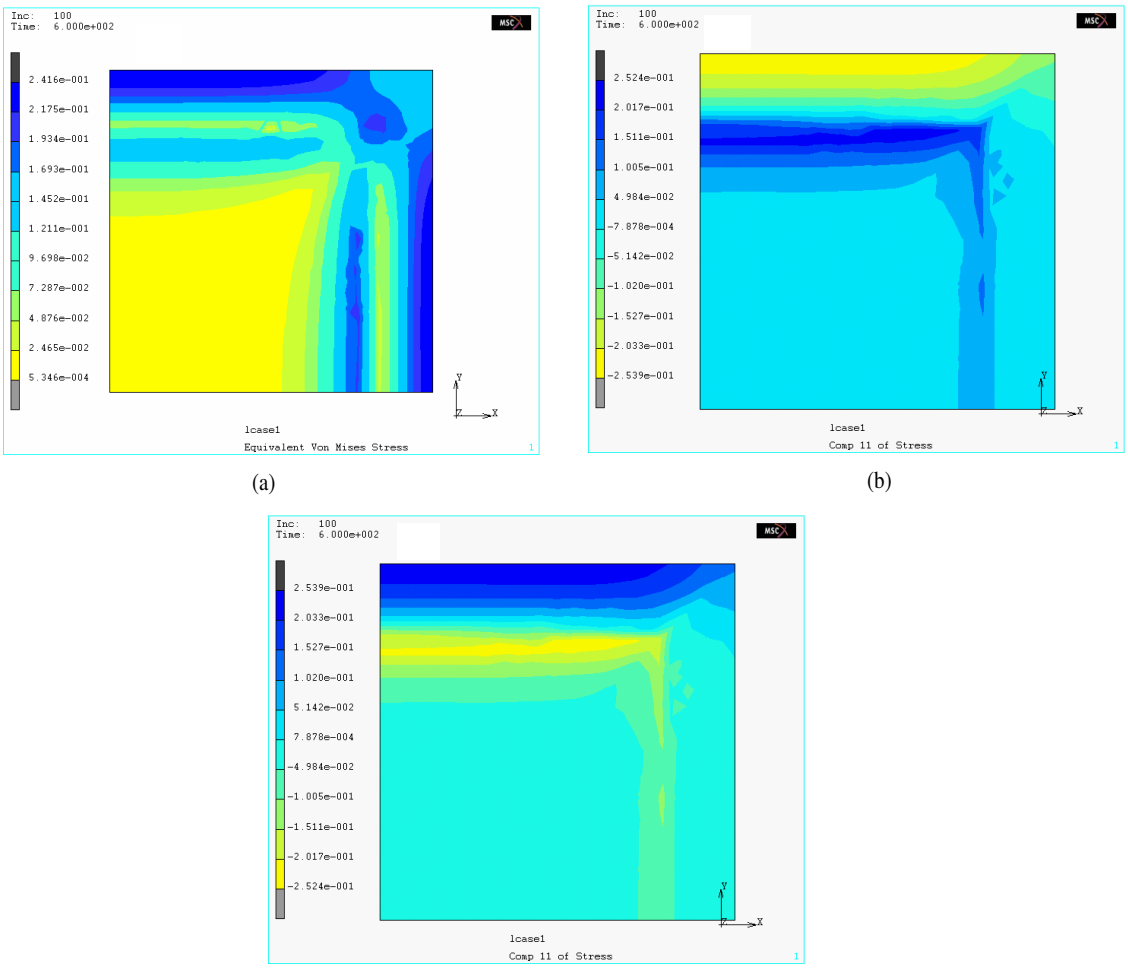
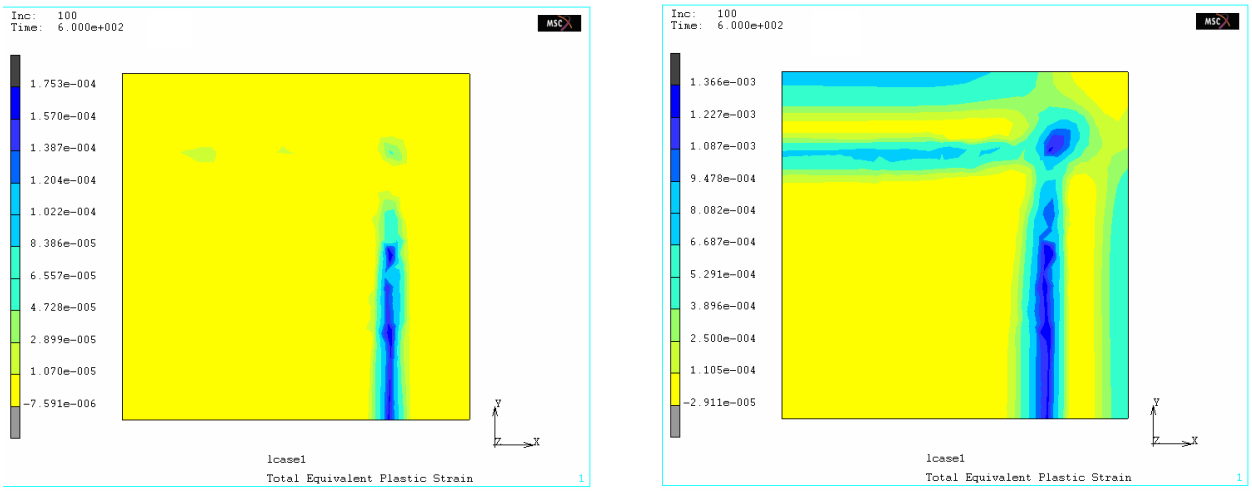


Figure 4: Stress distribution with thermal loading (stress unit: GPa): (a) equivalent Von Mises stress after a temperature cooling from initial stress-free temperature of 800°C to 0°C, (b) distribution of axial stress after a temperature cooling from initial stress-free temperature of 800°C to 0°C, (c) distribution of axial stress after a temperature heating from 0°C to 800°C.

temperature cooling and heating can be regarded as loading and unloading process.



(a) (b)

Figure 5: Distribution of total equivalent plastic strain: (a) after a temperature cooling from 800°C to 0°C, (b) after a temperature heating from 0°C to 800°C.

3.2. Effects of Plastic Flow

When the cobalt concentration difference in the specimens is equal to or greater than 0.3, there is distinct plastic flow in the cobalt gradient zone. Here cobalt concentration difference means the maximum value of cobalt content in cobalt rich zone minus the minimum value in the surface zone. Plastic deformation has a significant effect on the redistribution of thermal stress. Furthermore, there are different responses between the temperature cooling and heating for the material properties varied with temperature. Table 3 show the different calculated results in different cobalt concentration difference (ΔCo) and different thermal loading (cooling or heating). Figure 5(a) shows there is distinct strip shaped plastic flow in cobalt rich zone of axial gradient after a temperature drop from initial stress-free temperature of 800°C. Figure 5(b) shows there is distinct plastic flow in cobalt rich zone of axial and radial gradient when the temperature increase from 0°C to 800°C. The maximum value of total equivalent plastic flow as cooling down is 0.00018, but the maximum value as heating up is 0.0014.

Figure 6 shows the effects of plastic flow on the redistribution of thermal stress. It can be seen

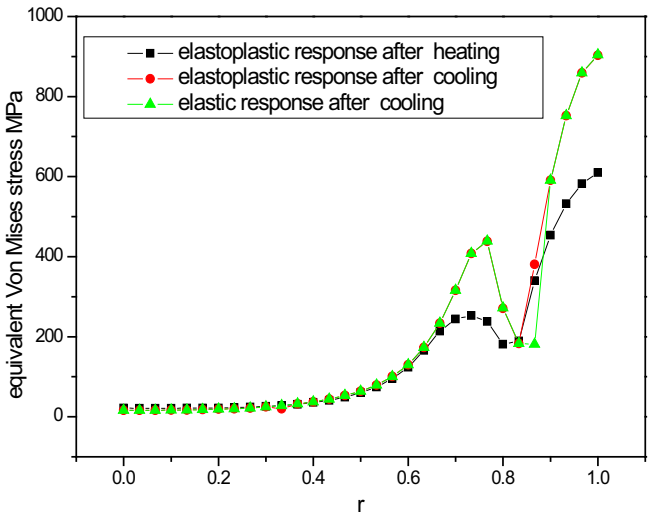


Figure 6: Distribution of equivalent stress along radial direction in different constitutive relations of materials.

that the difference between elastic and elastoplastic deformation after cooling is quite small, but there is a large difference of stress distribution between cooling and heating and between elastic and elastoplastic deformation. The difference of equivalent Von Mises stress between elastic and elastoplastic deformation is 20 MPa as temperature cooling and 530MPa as heating up.

Table 3: Responses of different cobalt concentration difference and thermal loading.

ΔCo	ΔT(°C)	Equivalent stress max (MPa)	Axial stress max (MPa)	Total plastic strain max
0.08	−800	241	258/−257	0
	+800	241	257/−258	0
0.3	−800	904	752/−957	0.00018
	+800	616	672/−561	0.0014

Remark: Co indicates the cobalt concentration difference. −800 denotes temperature cooling down from 800°C to 0°C, +800 means temperature heating up from 0°C to 800°C.

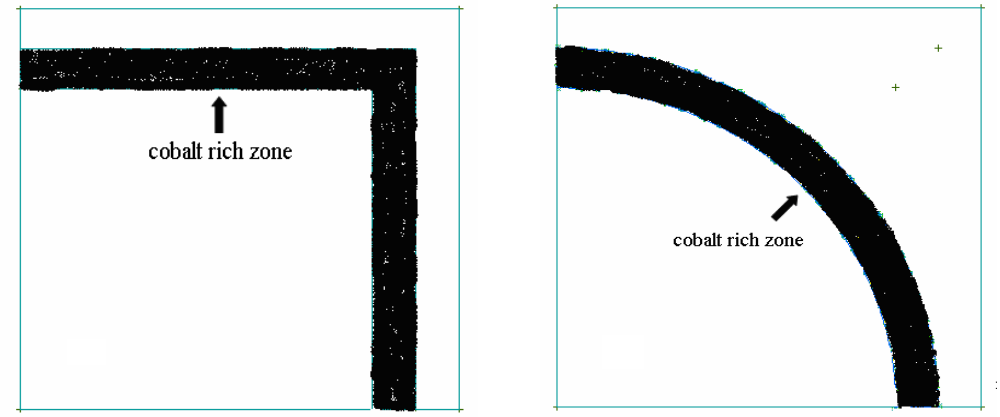


Figure 7: Different shapes of cobalt rich zone: (a) orthogonal gradient in the corner, (b) arc transition gradient.

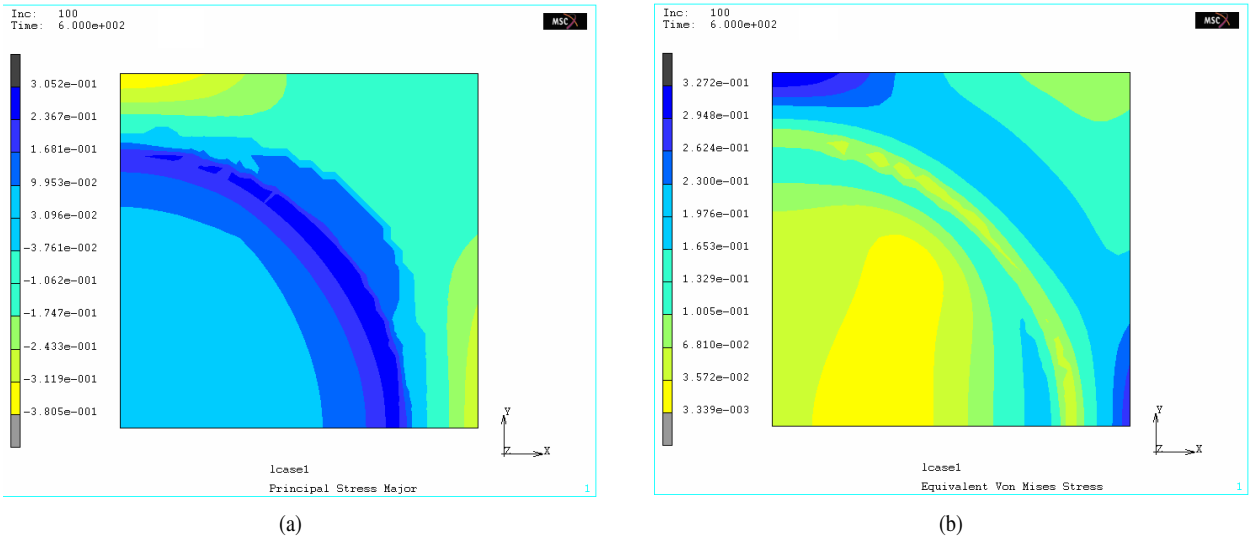


Figure 8: Stress distribution of arc transition gradient after a temperature drop from 800°C to 0°C: (a) principal stress major, (b) equivalent Von Mises stress.

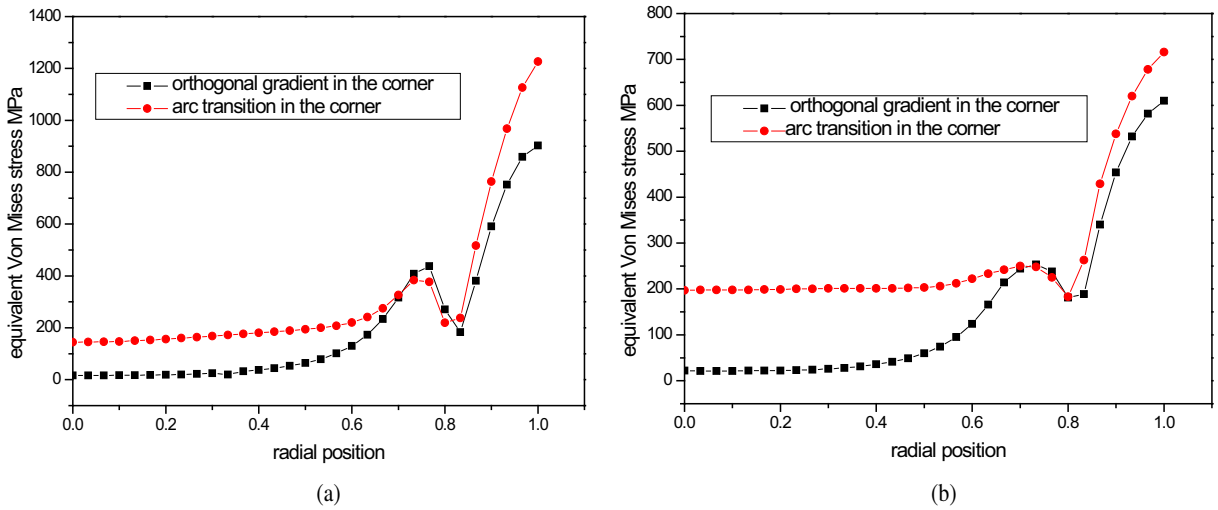


Figure 9: Comparison between orthogonal gradient and arc transition: (a) distribution of equivalent Von Mises stress from the core to the surface after a temperature cooling from 800°C to 0°C, (b) distribution of equivalent Von Mises stress from the core to the surface after a temperature heating from 0°C to 800°C.

3.3. Effects of Arc Transition in the Corner

In the above calculations, orthogonal gradient in the corner for the distribution of cobalt rich zone is assumed. In fact, experimental results show it is arc transition in the corner (shown in Figure 7). X-Ray diffraction was used to determine the surface residual stresses that develop in a functionally graded WC-Co composite. XRD stress measurement show the surface compressive stress of WC 201 is around −377 MPa. There is a difference of around 100 MPa between our computational results and the XRD results. Therefore the model of 2D-gradient with arc transition in the corner is used to determine the distribution of thermal stresses. Figure 8(a) shows the surface compressive stress is −380 MPa and the computational results have a good agreement with the results of XRD stress measurement. The surface compressive stress concentrate in the middle part of solid cylinders and the center of both cylinder ends. Figure 9 shows the distribution of equivalent Von Mises stress after cooling and heating. It is apparent that thermal stress values with arc transition are greater than the values in orthogonal gradient (shown in Figure 9).

4. CONCLUSIONS

1. The thermal residual stresses mainly concentrate in the cobalt gradient zone and in the core with middle cobalt content the stresses are very small. When temperature drops from the processing, there forms compressive stresses in the surface zone and tensile stresses in the cobalt rich zone. When temperature increases there forms the opposite distribution of stresses.

2. When temperature increase from 0°C to 800°C and the cobalt content is higher, plastic flow has a significant effects on redistribution of thermal stress reduction.

3. The arc transition gradient in the corner has greater values of residual stresses than the orthogonal gradient in the corner.

4. Detailed experiments are also necessary to develop an understanding of diffusion, microstructural stability and creep within the graded material during thermal fatigue. Nevertheless, the present results provide a framework for the design of critical experiments and for the development for more rigorous analytical/numerical models.

ACKNOWLEDGMENT

This work was supported by National Natural Science Foundation of China under contract NO. 50323008.

REFERENCES

1. Lengauer, W. and K. Dreyer, "Functionally graded hardmetals," *J. Alloy Compd.*, Vol. 338, 194–212, 2002.
2. Frykholm, R. and H.-O. Andrén. "Development of the microstructure during gradient sintering of a cemented carbide," *Mater. Chem. Phys.*, Vol. 67, 203–208, 2001.

3. Frykholm, R., M. Ekroth, B. Jansson, J. Agren, and H.-O. Andren, "A new labyrinth factor for modeling the effect of binder volume fraction on gradient sintering of cemented carbides," *Acta Mater.*, Vol. 51, 1115–1121, 2003.
4. Frykholm, R., M. Ekroth, B. Jansson, H.-O. Andren, and J. Andren, "Effect of cubic phase composition on gradient zone formation in cemented carbides," *Int. J. Refract. Met. H.*, Vol. 19, 527–538, 2001.
5. Narasimhan, K., S. P. Boppana, and D. G. Bhat, "Development of a graded TiCN coating for cemented carbide cutting tools-a design approach," *Wear*, Vol. 188, 123–129, 1995.
6. Su, Y. L., S. H. Yao, C. S. Wei, and C. T. Wu, "Analysis and design of a WC milling cutter with TiCN coating," *Wear*, Vol. 215, 9-66, 1998.
7. Bull, S. J., D. G. Bhat, and M. H. Staia, "Properties and performance of commercial TiCN coating, Part I: Coating architeoture and hardness modeling," *Surf. Coat. Tech.*, 163–164, 499-506, 2003.
8. Fischer, et al., US Pat. 4743515, 10 May 1988.
9. Akeraan, et al., US Pat. 5279901, 18 Jan. 1994.
10. Fischer, et al., US Pat. 4820482, 11 April 1989.
11. Williamson, R. L., B. H. Rabin, and J. T. Drake, "Finite element analysis of thermal residual stresses at graded ceramic-metal interfaces, Part I: Model description and geometrical effects," *J. Appl. Phys.*, Vol. 74, 1310–20, 1993.
12. Drake, J. T., R. L. Williamson, and B. H. Rabin, "Finite element analysis of thermal residual stresses at graded ceramic-metal interfaces, Part II. Interface optimization for residual stress reduction," *J. Appl. Phys.*, Vol. 74, 1321–26, 1993.
13. Olsson, M., A. E. Giannakopoulos, and S. Suresh, MIT report 94-3, submitted for publication.
14. Glannakopoulos, A. E., S. Suresh, M. Finot, and M. Olsson, "Elastoplastic analysis of thermal cycling: Layered materials with compositional gradients," *Acta Metal. Mater.*, Vol. 43, No. 4, 1335–1354, 1995.
15. Rohde, J., S. Schmauder, and G. Bao, "Mesoscopic modeling of gradient zones in hardmetals," *Comp. Mater. Sci.*, Vol. 7, 63–67, 1996.
16. Tamura, I., Y. Tomota, and H. Ozura, "Strength and ductility of Fe-Ni-C Alloys composed of austenite and martensite with various strength," *Proc. Third Int. Conf. Strength of Metals and Alloys*, Vol. 1, 611–615, Institute of Metals, Cambridge, 1973.
17. Drucker, D. C., *Hight-strength Materiak*, New York, 1965.

Inversion of the Wideband High-frequency Electromagnetic Data

S. T. Zheng[1], **Z. F. Zeng**[1,2], **S. X. Liu**[1], **and F. S. Liu**[2]

[1]College of Geo-exploration Science and Technology, Jilin University, Changchun, China
[2]ARMC, Delaware State University, DE 19901, USA

Abstract— To estimate the underground resistivity and permittivity, the nonlinear least-squares inverse method based on Monte-Carlo algorithms is applied in this paper to inverse the ellipse polarization data of wideband high-frequency electromagnetic (WHFEM) method. From the inversion cases, the nonlinear least-squares inverse method can execute well in the half-space or two layer earth. In order to reduce to unstable and increase convergence, the Monte-Carlo algorithm is used for multiple layers model. In the second part, we also adopt the phasor vectorgraph method for fast inversion of apparent resistivity and permittivity which is helpful to geology mapping or initialization for the least-squares inversion.

1. INTRODUCTION

For detection the low resistivity soil or layers contamination only with permittivity contrast in shallow depth, traditional electromagnetic method and GPR which usually does not penetrate deep enough, are both not efficiency in application. The high-frequency electromagnetic system operated at frequencies between GPR and traditional electromagnetic method has been increased popularity in near-surface geophysics especially in environment and engineering field. So the WFHEM data interpretation is urgent for field application.

The methods such as impendence method [1], neural network method [2] and least-squares method [3] are widely applied in inversing the data of EM detection. But the surface impendence method was widely used at low frequencies such as magnetotelluric (MT) and controlled-source audio-frequency magnetotelluric (CSAMT). A major advantage of these techniques is stable because the impedance response is plane wave and independent of the incident angle. One-dimensional layered-earth inversion is performed by neural network method using ellipticity data provided by Ralf Birken (1997) [2], the results of the sensitivity analysis are dependent on the training of several neural networks to determine the electrical properties of the subsurface and time consuming. A nonlinear least-squares method, using initial models with to fit the tilt angle and the ellipticity of magnetic fields and get resistivity as well as dielectric permittivity, is presented by Steward et al., (1994) [3].

The results using the complex image theory [5] and HFS Q-method [6] for the approximate inversion solution indicate that both the approximate inversions are more than an order-of-magnitude faster than numerical integration inversion. However, the approximate inversions could not always provide accuracy results for deeper layer parameters.

In this paper, a nonlinear least-squares method using high density linear filtering algorithm for forward modeling [7] which is an order-of-magnitude faster than numerical integration such as Chave's (1983) [4], Gauss integration is presented. And we get the equivalent accuracy as numerical integration. The Monte Carlo algorithm is applied to improve the convergence of multi-layer inversion. The phase vector graph [8] method for fast inversion of apparent resistivity and permittivity which is helpful to geology mapping or initialization for the least-squares method is also presented.

2. TWO LAYER CASE BY LEAST-SQUARES METHOD

According to Ward and Hohmann (1988) [9], vertical magnetic H_z and radial magnetic H_r expressions of the layer earth with loop source can be converted into Hankel function integration as followings:

$$f(r) = \int_0^\infty K(\lambda) J_n(r\lambda) d\lambda \tag{1}$$

To calculate the formula (1), the convenient linear filtering algorithm loses its accuracy in our high frequency. Chave's numerical integration although possesses enough accuracy, but executes too slowly for inversion. So the linear filtering algorithm with high sampling density where the filtering

coefficients is calculated by well known Johansen's algorithm et al., (1979) [10] with ln10/100 of sampling value. Then follow Bruce D. Smith et al., [11], polarization ellipticity can be calculated easily through the horizontal magnetic H_z and radical magnetic H_r.

As the following two layer case, in each station point, there are three VMD R-T coils offset r=1, 2, 4 m respectively with 35 frequencies point among 1 MHz∼50 MHz, and both transmitter and receiver loop are set by radius a=0.15 m on the earth surface.

Table 1: Table about true values, initial values and inversing results in a two layer earth.

Parameter	Resistivity(Ω·m)		Dielectric permittivity		Thickness (m)
True values	50	300	5	20	0.5
Initial values	100	500	10	15	0.2
Inversion 1 (exact)	49.65	304.53	4.96	19.92	0.501
Relative error 1	0.7%	1.51%	0.8%	0.4%	0.2%
Initial RMS is 2.71, and after 9 times iterative becomes 8.28×1.0^{-3}					
Inversion 2 (10% error)	50.93	317.78	4.94	19.84	0.505
Relative error 2	1.86%	5.93%	1.2%	0.8%	1%
Initial RMS is 2.71, and after 11 times iterative becomes 0.193					

In the Table 1, we use the true values term to forward the wide band and high-frequency measurement field data. The Initial values term is the initial model for our nonlinear least-squares inversion. The inversing values 1 term is inversion results without adding any noise, and the inversing values 2 term is inversion results after adding the forwarding data by random 10% noises. Figure 1 shows the forwarding curves with the initial model and true model respectively. Combining Table 1 and Figure 1, it shows that the nonlinear least-squares inversion is implemented with great efficiency and stability.

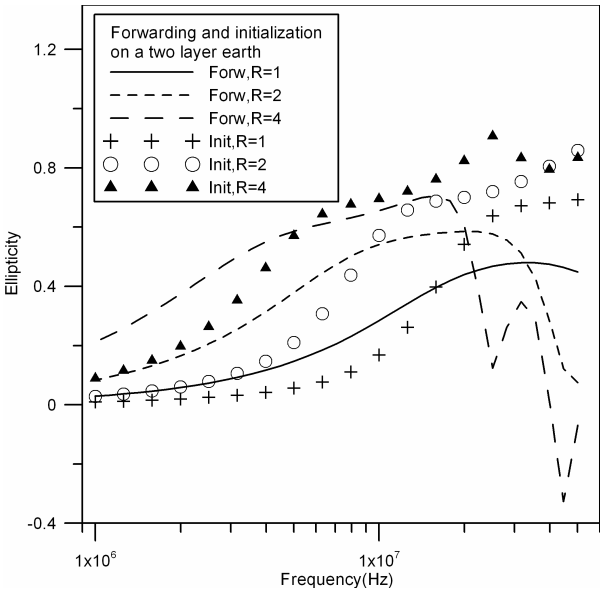


Figure 1: Ellipticity of both initialization and true models on a two layer earth with three R-T coils offset respectively, which each has 35 frequencies point.

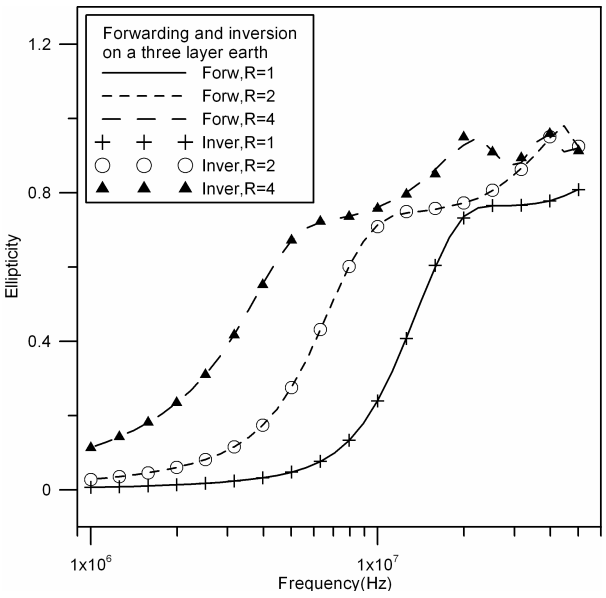


Figure 2: Ellipticity of both true and inversion models on a three layer earth with three R-T coils offset respectively, which each has 35 frequencies point.

3. MULTIPLE LAYER INVERSION RESULT

In the Table 2, we use the true values term to forward the wide band and high-frequency measurement field data for tree layer earth. The Initial values term is the initial model for our nonlinear least-squares inversion. Figure 2 shows the forwarding curves with the inversion model and true

Table 2: Table about true values, initial values and inversing results in a three layer earth.

Parameter	Resistivity (Ω·m)			Dielectric permittivity			Thickness (m)	
True values	300	200	100	25	20	15	2	2
Initial values	200	150	80	20	10	20	1.0	3.0
Inversion values	299.3	301.8	108.4	25	24.5	25.5	0.99	2.1
Relative error	0.23%	50.9%	8.4%	0.0%	22.5%	70%	50%	5%
Initial RMS is 2.55, and after 20 times iterative becomes 0.425								

model respectively, and both curves of two models are almost consistent with each other. Combining Table 2 and Figure 2, it shows that the nonlinear least-squares inversion could solve the upper layer electric parameter well, but produces a big relative error. The inversion is converged to the local minimum value 0.425 and can't be reduced further because of lacking measure information. The problem of local minimum can be resolved by providing more information such as more coil offsets, larger frequency range, and so on, but it's not always possible in practical use. So the Mont-Carlo algorithm is reduced into this paper for multiply layer before the least-squares inversion. For this case, the solver space is separated randomly by 40 groups, and the objective function of each group is calculated, then the group having the least objective function value is regarded as a initial model for the further least-squares inversion. In the Table 3, the initial model term is gained by the Monte-Carlo method. The inversion result proves that the inversion model is almost consistent with the true model except the lower layer dielectric permittivity (because of inadequate deep layer information). Noteworthiness, it requires the less iterative step of this inversion method to avoid jumping into another local minimum.

Table 3: Table about inversing parameters and results in a three earth after method of Monte-Carlo.

Parameter	Resistivity (Ω·m)			Dielectric permittivity			Thickness (m)	
True values	300	200	100	25	20	15	2	2
Initial values (gained by Monte-Carlo)	290	181	90	23	18	16	1.8	2.1
Inversion values	298.8	206.1	109.9	24.9	19.2	13.5	2.001	1.98
Relative error	0.4%	3.05%	9.9%	0.4%	0.40%	23%	0.05%	1%
Initial RMS is 0.18, and after 11 times iterative becomes 0.085								

Table 4: True and inversing resistivity, permittivity with relative error between them.

Resistivity	Permittivity	Apparent resistivity	Relative error	Apparent permittivity	Relative error
210	6	212.606	1.24%	5.939	1.02%
350	12	352.753	0.787%	11.948	0.43%
580	25	580.961	0.166%	25.053	0.212%
700	8	699.333	0.048%	8.040	0.5%
1000	30	1012.498	1.25%	29.812	0.627%

4. PHASE VECTORGRAPH METHOD FOR FAST INVERSION AND A HALF-SPACE CASE RESULT

Phase diagram[10, 12] of the normalized in-phase and quadrature components for the half space model for several resistivity and relative permittivity is usually used for the interpretation of helicopter-borne electromagnetic data. Analogously, we use polarization tilt angle and ellipticity components to make phase diagram as Figure 3 in our paper. The dielectric permittivity and

conductivity of a dielectric conductive half-space are obtained by transforming polarization tilt angle and ellipticity responses using a 2-D interpolation scheme. Table 4 is the half-space case results.

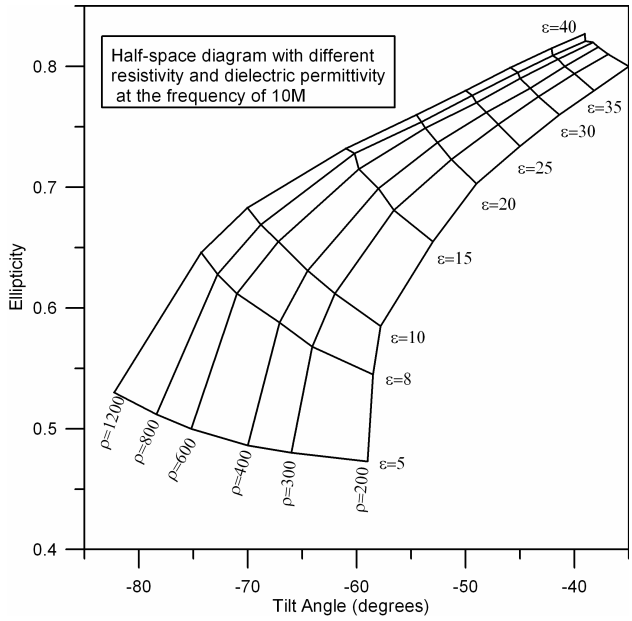


Figure 3: Half-space diagram between polarization ellipticity and polarization tilt angle in the frequency of 10 MHz.

5. CONCLUSIONS

From the case results, we can conclude that the nonlinear least-squares Inversion of the wideband and high-frequency electromagnetic data using the linear filtering with high sampling technique is a efficiency and stable tool for quantitative interpretation, and can be used to mapping the shallow resistivity and dielectric permittivity. The Monte-Carlo algorithm used for multiple layers model avoids local minimum. Phase vectorgraph as an approximate inversion method can solve apparent resistivity and permittivity fast, and is helpful to geology mapping or initialization for the least-squares inversion. Both inversions provide the useful information about layer earth and it will be applied in further engineering projection more increasingly.

ACKNOWLEDGMENT

The research work was supported by NSFC(40474042) of China, and a DoD Grant DAAD 19-03-1-0375 of USA, and the Grant-in-Aid for people who study abroad.

REFERENCES

1. Song, Y., H. J. Kim, and K .H. Lee, "High-frequency electromagnetic impedance method for subsurface imaging," *Geophysics*, Vol. 67, No. 2, 501–510, 2002.
2. Birken, R., "Neural network interpretation of electromagnetic ellipticity data in a frequency range from 1 kHz to 32 MHz," Ph. D. thesis, Arizona University, Tucson, 1997.
3. Stewart, D. C., W. L. Anderson, P. G. Thomas, et al., "Shallow subsurface mapping by electromagnetic sounding in the 300 KHz to 30 MHz range: Model studies and prototype system assessment," *Geophysics*, Vol. 59, No. 8, 1201–1210, 1994.
4. Chave, A. D., "Numerical integration of related Hankel transforms by quardrature and continued fraction expansion," *Geophysics*, Vol. 48, No. 12, 671–1686, 1983.
5. Anderson, W. L., "Approximate inversion of high-frequency electromagnetic soundings using complex image theory," *Geophysics*, Vol. 56, No. 7, 1087–1092, 1991.
6. Anderson, W. L., "Q-factor approximation of electromagnetic fields for high-frequency sounding in the 300 kHz to 30 MHz range over layered media," *Geophysics*, Vol. 60, No. 4, 1253–1258, 1995.

7. Zheng, S. T., Z. F. Zeng, D. G. Zhang, et al., "Calculation and analysis on high-frequency electromagnetic in a layer media," publishing, 2006.
8. Fraser, D. C., J. A. Stodt, and S. H. Ward, "The effect of displacement currents on the response of a high-frequency electromagnetic system," Ward, S. H. ed. *Geotechnical and Environmental Geophysics*, Vol. 2, "Environmental and Groundwater," *Soc. Expl. Geophysics*, 89–95, 1990.
9. Ward, S. H. and G. W. Hohmann, "Electromagnetic theory for geophysical applications," in Nabighian, M. N., Ed., "Electromagnetic methods in applied geophysics," *Soc. Expl. Geophys.*, Vol. 1, 131-311, 1987.
10. Johansen, H. K. and K. Sørensen, "Fast Hankel transforms," *Geophysical Prospecting*, Vol. 27, 876–901, 1979.
11. Smith, B. D. and S. H. Ward, "On the Computation of polarization ellipse parameters," *Geophysics*, Vol. 39, No. 6, 867–869, 1974.
12. Huang, H. P. and D. C. Fraser, "Dielectric permittivity and resistivity mapping using high-frequency, helicopter-borne EM data," *Geophysics*, Vol. 67, No. 3, 727–738, 2002.

Nondestructive Testing in Mechanical Engineering by Wave Inversion

Chien-Chang Lin[1], **Ganquan Xie**[2], **and Jianhua Li**[2]

[1]Center of Computational Nanotechnology, Dayeh University, Taiwan, China
[2]GL Geophysical Laboratory, USA

Abstract— In this paper, we present a new approach for the nondestructive testing for detecting cracks in materials by using wave parameter inversion that includes mechanical and electromagnetic wave parameter. We use singularity properties pulse function and convolution to find the Frechet derivative of the nonlinear response operator on the wave velocity. The Frechet derivative operator is weak ill posed integral geometry form in the manifolds. We use Newton method to solve the nonlinear equation for wave velocity inversion. Using the wave velocity inversion, we can detect the crack and properties anomalous in the materials.

1. INTRODUCTION

In the civil and mechanical engineering, nondestructive testing has great applications. It has been using in the satellite, space shuttle, flight, train, etc for their health and safety. However, most nondestructive testing instruments only used hardware measurement that can be used to detect the large crack existing and can not the small fracture and cracking developing and properties. Therefore, develop fast and high resolution nondestructive testing software is necessary. There are some works in the wave inversions. Professor P.D. Lax' inverse scattering theorem is describe in the his book [1]. Chen and Xie's paper [2] describe a PST method for nondestructive testing. Xie and Chen developed 2D elastic wave velocity inversion in [3]. Xie proposed a new iterative method for the coefficient inverse problem of wave equation [4]. Xie, Lin and Li proposed a new inversion method on the scattering wave propagation [5]. Xie et al proposed GILD for 3D electromagnetic modeling and inversion [6]. Qisu and Xie proposed a parallel computational algorithm for an inverse problem of low frequency: [7] In this paper, we developed a new approach for the nondestructive testing for detecting cracks in materials by using wave velocity inversion. We use singularity properties pulse function and convolution to find the Frechet derivative of the nonlinear response operator on the wave velocity. The Frechet derivative operator is weak ill posed integral geometry form in the manifolds. We use Newton method to solve the nonlinear equation for wave parameter inversion that includes mechanical and electromagnetic wave parameter. Using the wave velocity inversion, we can detect the crack and properties anomalous in the materials.

The organization of this paper is as follows. The wave equation and wave propagation in mechanical materials with crack is described in Section 2. In Section 3, we define a nonlinear response operator on the wave velocity. We derive the Frechet derivative in Section 4. In Section 5, we proposed an iteration to find the wave velocity. The applications are described in Section 6. Finally, we describe conclusion.

2. WAVE PROPAGATION IN MECHANICAL MATERIALS

The wave equation is expressed as

$$\frac{\partial^2 u}{\partial x^2} - \frac{1}{c(x)^2}\frac{\partial^2 u}{\partial t^2} = 0, \tag{1}$$

$$u(x,0) = 0, \tag{2}$$

$$\frac{\partial u}{\partial t}(x,0) = 0, \tag{3}$$

$$\frac{\partial u}{\partial x}(0,t) = \delta(t). \tag{4}$$

where the u in the Equation (1) is wave function $u(x,t)$, x is space variable, t is time variable, $c(x)$ is the wave velocity,

$$c(x) = \sqrt{\frac{\mu(x)}{\rho(x)}} \tag{5}$$

$\mu(x)$ is shear module parameter, $\rho(x)$ is density.

Wave propagation in mechanical materials with crack.

Equation (1) and its initial condition (2) and (3) and its exciting pulse wave on the surface boundary can governing the wave propagation in the elastic or plastic mechanical materials. Figure 1 shows the micro structure with power and crack and micro fractures distribution in a variable transverse axis. We can use macro elastic or plastic wave in $(1) \sim (5)$ to describe the axis with variable transverse and a crack in Figure 2.

We will discuss the wave propagation in the micro fracture shown in Figure 1. The micro molecular dynamic and plastic mixed module should use to describe the mechanical material properties. For simplicity, we can use plastic properties to include the molecular dynamic properties. Therefore, GILD wave propagation modeling can be used to the complex situation.

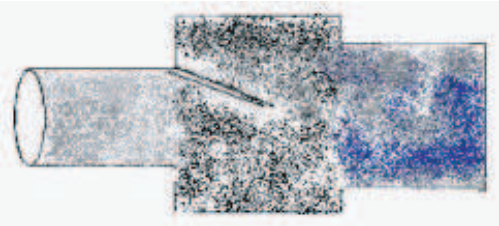


Figure 1: The crack and powder distributions in axis with variable transverse.

Figure 2: The crack distributions in variable transverse section.

3. WAVE VELOCITY INVERSION

3.1. Nondestructive Testing

There are several nondestructive testing instruments. The working principle is that impulses are exciting in some place in the surface boundary. The exciting wave is propagation in mechanical materials or parts. When wave meet the crack edge or crack surface, the reflection and scattering wave will be generated and can be propagate to surface. The instrument can measure the surface wave in time. Researcher can detect the crack from the surface measure wave roughly.

3.2. Measured Surface Response

We use the nondestructive testing instrument to measure the impulse wave response on the surface boundary, $x = 0$, as follows

$$u(0,t) = f(t). \tag{6}$$

From engineering point of view, the $f(t)$ is impulse wave response on the surface. In physical mathematical, we define the (6) to be nonlinear response operator equation in some functional Hilbert or Banach space. In general, we define the Hilbert Sobolev spaces

$$H_1(f) = \left\{ all\, f(t),\ \int_0^\infty \left(f^2 + (f)^2\right) dt < \infty \right\}. \tag{7}$$

3.3. Nonlinear Response Operator on the Wave Velocity

From the wave Equation (1), initial conditions (2) and (3), boundary exiting impulse condition (4), we can solve the forward wave propagation modeling by analytical or numerical methods. We can obtain the wave field $u(x,t)$. Moreover, the $u(x,t)$ is nonlinear functional of the $C(x)$. We write it $u(x,t;C)$.

Let

$$L(C) = \left\{ all\, C(x),\ \int_L c^2(x)\, dx < \infty \right\}. \tag{8}$$

3.4. Nonlinear Wave Velocity Inversion

In this paper, we reduce the three dimensional non destructive testing in Figure 1 and Figure 2 into one dimension wave velocity inversion which is demonstrated in Figure 3.

For any given wave velocity $c(x) \in L(C)$, the wave Equation (1) and initial condition (2) and (3) and the boundary condition (4) map $\mathbf{W}(\mathrm{c})$ is mapping the $c(x)$ into the wave function $u(x,t;c)$

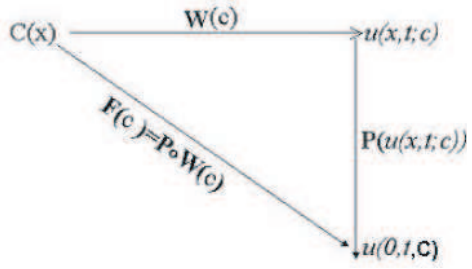


Figure 3: The nonlinear functional commute diagraph of $\mathrm{F(c)=P\gamma W(c)}$.

that is a nonlinear functional on the $c(x) \in L(C)$. The project operator map $\mathbf{P}$ is mapping $u(x,t;c)$ into the boundary surface response $u(0,t;c)$. The nonlinear functional commute diagraph of $\mathrm{F(c)=P\gamma W(c)}$ is shown in Figure 3. The composite map product induce a diagonal nonlinear functional $\mathbf{F(c)}=\mathbf{P\gamma W(c)}=u(0,t;c)$. Therefore, we define the nonlinear function equation

$$F(c) = u(0,t;c) = f(t). \tag{9}$$

The wave velocity inversion is to solve the nonlinear functional Equation (9) to find the wave velocity $c(x)$ from the measured boundary response $f(t)$.

4. FRECHET DERIVATIVE OPERATOR OF F(C)

4.1. Convolution Product

We define the convolution product of two function $f(t)$ and $g(t)$ to be

$$f * g(t) = \int_{-\infty}^{\infty} f(t-\xi)g(\xi)d\xi. \tag{10}$$

4.2. The Properties of the Convolution Product

(a) if $f(t) = 0,\ t < 0$ and $g(t) = 0,\ t < 0$, then

$$f * g(t) = \int_{0}^{t} f(t-\xi)g(\xi)d\xi. \tag{11}$$

(b) for any $f(t)$ defined in $[0,\infty)$, we have

$$f * \delta(t) = f(t). \tag{12}$$

Using time convolution, we derive the Frechet derivative operator of the nonlinear function as follows [3–5]

$$F_c\left(\delta c\right) = 2\int_{0}^{\infty} \frac{\delta c}{c^3}\frac{\partial u}{\partial t} *_t \frac{\partial u}{\partial t}dx \tag{13}$$

5. NEW ITERATION FOR WAVE VELOCITY INVERSION

We describe a new iteration for solving wave velocity inverse problem as follows,

(1) Suppose that we obtained $c_n(x)$ in the before $(n-1)$th iteration,
(2) Upon substituting the $c_n(x)$ in the wave Equation (1) for the $c(x)$, then we solve the wave Equation (1) with initial boundary conditions (2) and impulse boundary condition (4) numerically. We obtained the wave field $u(x,t;c_n)$ and its boundary response $u(0,t;c_n)$.
(3) To solve the following integral equation

$$2\int_{0}^{\infty} \frac{\delta c}{c^3}\frac{\partial u}{\partial t} *_t \frac{\partial u}{\partial t}dx = f\left(t\right) - u\left(0,t;c_n\right) \tag{14}$$

the increment velocity $\delta c(x)$ is obtained.

(4) To update

$$c_{n+1}(x) = c_n(x) + \delta c(x) \tag{15}$$

The steps $(1) \sim (4)$ are build a new iteration for the wave velocity inversion.

6. SIMULATION AND APPLICATION

We designed two models to test our new wave velocity inversion. The first model is shown in Figure 4. In the first model, the velocity function is as follows,

$$c(x) = \begin{cases} 11.0 & 0 \le x < 4 \\ 1.0 & \text{otherwhere} \end{cases} \tag{16}$$

The second model is shown in Figure 5, in the second model, the velocity function is as follows,

$$c(x) = \begin{cases} 6.0 & 0 \le x < 2 \\ 11.0 & 2 \le x < 4 \\ 1.0 & \text{otherwhere} \end{cases}, \tag{17}$$

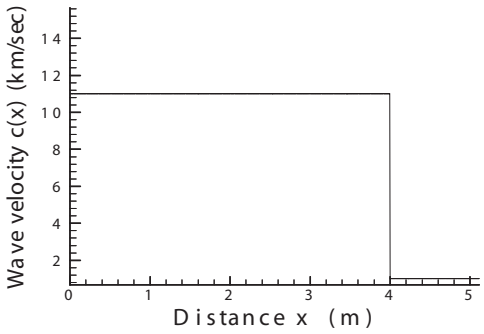


Figure 4: The first velocity model.

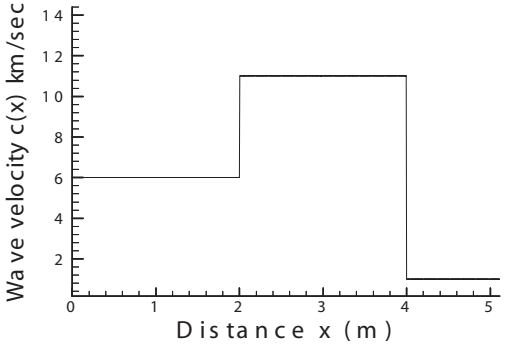


Figure 5: The second velocity model.

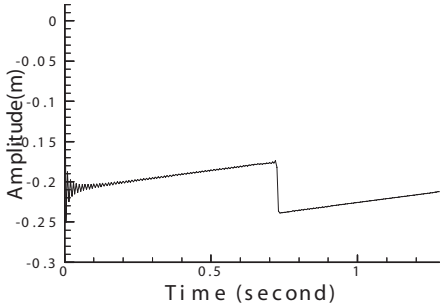


Figure 6: The boundary response of first model.

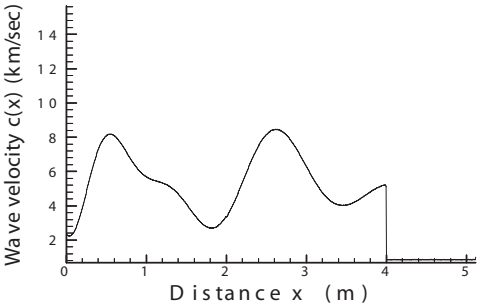


Figure 7: The first iteration velocity $c_1(x)$.

Upon substituting the velocity function $c(x)$ of the first model in (16) into the wave Equation (1), we solve wave Equation (1) with initial conditions (2) and (3) and impulse boundary condition (4) and obtained the wave field $u(x,t)$ and the response $U(0,t) = f(t)$ on the boundary $x = 0$ by numerical method. The $f(t)$ is plotted in the Figure 6. Then, we used using our new iteration in $(1) \sim (4)$ steps in the Section 5. The starting model is one Km per second. The velocity in first iteration is shown in Figure 7, The velocity in 4th iteration is shown in Figure 8, The velocity in 20th iteration is shown in Figure 9, The velocity in 25th iteration is shown in Figure 10, we obtained very accurate wave velocity only using 25 iteration. The residual after 25 iteration is 10^{-9}. Similarly, the second model in (16) and shown in Figure 5. We calculate the boundary response $U(0,t) = f(t)$ in $x = 0$ of the second model. The graph of $f(t)$ is shown in Figure 11. The starting iterative model is setting to background model, say 1 km per second everywhere. After 30 iterations, we obtained an accurate wave velocity in Figure 15. The residual after 30 iterations is 5×10^{-9}. The approximate wave velocity after first iteration is shown in Figure 12.

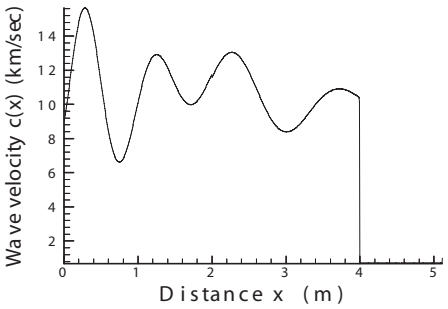


Figure 8: The 4'th iteration velocity $c_4(x)$.

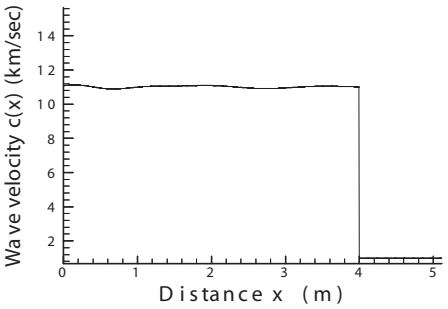


Figure 9: The 20'th iteration velocity $c_{20}(x)$.

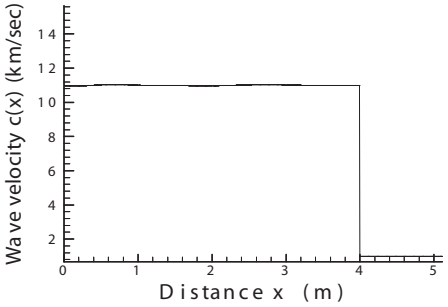


Figure 10: The 25'th iteration velocity $c_{25}(x)$.

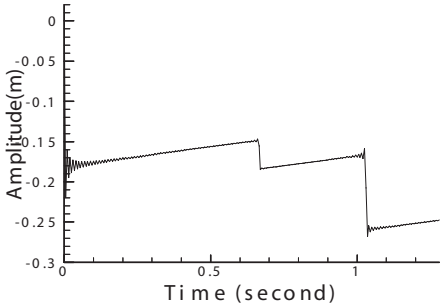


Figure 11: The boundary response of second model.

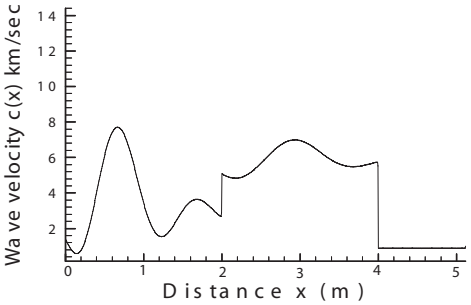


Figure 12: The first iteration velocity $c_1(x)$.

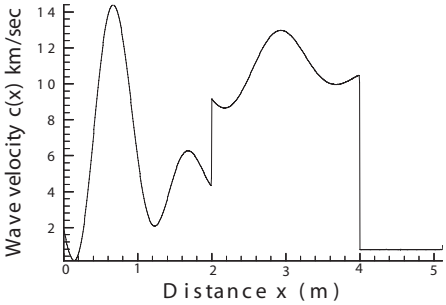


Figure 13: The 4'th iteration velocity $c_4(x)$.

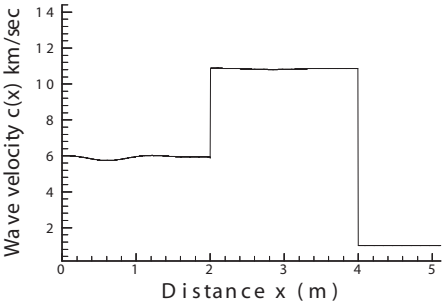


Figure 14: The 25'th iteration velocity $c_{25}(x)$.

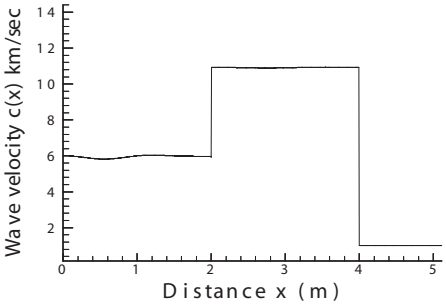


Figure 15: The 30'th iteration velocity $c_{30}(x)$.

Figure 13 shows the wave velocity after 4 iterations. The approximate wave velocity after 25 iterations is shown in Figure 14.

We also test our new wave velocity inverse iteration in many models. All of the tests show our method fast, stable, and high resolution. We will use our method for the project of high speed train vibration and nondestructive testing, macro and micro mechanical nondestructive and micro fracture analysis. Also our wave velocity inverse iteration can be useful for Taiwan earthquake exploration, seismic engineering and high speed train vibration etc.

7. CONCLUSION

The new method for solving the wave velocity inversion is fast, stable, and high resolution. The method can be extended to two dimensional and three dimensional wave velocity inversions. We have used the new inverse method to nondestructive testing and Chi Chi earthquake exploration in Taiwan. In this paper, we developed one dimensional wave velocity inversion for non destructive

testing. The extensive two and three dimensional non destructive inversion will be developing in next papers.

ACKNOWLEDGMENT

Authors thank National Science Council for financial support. Authors thank Professor C. C. Wu and Professor C. S. Yeh at Institute of Applied Mechanics of National Taiwan University for their help.

REFERENCES

1. Lax, P. D and R. S. Philips, *Scattering Theory*, Revised edition, Academic Press, New York, 1989.
2. Chen, Y. M. and G. Xie, "A numerical method for simultaneous determination of bulk modulus, shear modulus and density variations for nondestructive evaluation," *Nondestructive Testing Commun.*, Vol. 1, 125–135, 1984.
3. Xie, G. and Y. M. Chen, "Two dimensional elastic wave inversion," *Applied Numerical Mathematics*, Vol. 1, No. 3, 217–237, 1985.
4. Xie, G., "New iteration for solving coefficient inverse problem," *Communication Pure and Applied Mathematics*, Vol. 39, 307–322, 1986.
5. Xie, G., C. C. Lin, and J. Li, "New inversion method on the scattering wave propagation," submitted to inverse problem, 2002.
6. Xie, G., J. Li, E. Majer, D. Zuo, and M. Oristaglio, "3-D electromagnetic modeling and nonlinear inversion," *Geophysics*, Vol. 65, No. 3, 804–822, 2000.
7. Zou, Q. and G. Xie, "A parallel computational algorithm for an inverse problem of low frequency," *Computer Physics Communications*, Vol. 65, 337–332, 1991.

Particular Properties in the Dielectric Response of Negative-permittivity Scatterers

A. Sihvola

Electromagnetics Laboratory, Helsinki University of Technology
P. O. Box 3000, Espoo 02015–TKK, Finland

Abstract— In this paper, some interesting observations about the peculiarities of the dipole response of spherical scatterers of negative permittivity will be discussed.

The essence of the electromagnetic response of a small scatterer is contained in its field-induced dipole moment, determined by the polarizability which is a function of the geometry and permittivity of the particle. In applications like modeling of materials, the dipole polarizability is the most important of the multipolar parameters. All higher-order multipole fields decay more strongly with distance than those of the dipole.

The normalized polarizability of a small, homogeneous dielectric sphere is

$$\alpha = \frac{\alpha_{\text{abs}}}{\epsilon_0 V} = 3\frac{\epsilon - 1}{\epsilon + 2} \tag{1}$$

where ϵ is the relative permittivity of the sphere (the absolute polarizability is made dimensionless through division by the free-space permittivity ϵ_0 and the volume of the sphere V). Obviously the dielectric response of a small sphere behaves very strangely if the permittivity is allowed to be negative. The polarizability grows without limit if the permittivity approaches the value -2. This is the so-called Fröhlich mode [1]. One could also call it the electrostatic resonance [2].

Negative-valued permittivities are no anomalies. Several metals display such values for optical or infrared frequencies. And indeed, the present interest in so-called metamaterials which are being studied for their possible potential in microwave and higher-frequency imaging applications [3, 4] are often based on the simultaneously negative permittivity and permeability parameters [5]. Such artificial materials can be fabricated by embedding various resonating "molecules" into a neutral matrix.

A generalization of (1) into other spatial dimensions is obvious. Depending on the dimension D of the problem, the "sphere" polarizability is

$$\alpha = D\frac{\epsilon - 1}{\epsilon + D - 1} \tag{2}$$

and therefore, the polarizability of a circle (a two-dimensional sphere) is

$$\alpha = 2\frac{\epsilon - 1}{\epsilon + 1} \tag{3}$$

which has a singularity when the relative permittivity attains the value $\epsilon = -1$.

For a two-layer sphere with core of ϵ_2 and shell of ϵ_1, the polarizability, generalization of (1), reads [6]

$$\alpha = 3\frac{(\epsilon_1 - 1)(\epsilon_2 + 2\epsilon_1) + \beta(2\epsilon_1 + 1)(\epsilon_2 - \epsilon_1)}{(\epsilon_1 + 2)(\epsilon_2 + 2\epsilon_1) + 2\beta(\epsilon_1 - 1)(\epsilon_2 - \epsilon_1)} \tag{4}$$

where $\beta = a_2^3/a_1^3$ is the volume fraction of the core of the total volume.

The corresponding polarizability value for the two-dimensional case (concentric circles, i. e., cross-cuts of cylinders) obeys the formula

$$\alpha = 2\frac{(\epsilon_1 - 1)(\epsilon_2 + \epsilon_1) + \beta(\epsilon_1 + 1)(\epsilon_2 - \epsilon_1)}{(\epsilon_1 + 1)(\epsilon_2 + \epsilon_1) + \beta(\epsilon_1 - 1)(\epsilon_2 - \epsilon_1)} \tag{5}$$

where now β refers to the fractional area a_2^2/a_1^2.

A look at the relation (3) gives rise to an interesting observation concerning two-dimensional scatterer responses. The polarizabilities of "complementary circles," meaning circles with inverse permittivities (ϵ and $1/\epsilon$) are opposite numbers:

$$\alpha_{1/\epsilon} = 2\,\frac{1/\epsilon - 1}{1/\epsilon + 1} = -2\,\frac{\epsilon - 1}{\epsilon + 1} = -\alpha_\epsilon \tag{6}$$

This study points out some interesting observations about the peculiarities of the dipole response of spherical scatterers of negative permittivity. In particular, the focus is on how closely a very high and very low observability of a small particle are connected when its permittivity approaches certain negative values. The invisibility or at least low observability of scatterers with negative material parameter values has been studied also in the full-wave regime [7], and recent studies have appeared dealing with even a complete cloaking of arbitrary objects by tracing and dragging wave rays past it [8, 9].

REFERENCES

1. Bohren, C. F. and D. R. Huffman, *Absorption and Scattering of Light by Small Particles*, Wiley, New York, 1983.
2. Fredkin, D. R. and I. D. Mayergoyz, "Resonant behavior of dielectric objects (electrostatic resonances)," *Physica Review Letters*, Vol. 91, No. 25, Article 253902, 2003.
3. Pendry, J. B., Negative refraction makes a perfect lens, *Physical Review Letters*, Vol. 85, 3966–3969, 2000.
4. Smith, D. R., J. B. Pendry, and M. C. K. Wiltshire, "Metamaterials and negative refractive index," *Science*, Vol. 305, 788–792, 2004.
5. 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.
6. Sihvola, A., *Electromagnetic Mixing Formulas and Applications*, IEE Publishing, London, 1999.
7. Alù, A. and N. Engheta, "Achieving transparency with plasmonic and metamaterial coatings," *Physical Review E*, Vol. 72, 016623, 2005.
8. Pendry, J. B., D. Schurig, and D. R. Smith, "Controlling electromagnetic fields," *Science*, 10.1126/science.1125907, May 25, 2006.
9. Leonhardt, U., "Optical conformal mapping," *Science*, 10.1126/science.1126493, May 25, 2006.

Accurate Analysis of Metamaterials Involving Finite Arrays of Split-ring Resonators and Thin Wires

Levent Gürel[1,2], **Özgür Ergül**[1], **and Alper Ünal**[1]

[1]Department of Electrical and Electronics Engineering, Bilkent University
TR-06800, Bilkent, Ankara, Turkey
[2]Computational Electromagnetics Research Center (BiLCEM), Bilkent University
TR-06800, Bilkent, Ankara, Turkey

Abstract— In order to gain physical insight into how some metamaterial structures behave, we report our results obtained by accurate numerical solutions of electromagnetic problems related to various constructions of split-ring-resonators (SRRs) and thin wires (TWs). Single and multiple layers of arrays of SRRs and TWs are investigated in detail using the electric-field integral equation. Without utilizing any homogenization techniques, we accurately model large numbers of unit cells that translate into very large computational problems, which are solved efficiently by employing multilevel fast multipole algorithm.

1. INTRODUCTION

Since they were first proposed theoretically by Veselago in 1968 [1], metamaterials (MMs) have attracted a great amount of interest because of their unusual electromagnetic properties. MMs are usually constructed by periodical arrangements of unit cells such as split ring resonators (SRRs) and thin wires (TWs). In this paper, we specifically investigate the MM structures consisting of SRR and TW arrays, as depicted in Fig. 1. These two types of arrays are also included in the same medium to obtain composite metamaterials (CMMs). In our modelling and simulations, we take into account that these structures actually have finite extent and they exhibit interface properties. Without using the symmetry and periodicity of these structures, we accurately model large numbers of unit cells to understand their transmission properties. The surfaces are modelled by perfectly conducting sheets and the scattering problems are formulated by the electric-field integral equation (EFIE). With the triangulation of the conducting surfaces, we employ Rao-Wilton-Glisson (RWG) [2] basis functions to expand the unknown surface current density. Accurate modelling of MMs translate into very large computational problems, which can be solved with the aid of advanced acceleration techniques, such as the multilevel fast multipole algorithm (MLFMA) [3].

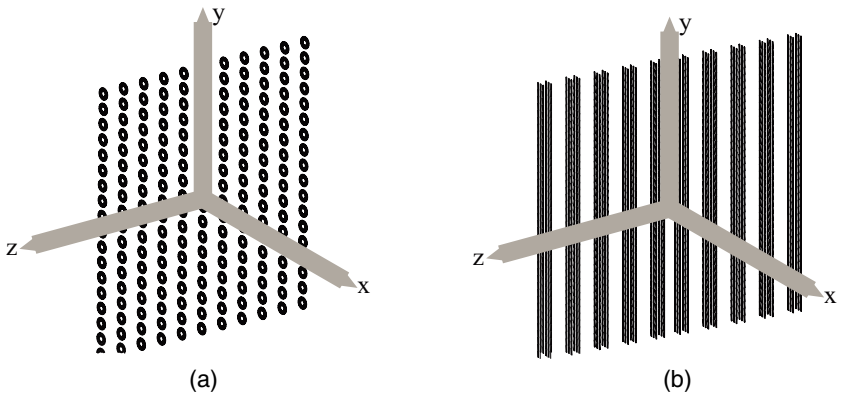


Figure 1: (a) Single-layer SRR array obtained by the arrangement of 11×18 SRRs. (b) TW array having the same dimensions as the SRR array in Fig. 1(a).

In this work, dimensions of the unit cells are in the order of microns to obtain negative effective permeability around 100 GHz [4]. Dimensions of a single SRR are as follows: The smaller ring has 43 μm inner radius and 67.2 μm outer radius, the larger ring has 80.7 μm inner radius and 107.5 μm outer radius, and the gap width is 7.2 μm. The SRR array depicted in Fig. 1(a) is constructed by the arrangement of 11×8 SRRs. Due to the negative effective permittivity stimulated in the medium, the transmission through the array is expected to decrease significantly around the

resonance frequency. Dimensions of the unit cells of the TW array depicted in Fig. 1(b) are compatible with the dimensions of the SRRs and the array exhibits negative effective permittivity in a wide range of frequencies including 100 GHz. As a result, CMM structures obtained by the combination of the SRR and TW arrays in Fig. 1 are expected to show double-negative property around 100 GHz. In this paper, examples on the SRR and CMM arrays are provided to confirm the theoretical findings on practical cases.

2. SIMULATION ENVIRONMENT

In this paper, both the SRR and TW geometries are modelled by open surfaces with zero thickness. Therefore, the scattering problems related to MMs are formulated by EFIE, which is applicable to open geometries. We solve the problems in frequency domain using phasor notation with the e^{-iwt} convention. For conducting surfaces, EFIE can be written as

$$\hat{\boldsymbol{t}} \cdot \int_S d\boldsymbol{r}' \boldsymbol{J}(\boldsymbol{r}') \cdot \left(\bar{\mathbf{I}} - \frac{\nabla \nabla'}{k^2} \right) g(\boldsymbol{r}, \boldsymbol{r}') = \frac{i}{k\eta} \hat{\boldsymbol{t}} \cdot \boldsymbol{E}^{\boldsymbol{i}}(\boldsymbol{r}) \tag{1}$$

directly from the boundary condition for the tangential electric field, where k and η are the wavenumber and wave impedance associated with the host medium. In our simulations, we assume that the relative permittivity of the host medium is 4.8, as it is commonly used in experimental setups [4]. In (1), scattered electric field is expressed in terms of the induced (unknown) surface current $\boldsymbol{J}$ with the aid of the free-space Green's function

$$g(\boldsymbol{r}, \boldsymbol{r}') = \frac{e^{ikR}}{4\pi R} \qquad \left(R = \left| \boldsymbol{r} - \boldsymbol{r}' \right| \right), \tag{2}$$

where $\boldsymbol{r}$ is an observation point on the surface, and $\hat{\boldsymbol{t}}$ is the tangential vector at the observation point. On the right-hand side of (1), $\boldsymbol{E}^{\boldsymbol{i}}(\boldsymbol{r})$ denotes the incident electric field.

For the numerical solutions of EFIE, the unknown current induced on the conducting surfaces is expanded in a series of basis functions $\boldsymbol{b}_n$, i.e.,

$$\boldsymbol{J}(\boldsymbol{r}) = \sum_{n=1}^{N} a_n \boldsymbol{b}_n(\boldsymbol{r}), \tag{3}$$

where a_n represents the unknown coecients for $n = 1, 2, \ldots, N$ and N is the number of unknowns. Then, the application of the method of moments leads to $N \times N$ dense matrix equation

$$\sum_{n=1}^{N} Z_{mn}^{E} a_n = v_m^{E}, \qquad m = 1, \ldots, N, \tag{4}$$

where

$$Z_{mn}^{E} = \int_{S_m} d\boldsymbol{r} \boldsymbol{t}_m(\boldsymbol{r}) \cdot \int_{S_n} d\boldsymbol{r}' \left(\bar{\mathbf{I}} - \frac{\nabla \nabla'}{k^2} \right) g(\boldsymbol{r}, \boldsymbol{r}') \cdot \boldsymbol{b}_n(\boldsymbol{r}') \tag{5}$$

represents the matrix element, and

$$v_m^{E} = \frac{i}{k\eta} \int_{S_m} d\boldsymbol{r} \boldsymbol{t}_m(\boldsymbol{r}) \cdot \boldsymbol{E}^{\boldsymbol{i}}(\boldsymbol{r}) \tag{6}$$

represents the mth element of the excitation vector. We apply the Galerkin scheme by choosing the testing functions t_m also as RWG. In (5) and (6), S_m and S_n symbolize the spatial supports of the mth testing and nth basis functions, respectively.

The matrix equation in (4) is solved iteratively, where the matrix-vector products are accelerated by MLFMA. The fundamental idea in MLFMA is to replace the element-to-element interactions with cluster-to-cluster interactions in a multilevel scheme. This computational scheme relies on the factorization of the Green's function, which is valid only for basis and testing functions that are far from each other. By employing MLFMA, matrix-vector multiplications can be performed in $O(N \log N)$ processing time using $O(N \log N)$ memory as detailed in [5].

EFIE usually produces ill-conditioned matrix equations that are dicult to solve by an iterative algorithm [6]. In addition, MM structures usually present numerical resonances, which further

inhibits a quick convergence without preconditioning. To obtain a convergence in a reasonable number of iterations, we employ near-field preconditioner (NFP) obtained by retaining all of the available near-field interactions. NFP is not commonly used in the solutions of large scattering problems, since it has a complexity larger than MLFMA. However, for all problems investigated in this work, we observe that the additional time required for the factorization and use of NFP is smaller than the time gained by the reduction of the iterations. In addition to NFP, we also employ more robust and efficient preconditioners based on approximate MLFMA inserted in flexible solvers consisting of inner and outer iterations.

3. RESULTS

Figure 2 presents the results for the SRR array in Fig. 1(a), where the power transmission is plotted at various frequencies, i.e., 95 GHz, 100 GHz, and 110 GHz. The transmission is calculated at different points in the $z = 0$ plane and the SRR arrays are depicted in the plots. The excitation is a Hertzian dipole oriented in the y direction as also indicated in the plots and the transmission region is on the left of the arrays. At 95 GHz and 105 GHz, the power transmission through the 1-layer SRR array is almost unity, which corresponds to 0 decibels (dB). On the other hand, around the resonance frequency (100 GHz), the transmitted power drops dramatically due to the shadowing effect of the SRR array. In other words, negative effective permeability is stimulated in the medium at 100 GHz. In the 2-layer case, the frequency range for the resonance effect is extended and the transmitted power is slightly blocked also at 95 GHz and 105 GHz.

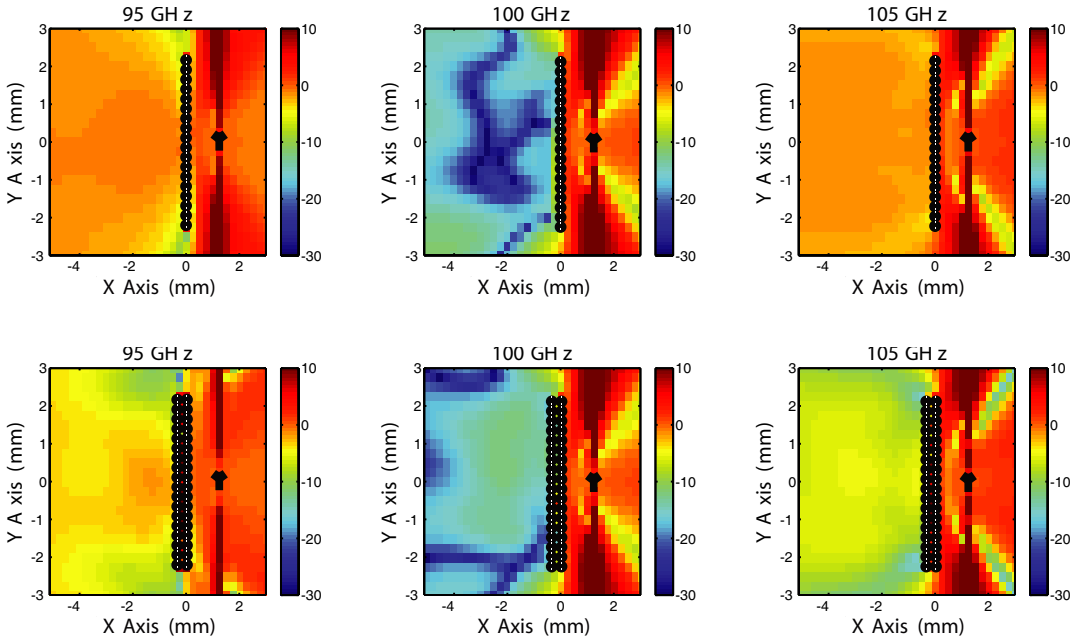


Figure 2: Power transmission for the 1-layer and 2-layer 11 × 18 SRR arrays at 95 GHz, 100 GHz, and 105 GHz. SRRs and the ideal (Hertzian) dipole are also depicted in the figures and the transmission region is on the left of the array.

Figure 3 presents the power transmission for the 1-layer and 2-layer CMM arrays constructed by employing the SRR and TW arrays in Fig. 1. At 95 GHz and 105 GHz, we observe that the array blocks the fields and this is mainly due to the negative effective permittivity introduced by the TWs. On the other hand, the transmission through the CMM array increases at 100 GHz, which is more visible in the 2-layer case. The reason is that the SRRs resonate around 100 GHz as depicted in Fig. 2. Then, both the effective permittivity and permeability stimulated in the medium become negative. We note that increasing the number of layers provides an extended range for the negative effective permeability introduced by the SRRs and this leads to increased double-negativity effect, i.e., more power transmission around the resonance frequency, obtained from the CMM structure.

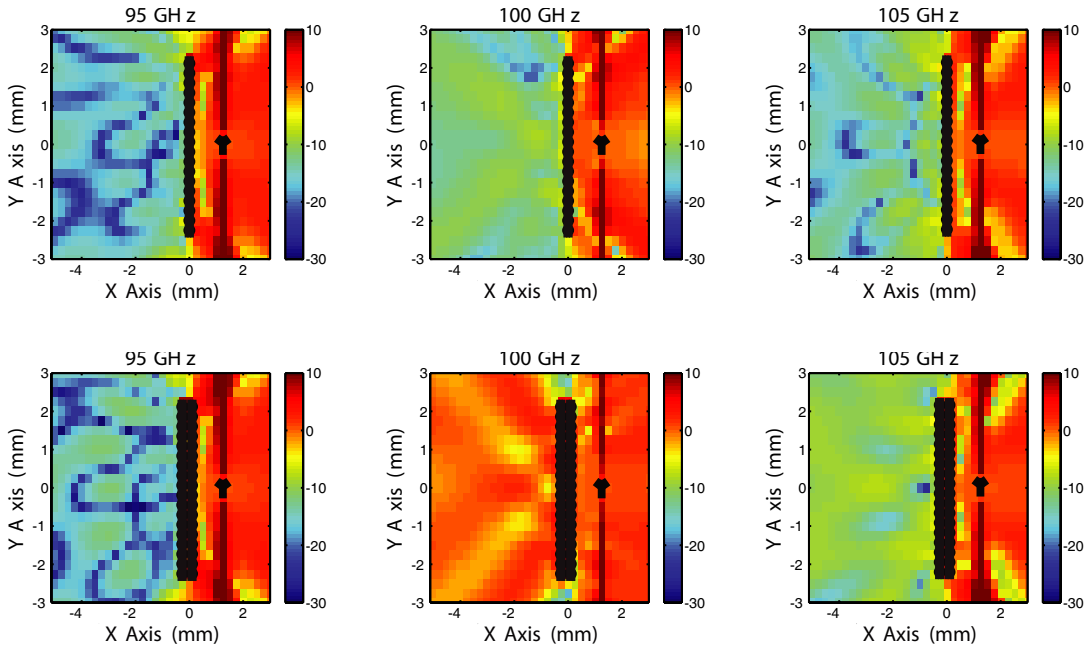


Figure 3: Power transmission for the 1-layer and 2-layer CMM arrays at 95 GHz, 100 GHz, and 105 GHz. The arrays and the ideal (Hertzian) dipole are also depicted in the figures and the transmission region is on the left of the arrays.

ACKNOWLEDGMENT

This work was supported by the Scientific and Technical Research Council of Turkey (TUBITAK) under Research Grant 105E172, by the Turkish Academy of Sciences in the framework of the Young Scientist Award Program (LG/TUBA-GEBIP/2002-1-12), and by contracts from ASELSAN and SSM.

REFERENCES

1. Veselago, V. G., "The electrodynamics of substances with simultaneously negative values of permittivity and permeability," *Sov. Phys. Usp.*, Vol. 10, No. 4, 509–514, 1968.
2. Rao, S. M., D. R. Wilton, and A. W. Glisson, "Electromagnetic scattering by surfaces of arbitrary shape," *IEEE Trans. Antennas Propagat.*, Vol. AP-30, No. 3, 409–418, 1982.
3. Lu, C.-C. and W. C. Chew, "Multilevel fast multipole algorithm for electromagnetic scattering by large complex objects," *IEEE Trans. Antennas Propagat.*, Vol. 45, No. 10, 1488–1493, 1997.
4. Gokkavas, M., K. Güven, I. Bulu, K. Aydın, R. S. Penciu, M. Kafesaki, C. M. Soukoulis, and E. Özbay, "Experimental demonstration of a left-handed metamaterial operating at 100 GHz," *Phys. Rev. B.*, Vol. 73, No. 193103, 2006.
5. Chew, W. C., J.-M. Jin, E. Michielssen, and J. Song, *Fast and Efficient Algorithms in Computational Electromagnetics*, Artech House, Boston, MA, 2001.
6. Gürel, L. and Ö. Ergül, "Comparison of FMM implementations employing different formulations and iterative solvers," *2003 IEEE AP-S International Symposium*, Columbus, Ohio, Vol. 1, 19–22, June 2003.

A Magnetic Metamaterial Composed of Randomly Oriented SRRs

L. Jelínek, J. Macháč, and J. Zehentner

Czech Technical University in Prague, Technická 2, 166217 Prague 6, Czech Republic

Abstract— The concept, manufacture and testing of a volumetric isotropic magnetic metamaterial composed of randomly distributed broadside coupled split ring resonators is presented in this paper. An estimate of the isotropy based on the measured transmissions is made in terms of the arithmetic mean value and the standard deviation of the sample. We offer conclusions and recommendations for the development of a genuine bulk isotropic medium of various densities and distributions of the particles, based on the 64 measured frequency responses of each sample.

1. INTRODUCTION

In recent years metamaterials have attracted great attention among scientists, researchers and engineers [1, 2]. The main characteristics of metamaterials, such as left-handed behavior of the waves, negative permeability, negative permittivity and negative index of refraction, can be observed in bulky media and also on periodically loaded transmission lines from microwaves up to optics. The concept of a left-handed transmission line is well known [3]. The main research effort is devoted to the design of practically realizable unit cells in available technology meeting the desired relatively wide-band frequency characteristics. Solving the same task in a bulky medium is much more difficult. This is the challenge that has inspired our work.

Generally, metamaterials are anisotropic media. However, for many applications isotropic materials are needed. Two ways can be used to achieve isotropy. The first is to use a specific form of unit cells, periodically arranged in such a way that the system satisfies selected crystallographic groups of symmetry. The 3D structures introduced in [4] provide an example of this approach. The second way leading to an isotropic medium is based on randomly located cells in the volume of the host. We will follow this approach and will try to verify its usefulness for producing bulky isotropic magnetic metamaterials. The first results are presented here.

2. ARCHITECTURE AND FABRICATION OF A 2D ISOTROPIC METAMATERIAL

3D random location of particles is a great technological problem, and therefore the 2D random system will be studied first. The system will consist of the host material, which has the shape of a parallelepiped cut into three slices that are equal in height, Fig. 1. In each slice an equal number of particles, Broadside Coupled Split Ring Resonators (BC-SRR) [5], is inserted with

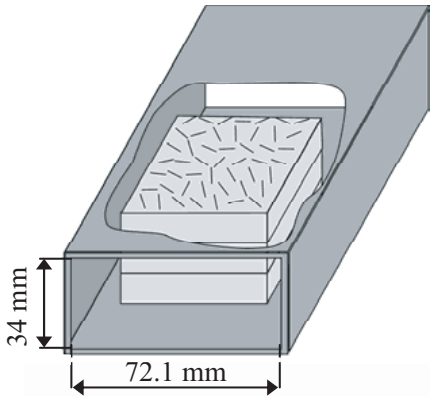


Figure 1: Waveguide loaded with parallelepiped composing of three slices with randomly distributed BC-SRRs.

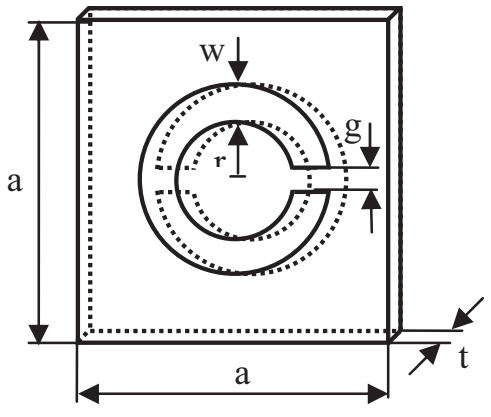


Figure 2: Broadside-coupled split ring resonator.

random orientation. The BC-SRR consists of two split rings placed on the opposite sides of a thin dielectric substrate so that the particle exhibits inversion symmetry, as shown in Fig. 2. Note also that BC-SRR is not a bianisotropic particle [5] and the orientation of its ring-gaps does not play a

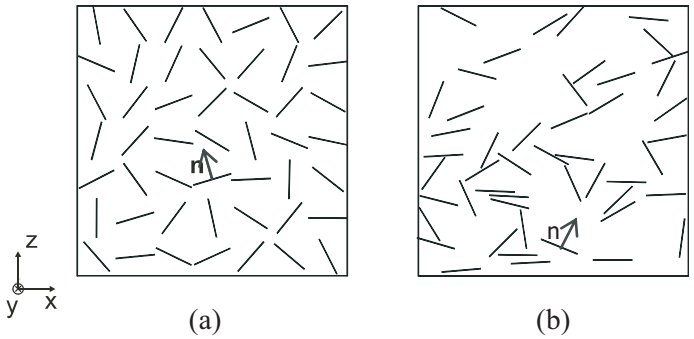


Figure 3: (a) Periodic BC-SRR positions in the nodes of the squared net, and (b) their random positions in the slice.

role when it is excited by a magnetic field. Two different random arrangements, depicted in Fig. 3, representing the top view of a slice, were used and tested. The surface normals of the SRRs, and thus also the magnetic moments, lie only in the x-z plane, forming the desired 2D system. The centers of particular resonators are either positioned periodically in the nodes of a squared net, Fig. 3(a), or their position varies randomly, Fig. 3(b). Randomness is obtained by choosing random angles between the normals of the resonators and the x-axis. The distribution of these angles and also of the particle positions on each slice was generated independently. Consequently, by changing the mutual orientations of the slices in the parallelepiped by 90-degree rotations, there are 64 different parallelepipeds with randomly distributed particles inside, assuming that the vertical order of the slices is kept. The interchange of the vertical order of the slices, with their three rotations, offers a total of 384 combinations of particle locations. However, we utilized only the first 64 combinations in the experiment.

Standard printed circuit technology was used for etching the BC-SRRs on Rogers RT/duroid 5880 substrate of thickness $t = 0.127\,\text{mm}$ with $\varepsilon_r = 2.20$ and $0.017\,\text{mm}$ copper cladding. Each resonator resides on a squared leaf with an edge $a = 7\,\text{mm}$. Froth polystyrene was chosen as the host dielectric. The parallelepiped base edge, equal to the width of the squared slice, is 50 mm, while the height of each slice is 10 mm. The proportions of the BC-SRR are $r = 2.5\,\text{mm}$, $w = 0.7\,\text{mm}$ and $g = 0.3\,\text{mm}$.

3. EXPERIMENTS, OBSERVATIONS AND RESULTS

The proportions of the BC-SRRs were designed in such a way that their resonant frequencies occurred around 3 GHz. Consequently, measurements were carried out in the rectangular waveguide R32. The HP 8510B network analyzer measured the scattering coefficients. First we checked the resonant frequency of the single resonator positioned in the middle E plane of the waveguide. A typical transmission is depicted in Fig. 4. It turned out that the resonant frequencies of the individual resonators differed in the 50 MHz interval due to manufacturing imperfections.

In the next step, the isotropy of the above-mentioned parallelepipeds, stacks of three slices, was tested. The stacks were inserted into the rectangular waveguide R32, as depicted in Fig. 1. If the parallelepiped is isotropic, then the measured scattering parameters should stay invariant for any 90° rotations of each slice of the parallelepiped around its central vertical axis. Due to the random arrangement of the particles, this last presumption should hold in a statistical sense and therefore the arithmetic mean value $\bar{x}$

$$\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i$$

and the sample standard deviation σ

$$\sigma = \sqrt{\left(\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2\right)}$$

were calculated from all the 64 measurements. This resulted in the records depicted in Figs. 5–7.

All three measured samples can be described by means of three characteristics. The first is the isotropy, which in the measurements presented here is directly proportional to the disperse σ of

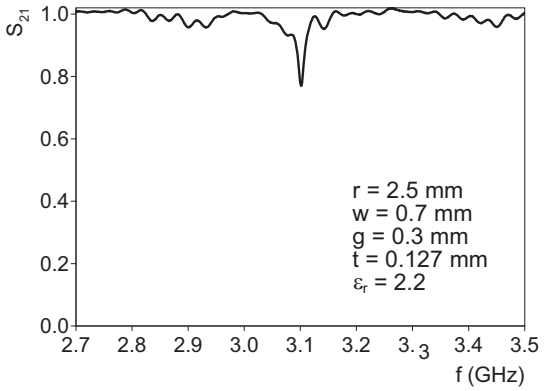


Figure 4: Transmission of one BC-SRR in the R32 waveguide, $f_0 = 3.1015\,\mathrm{GHz}$.

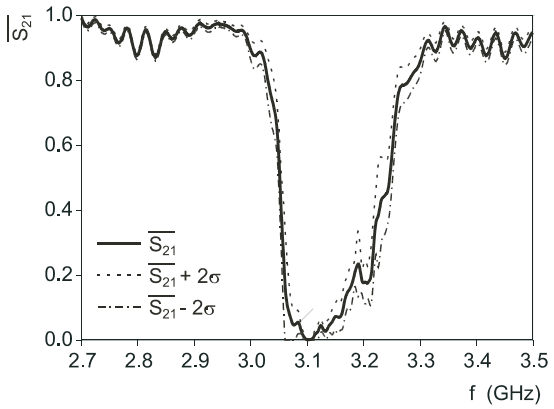


Figure 5: Arithmetic mean value $\bar{s}_{21}$ of the transmission through the sample with 147 BC-SRRs and its disperse $\bar{s}_{21} \pm 2\sigma$ when the resonators are in the nodes of the squared net.

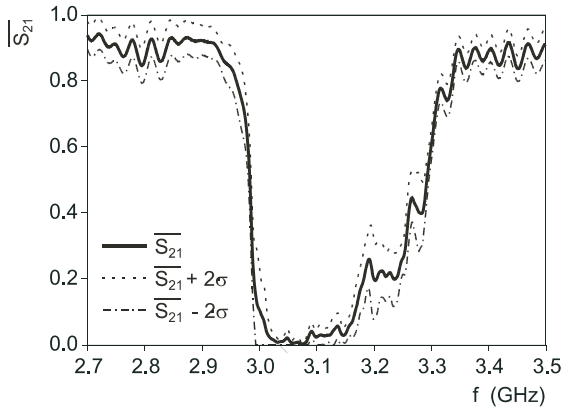


Figure 6: Arithmetic mean value $\bar{s}_{21}$ of the transmission through the sample with 243 BC-SRRs and its disperse $\bar{s}_{21} \pm 2\sigma$ when the resonators are in the nodes of the squared net.

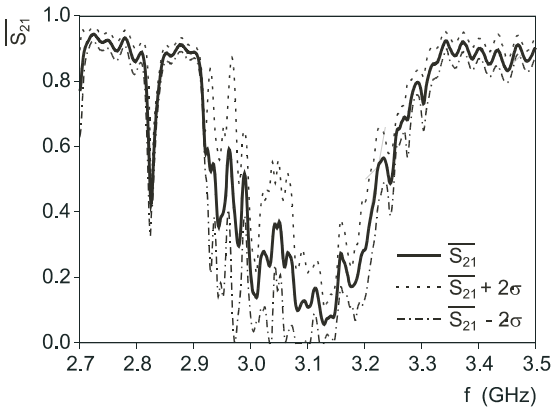


Figure 7: Arithmetic mean value $\bar{s}_{21}$ of the transmission through the sample with 243 BC-SRRs and its disperse $\bar{s}_{21} \pm 2\sigma$ when the resonators are randomly distributed.

the transmission coefficient. Furthermore, it was experimentally observed that almost all measured data lies in the interval of $\pm 2\sigma$ around the arithmetic mean value of s_{21}. It is seen from Figs. 5 and 6 that the isotropy depends only slightly on the particle density, as the disperse of the transmission coefficient is nearly the same in both cases. However, the isotropy is very sensitive to the particle distribution, as follows from Fig. 7. This figure shows that the particle distribution from Fig. 3(b) is not very suited for isotropic materials, while the distribution from Fig. 3(a) provides a material with acceptable isotropy. The second and third important characteristics are the bandwidth and the insertion losses at the resonance. It is worth noting that the higher the insertion losses, the higher, in absolute values, are the effective material parameters. It is apparent from Figs. 5 and 6 that widening of the bandwidth can be achieved by using higher density of the particles, and the insertion losses are preserved. Using random positions of the particles, as in Fig. 3(b), leads also to the bandwidth widening shown in Fig. 7. However, in this case the insertion losses are reduced.

Lastly the measured transmissions of the randomly distributed SRRs were compared with transmissions of the sample with regular particle distribution. A parallelepiped was produced with all resonators aligned in such a way that their magnetic moments were oriented along the exciting transversal magnetic field. This parallelepiped consists of three slices, as in the former cases. Regular distribution of the resonators, as depicted in Fig. 8, was used in each slice. Now the resonators in one row occupied positions shifted by the half-pitch with regard to resonators standing in the closest neighboring rows in order to achieve the same density of the particles, i.e., 147 pieces per parallelepiped. The measured transmission coefficient through the waveguide loaded with this highly anisotropic sample is depicted in Fig. 9. Comparing Figs. 5 and 9, we can conclude that the

random distribution of particles, which leads to an isotropic response, not only leaves the insertion losses unaffected, but also remarkably widens the bandwidth.

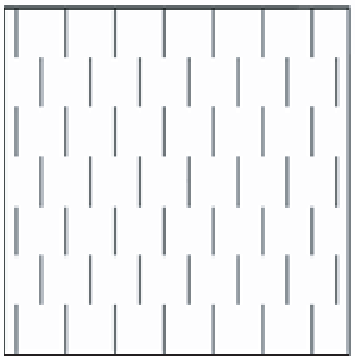


Figure 8: Regular positions of particles in one slice.

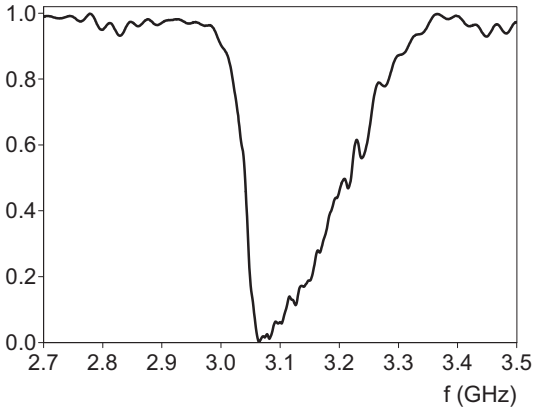


Figure 9: Transmission through 147 BC-SRRs periodically located and axially aligned.

4. CONCLUSIONS

To obtain a genuine 3D isotropic medium the space density of the particles has to be sufficiently great to enable averaging of the responses of all involved effects and couplings participating in the homogenization, and the size of the particles must be kept negligible compared with the wavelength. It turned out in our experiments that the particle distribution is also essential. While random distribution of randomly oriented particles does not lead to an isotropic material, regular distribution of randomly oriented particles shows very good results and with higher particle density it also provides an opportunity to increase the bandwidth of the resonance. The latter particle distribution is thus very promising and should also be used for manufacturing the 3D isotropic material.

ACKNOWLEDGMENT

This work has been supported by the Grant Agency of the Czech Republic under project 102/06/1106 "Metamaterials, nanostructures and their applications".

REFERENCES

1. Engheta, N. and R. W. Ziolkowski Ed., *Metamaterials: Physics and Engineering Explorations*, Willey-IEEE Press, 2006.
2. Eleftheriades, G. V. and K. G. Balmain Ed., *Negative-Refraction Materials: Fundamental Principles and Applications*, Willey-IEEE Press, 2006.
3. Caloz, C. and T. Itoh, *Electromagnetic Metamaterials, Transmission Line Theory and Microwave Applications*, Willey - Interscience, 2005.
4. Baena, J. D., L. Jelinek, R. Marqués, and J. Zehentner, "On the design of isotropic magnetic materials," *Proceedings of MELECON 2006*, 250–254, Malaga, Spain, May 2006.
5. Marqués, R., F. Medina, and R. Rafii-El-Idrisi, "Role of bianisotropy in negative permeability and left-handed metamaterials," *Physical Review B*, Vol. 65, 144440-(1–6), 2002.

Compact Rat-race Hybrid Based on Complementary Split Rings Resonators

G. Sisó, J. Bonache, M. Gil, I. Gil, J. García-García, and F. Martin

GEMMA/CIMITEC, Department Enginyeria Electrònica
Universitat Autònoma de Barcelona 08193 BELLATERRA, Spain

Abstract— In this work, the application of Complementary Split Rings Resonators (CSRRs) to the design of a rat-race hybrid in microstrip technology is analyzed. The use of a left-handed (LH) cell composed by a CSRR etched in the ground plane and series capacitive gap reduces the length of the required 270° line by a factor of 5. Thanks to this length reduction, the area of the final prototype is 4 times smaller than the conventional distributed design. The measured results show satisfactory behavior in the frequency of interest.

1. INTRODUCTION

Complementary split rings resonators (CSRRs) have been recently reported by some of the authors [1] as key particles for the synthesis of negative permittivity metamaterial microstrip lines. Combined with series capacitive gaps, CSRRs have been also used for the design of narrow bandpass structures with backward (or left-handed — LH) wave propagation in the allowed band [2], or other microwave devices such as power dividers [3], impedance inverters, or backward couplers [4]. In this work, a left-handed cell based on CSRRs is applied to the design of a microwave hybrid coupler in "rat-race" configuration, implemented by using a fully planar fabrication process, i.e., without any external lumped component. Taking advantage of the propagation characteristics of this type of structures, important size reduction will result. This same principle has been previously applied by using the dual (or composite right/left handed) approach (non resonant) for the design of this type of devices [5]. In that case, however, external lumped elements where necessary to achieve the desired behavior.

2. DEVICE DESIGN AND RESULTS

The device consists on square-shaped combination of three (right-handed) conventional lines with $-90°$ phase shift, and one CSRR-based metamaterial cell that exhibits a positive phase shift of $+90°$ (the characteristic impedance of all the lines is 70.71 Ω). The metamaterial cell presents the same response (in the frequency of operation) than a $-270°$, but with an important size reduction as compared to the distributed approach. This cell is composed by a CSRR etched in the ground plane of a microstrip line combined with a series capacitive gap. The layout of this cell can be seen in Fig. 1(a), where the bottom layer of the cell is depicted in grey color and the top layer is depicted in black. The electrical model of this cell appears in Fig. 1(b) [6]. The series gap is modeled by the capacitance C_g, while the CSRR is described by the parallel resonant tank (with inductance L_c and capacitance C_c), which is electrically coupled to the host line through the capacitance C; the line inductance is modeled by L. In order to design the metamaterial cell, we must force $+90°$ phase shift (positive instead of, negative due to the LH nature of the cell) and 70.71 Ω characteristic impedance, at the operating frequency of the device. This can be achieved by using the following expressions:

$$Z_B = \sqrt{Z_S(j\omega)[Z_S(j\omega) + 2Z_P(j\omega)]} \tag{1}$$

$$\cos(\beta l) = 1 + \frac{Z_S(j\omega)}{Z_P(j\omega)} \tag{2}$$

where Z_B, βl, Z_s and Z_p are the characteristic impedance, the electrical length, the series and shunt impedance of the T circuit model of the LH the cell. These two conditions are not enough to completely determine all the elements of the electrical model of Fig. 1(b) (in fact we need three additional conditions, but taking into account that Z_s is dominated by C_g the contribution of L has not been considered). We can also force the limits of the frequency band where the LH wave propagation is possible by using the procedure shown in [7]. In this device we have set the operation frequency to 4.8 GHz, with a LH band comprised between 4.6 GHz and 5.1 GHz. With

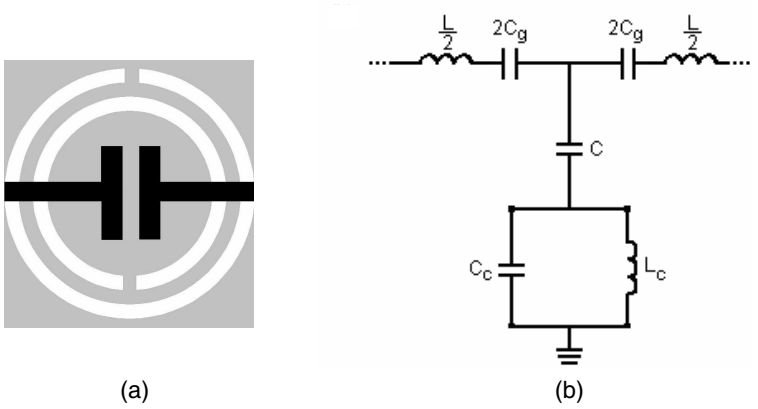


Figure 1: Topology of the LH cell (a) and its lumped element equivalent circuit model. (b) The upper metallization is depicted in black, whereas the bottom slot regions are depicted in grey.

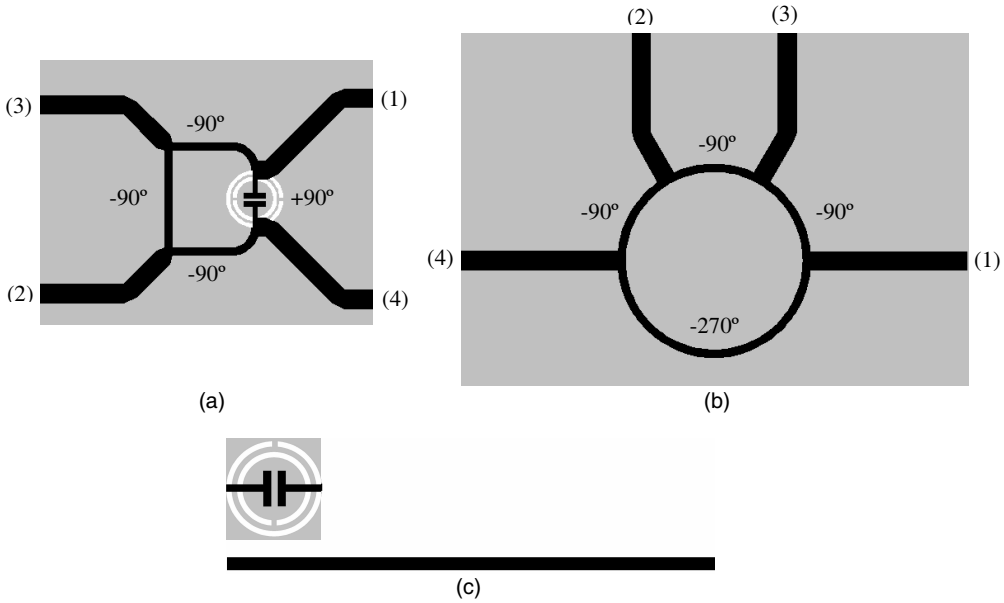


Figure 2: CSRR-based hybrid coupler (a) conventional hybrid coupler (b) and comparison between a +90° LH cell and conventional −270° transmission line.

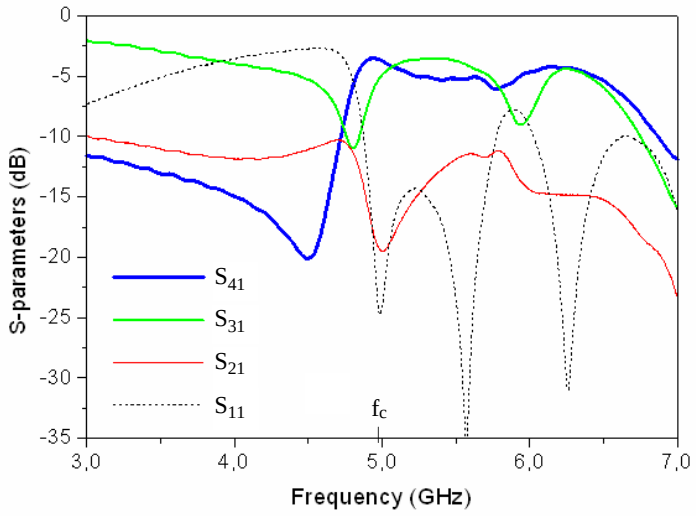


Figure 3: Measured frequency response of the CSRR-based hybrid coupler.

this information, we have obtained the element values of the electrical model of the cell. From these values we have obtained final layout dimensions of the cell by using the expressions that appears in [6] as a first approximation, and by using the extraction parameter method presented in [8]. The final layout of the device is depicted in Fig. 2(a), together with the layout of a conventional rat-race hybrid, whereas Fig. 2(b) shows a size comparison between the $+90^\circ$ LH cell and the conventional TL with an electrical length of -270°. It can be seen a significant reduction of the physical length of the metamaterial cell as compared with the conventional line. This induces an important size reduction in the final prototype, as compared to the distributed approach. The frequency response of the final prototype appears in Fig. 3(a), where it can be seen the expected coupling in the frequency of interest. The resulting bandwidth is smaller than those achievable with conventional rat-race hybrids, due to the characteristics of the cell used. This narrow band behavior is expected to be improved by using other types of cells based on CSRRs, which are able to produce extremely wide bandwidths [9]. Nevertheless the proposed device is small, fully planar and can be of interest for narrowband applications.

3. CONCLUSION

In this work, the application of a left-handed CSRR-based cell to the design of a microwave hybrid has been analyzed. The device is composed of three sections of conventional transmission lines, and one section formed by a left-handed cell. The propagation characteristics of the left-handed cell allows for a significant size reduction as compared to the conventional distributed design. As a difference between other solutions based on metamaterials, this hybrid has been implemented without any external lumped component. The measured results exhibit satisfactory behavior for narrowband applications.

REFERENCES

1. Falcone, F., T. Lopetegi, J. D. Baena, R. Marqués, F. Martín, and M. Sorolla, "Effective negative-ε stop-band microstrip lines basedon complementary split ring resonators," *IEEE Microw. Wirel. Compon. Lett.*, Vol. 14, 280–282, 2004.
2. Falcone, F., T. Lopetegi, M. A. G. Laso, J. D. Baena, J. Bonache, M. Beruete, R. Marqués, F. Martí, and M. Sorolla, "Babinet principle applied to the design of metasurfaces and metamaterials," *Phys. Rev. Lett.*, Vol. 93, 197401, 2004.
3. Gil, M., J. Bonache, I. Gil, J. García-García, and F. Martín, "Miniaturization of planar microwave circuits by using resonant-type left handed transmission lines", *IEE. Proc. on Antennas and Propagation*, Submited.
4. Jarauta, E., M. A. G. Laso, T. Lopetegi, F. Falcone, M. Beruete, J. D. Baena, J. Bonache, I. Gil, J. Garcia-Garcia, A. Marcotegui, F. Martin, R. Marques, and M. Sorolla, "Metamaterial microstrip backward couplers for fully planar fabrication techniques," *Infrared and Millimeter Waves*, 2004; and *12th International Conference on Terahertz Electronics, 2004*, *Conference Digest*, 185–186, 27 Sept.–1 Oct. 2004.
5. Okabe, H., C. Caloz, and T. Itoh, "A compact enhanced-bandwidth hybrid ring using an artificial lumped-element left-handed transmission-line section," *IEEE Microw. Wirel. Compon. Lett.*, Vol. 52, 898–804, 2004.
6. Baena, J. D., J. Bonache, F. Martín, R. Marqués, F. Falcone, T. Lopetegui, M. A. G. Laso, J. García-García, I. Gil, P. M. Flores, and M. Sorolla, "Equivalent-circuit models for split-ring resonators and complementary split-ring resonators coupled to planar transmission lines," *IEEE Trans. on Microwave Theory and Techniques*, Vol. 53, No. 4, 1451–1461, April 2005.
7. Gil, M., J. Bonache, I. Gil, J. García-García, and F. Martín, "On the transmission properties of left handed microstrip lines implemented by complementary split rings resonators," *Int. Journal Numerical Model.*, Vol. 19, 87–103, 2006.
8. Bonache, J., J. Martel, I. Gil, M. Gil, J. García-García, F. Martín, I. Cairó, and M. Ikeda "Super compact ($< 1\,\mathrm{cm2}$) band pass filters with wide bandwidth and high selectivity at C-band", *European Microwave Conf.*, Manchester, UK, 2006, Accepted.

High-pass Filters Implemented by Composite Right/Left Handed (CRLH) Transmission Lines Based on Complementary Split Rings Resonators (CSRRs)

M. Gil, J. Bonache, J. Selga, J. García-García, and F. Martín

GEMMA/CIMITEC (Departament d'Enginyeria Electrònica)
Universitat Autònoma de Barcelona, 08193 BELLATERRA (Barcelona), Spain.

Abstract— It was recently shown [1] that it is possible to achieve broadband composite right/left handed (CRLH) transmission lines by using complementary split rings resonators (CSRRs). Making use of balanced CRLH/CSRR-based metamaterial transmission lines, high pass filters are designed. Good performance and small dimensions are achieved. The results open the door to new design possibilities in microwave circuitry by using this type of CRLH metamaterial transmission lines.

In microwave engineering, two main approaches exist to obtain left-handed transmission lines. One of them is the dual transmission line approach [2–4], which is based on loading transmission lines with series capacitances and shunt inductances. The other one is the resonant-type approach [5, 6], which combines resonant sub-wavelength particles, such as split rings resonators (SRRs) or their dual counterparts (CSRRs), with shunt inductances or series capacitances etched on a host transmission line. Traditionally, it was supposed that the use of structures based on the resonant-type model was limited to narrow band applications. However, as occurs in the dual transmission line model, resonant-based left handed structures exhibit also a right-handed transmission band at higher frequencies due line parasitics [7]. This allows us to obtain a composite right/left handed (CRLH) behaviour. The line can be tailored in order to collapse the typical frequency gap present between the left handed and right handed transmission bands. Namely, it is possible to make coincident the upper and the lower limits of the left and the right-handed bands, respectively (balance case), with the result of a continuous transition between both transmission bands. Although up to now it was not exploited, the balance case is also achievable by means of the resonant approach, what makes possible to obtain broad-band responses by using CRLH resonant-type transmission lines.

The basic cell of the structure is represented in Fig. 1(a). It consists on a CSRR etched on the ground plane of a microstrip line, and located underneath a capacitive gap etched on the top layer (conductor strip). The CSRR provides the negative value of the dielectric permittivity in the vicinity of its resonance frequency, whereas the negative value of the magnetic permeability needed to obtain the left-handed behaviour is related to the effect of the series capacitance due to the gap. This structure is modelled by the equivalent T-circuit model shown in Fig. 1(b) [8]. The CSRR is modelled by the resonant tank L_C-C_C, which is electrically coupled to the line through the capacitance C. The capacitance C_g models the series gap, and L accounts for the line inductance. The right handed transmission band is due to line parasitics (L and C), but also to the CSRR (which behaves capacitively at high frequencies).

The analysis of the equivalent circuit model reveals the behaviour of the structure. The phase shift per cell, ϕ, (dispersion relation) and Bloch impedance, Z_B, are given by:

$$\cos\phi = 1 + \frac{Z_s(j\omega)}{Z_p(j\omega)} \tag{1}$$

$$Z_B = \sqrt{Z_s(j\omega)\left[Z_s(j\omega) + 2Z_p(j\omega)\right]} \tag{2}$$

where Z_s and Z_p are the series and shunt impedances of the equivalent circuit model. Propagation is allowed at those frequencies in which ϕ and Z_B take real values. That occurs in two frequency bands. The left-handed band begins at f_L, where $\phi(f_L) = -180^\circ$ and $Z_B(f_L) = 0$, and finishes at the resonance frequency of the CSRR, f_p, where $\phi(f_p) = 0^\circ$ and $Z_B(f_p) \to \infty$ The series elements L and C_g do also present a resonance frequency, f_s, beyond which appears the right-handed transmission band. The balance case occurs when both, the series impedance and the parallel admittance are

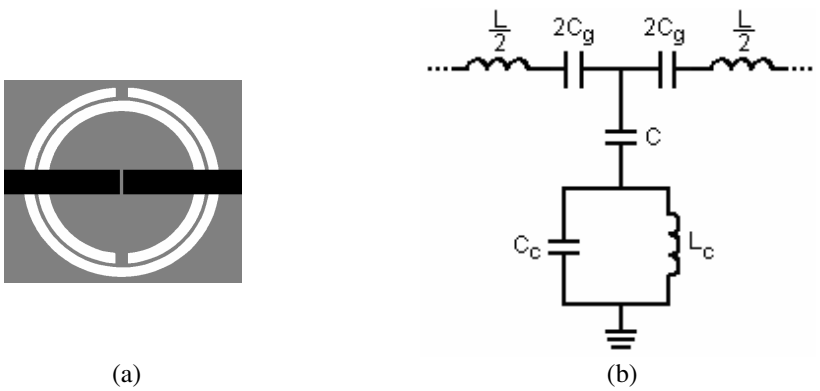


Figure 1: Topology (layout) of the basic cell (a) and its equivalent T-circuit model, (b) The layout represents in black the microstrip line on the top layer whereas the metallic parts on the bottom layer are depicted in grey.

null at the same frequency f_0 so that the two resonance frequencies are identical ($f_0 = f_p = f_s$). In this case, the frequency gap between the two different transmission bands disappears and, as a result, there is a broad composite right-left-handed transmission band.

Taking advantage of this possibility, several balanced cells have been designed and fabricated. Broad band responses have been obtained [1]. It is also possible to fabricate multi-cell devices that likewise reproduce this composite right-left-handed behaviour. Taking this into account, the possibility of designing a resonant-type high pass filter formed by several cells can be exploited. As an illustrative example, Fig. 2 shows the layout of a three-cell structure.

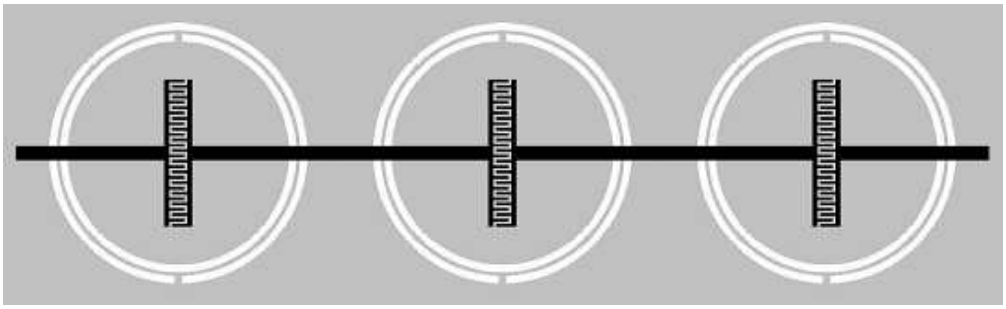


Figure 2: Layout of the filter formed by three cells. The metallic parts are depicted in black in the top layer, and in grey in the bottom layer. The rings are etched on the bottom layer. Dimensions are: total length $l = 55\,\text{mm}$, line width $W = 0.8\,\text{mm}$, external radius of the outer rings $r = 7.3\,\text{mm}$, rings width $c = 0.4\,\text{mm}$ and rings separation $d = 0.2\,\text{mm}$; the interdigital capacitors, formed by 28 fingers separated 0.16 mm, have been used to achieve the required capacitance value.

In Fig. 3, the simulated (through the commercial software *Agilent Momentum*) frequency responses obtained in three different filters formed by two, three, and four cells are shown. The parameters of the *Rogers RO3010* substrate have been considered (thickness $h = 1.27\,\text{mm}$, dielectric constant $\varepsilon_r = 10.2$). As expected, the larger the number of cells is, the more abrupt the transition band becomes. It is also clearly visible how stop band rejection increases with the number of cells.

The three structures exhibit low in-band losses and reasonable return losses. It is worth mentioning the large bandwidth of the high pass filters achieved. The cutoff frequency is located at 0.75 GHz, whereas transmission is truncated at high frequencies, as occurs in any planar filter implementation. These results demonstrate the possibility of obtaining broadband responses by means of the application of these resonant-type CRLH structures. By controlling the cut-off at higher frequencies, the application to band pass filters is also possible.

As conclusion, in this work we have presented results that confirm the possibility of obtaining broadband responses by using resonant-type CRLH structures. We have described and applied the balance case to the design of high pass filters. Dimensions and performance are good. Work is in progress in order to apply this new method to obtain new compact broadband devices.

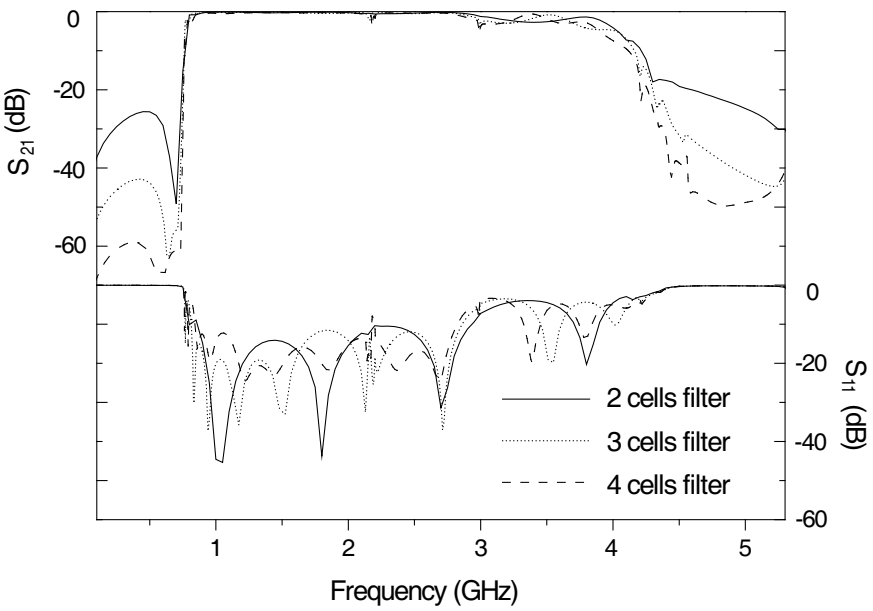


Figure 3: Simulated frequency response of three filters formed by two, three and four CSRR CRLH cells.

ACKNOWLEDGMENT

This work has been supported by Spain-MEC (project contract TEC2004-04249-C02-01). MEC has also given an FPU grant to Marta Gil (Ref. AP2005-4523).Thanks are also given to the European Union for funding the Network of Excellence NoE METAMORPHOSE, and to the Catalan Government (CIDEM) for funding CIMITEC.

REFERENCES

1. Gil, M., J. Bonache, J. Selga, J. García, and F. Martín, "Broadband resonant-type metamaterial transmission lines," *IEEE Microwave and Wireless Components Letters*, (submitted).
2. Iyer, A. K. and G. V. Eleftheriades, "Negative refractive index metamaterials supporting 2-D waves," *IEEE-MTT Int'l Symp.*, Vol. 2, 412–415, Seattle, WA, June 2002.
3. Oliner, A. A., "A periodic-structure negative-refractive-index medium without resonant elements," *URSI Digest*, *IEEE-AP-S USNC/URSI National Radio Science Meeting*, 41, San Antonio, TX, June 2002.
4. Caloz, C. and T. Itoh, "Application of the transmission line theory of left-handed (LH) materials to the realization of a microstrip LH transmission line," *Proc. IEEE-AP-S USNC/URSI National Radio Science Meeting*, Vol. 2, 412–415, San Antonio, TX, June 2002.
5. Martín, F., F. Falcone, J. Bonache, J. R. Marqués, and M. Sorolla, "A new split ring resonator based left handed coplanar waveguide," *Appl. Phys. Lett.*, Vol. 83, 4652–4654, December 2003.
6. Falcone, F., T. Lopetegi, M. A. G. Laso, J. D. Baena, J. Bonache, R. Marqués, F. Martín, and M. Sorolla, "Babinet principle applied to the design of metasurfaces and metamaterials," *Phys. Rev. Lett.*, Vol. 93, 197401–4, November 2004.
7. Sanada, A., C. Caloz, and T. Itoh, "Characteristics of the composite rigth/left handed transmisión lines," *IEEE Microwave Wireless Comp. Lett.*, Vol. 14, 68–70, Feb. 2004.
8. Baena, J. D., J. Bonache, F. Martín, R. M. Sillero, F. Falcone, T. Lopetegi, M. A. G. Laso, J. García-García, I. Gil, M. F. Portillo, and M. Sorolla, "Equivalent-circuit models for split-ring resonators and complementary split-ring resonators coupled to planar transmission lines," *IEEE Trans. on Microwave Theory and Techniques*, Vol. 53, Issue 4, 1451–1461, April 2005.

Novel Reconfigurable Left-handed Unit Cell for Filter Applications

Branka Jokanovic[1] **and Vesna Crnojevic-Bengin**[2]

[1]Institute IMTEL, Belgrade, Serbia
[2]Faculty of Technical Sciences, Novi Sad, Serbia

Abstract— In this paper a novel, super compact, left-handed unit cell, called ForeS is proposed. ForeS consists of Four Spirals, which are both mutually coupled and end-coupled with microstrip line. Single unit cell is studied together with its modifications that could be realized with simple diode arrangements. By changing bias voltage, different resonant frequencies, insertion losses and spurious responses can be obtained. Simulated characteristics of ForeS exhibit a smaller attenuation, better stop-band characteristic and considerable size reduction in comparison to zeroth-order resonator (ZOR).

1. INTRODUCTION

Recently, revolutionary results were obtained in the field of metamaterials, artificial structures that exhibit electromagnetic properties generally not found in nature. Metamaterials are designed using sub-wavelength particles, whose size is typically smaller then one tenth of the propagating signal wavelength. Due to this fact, quasi-static analysis can be performed and concept of artificial effective media applied. Structures obtained by using this approach can be considered as a continuous medium with effective parameters, namely effective dielectric permittivity and effective magnetic permeability. By properly choosing type and geometrical arrangement of constituent sub-wavelength particles, effective parameters of metamaterials can be made arbitrarily small or large, or even negative.

In the realization of metamaterials two main approaches are widely accepted in microwave community: resonant-type left-handed metamaterials based on thin-wire structure and split-ring resonators (SRRs) [1, 2] which are lossy and narrow-banded and non-resonant transmission line (TL) approach [3, 4] that provides useful tool for design of simultaneously low loss and broad bandwidth structures.

Following the TL approach, a zeroth-order resonator (ZOR) [5] was proposed whose resonant frequency depends only on the circuit elements of the unit cell and not on the physical length of the resonator. This suggests that a zeroth-order resonator could be made arbitrarily small.

In this paper we proposed a LH unit cell (ForeS) that consists of four grounded spirals end-coupled with microstrip line. Overall dimension of this novel particle is $\lambda_g/13$ by $\lambda_g/13$, where λ_g is the wavelength in microstrip line. Generation of ForeS is shown in Fig. 1. This structure is also

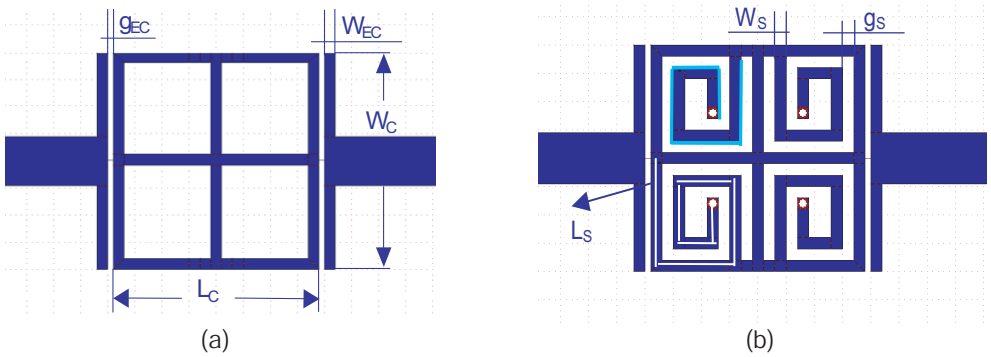


Figure 1: (a) First step in ForeS generation. (b) ForeS view.

very suitable as a constituent element for metasurfaces. Due to its unique construction, the ForeS is possible to combine with diodes providing the structure reconfigurability with proper voltage biasing.

2. SIMULATION RESULTS

In order to compare characteristics of the proposed structure with ZOR, ForeS is simulated on Rogers 5880 substrate $\varepsilon_r = 2.17$, $h = 1.575\,\text{mm}$ and $\text{tg}\delta = 0.0009$. Performances of the structures are determined using IE3D Power Pack Version 10.0 with conductor losses included through $\sigma = 58\,\text{MS/m}$. Dimensions of the cell are optimized to provide an equal resonant frequency as ZOR.

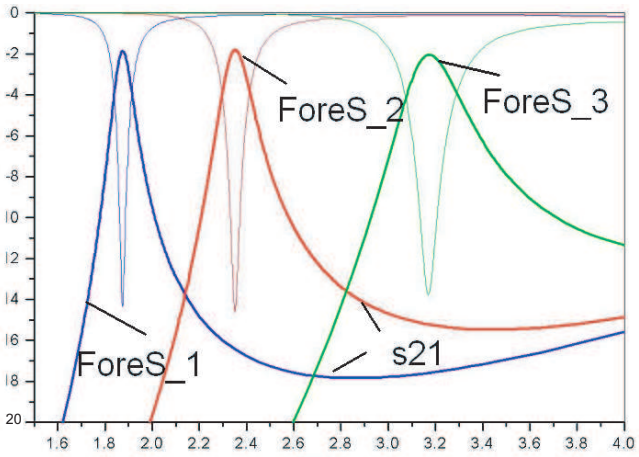


Figure 2: Simulated s_{11} and s_{21} for three different ForSs occupying the same area.

Firstly, we analyzed the ForeS (Fig. 1(b)) having the following dimensions: $L_c = 9.0\,\text{mm}$, $W_c = 9.0\,\text{mm}$, spiral line width $W_s = 0.1\,\text{mm}$, end-coupled gap $g_{EC} = 0.1\,\text{mm}$, spiral gap $g_S = 0.1\,\text{mm}$ and via diameter $d = 0.3\,\text{mm}$. Keeping unit cell length (L_C) and width (W_C) constant as well as end-coupled gap (g_{EC}), width of spiral lines (W_S) and the gap (g_S) are varied to investigate the flexibility of the design. Results of simulations are shown in Table 1. and in Fig. 2.

Table 1: Simulated results of the three different ForeSs occupying the same area.

Resonator	***ForeS_1***	***ForeS_2***	***ForeS_3***
L_C [mm]	9.0	9.0	9.0
W_C [mm]	9.0	9.0	9.0
W_S [mm]	0.4	0.5	0.6
g_S [mm]	0.1	0.1	0.2
g_{EC} [mm]	0.1	0.1	0.1
L_S [mm]	36.7	30.7	22.5
f_{R1} [GHz]	1.876	2.35	3.175
s_{21}^1 [dB]	−1.86	−1.8	−2.05
s_{11}^1 [dB]	−14.17	−14.5	−13.79
B [MHz]	95.5	147.67	290
Q_L	19.64	15.91	10.95
Q_0	33.26	27.32	17.85
f_{R2} [GHz]	7.75	7/8.15	7.9
S_{21}^2 [dB]	−1.12	−2/−1.43	−1.3

By introducing small changes in dimensions of ForeS, significantly different resonant frequencies can be obtained in range of 1.87 GHz to 3.18 GHz that is 59% in respect to the highest frequency. In the same time, all configurations exhibit very small insertion loss, compared to other LH structures. It is seen that ForeS_1 has the best performances: the lowest resonant frequency and the highest Q-factor. Its insertion loss is somewhat higher (−1.86 dB), due to smaller spiral width ($W_S = 0.4\,\text{mm}$). Generally, due to very small gap between cell and input microstrip line (g_{EC}), simulated ForeSs exhibit low insertion loss at the resonance in range −1.8 dB to −2.05 dB and also very good reflection of about −14.5 dB to −13.8 dB.

Simulated results for three different ForeSs in a wide frequency range are presented in Fig. 3. It is shown that variation of spiral width and gap only influences the first resonance, while the second one remains almost unchanged, around 8.0 GHz.

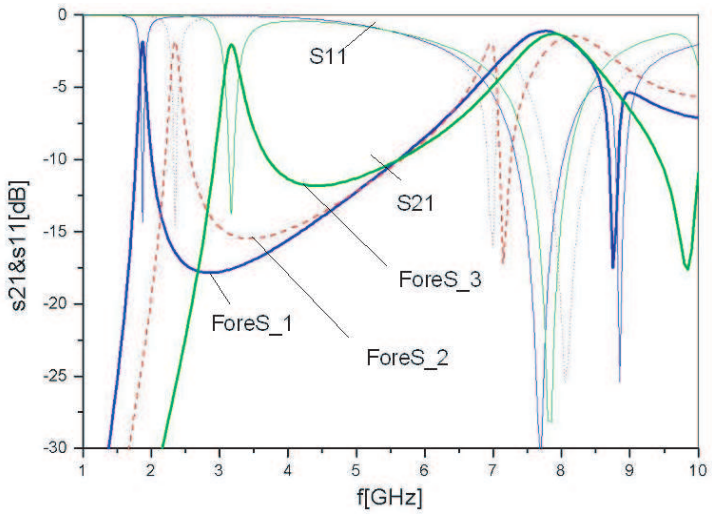


Figure 3: Simulated results for three different ForeSs in wide frequency range.

3. RECONFIGURABILITY OF THE FORES

Figure 4 shows some possible modifications of basic ForeS (Fig. 1). In the simulation, light green patches are considered as metal patches. Electronic reconfigurabilty of the ForeS is possible by

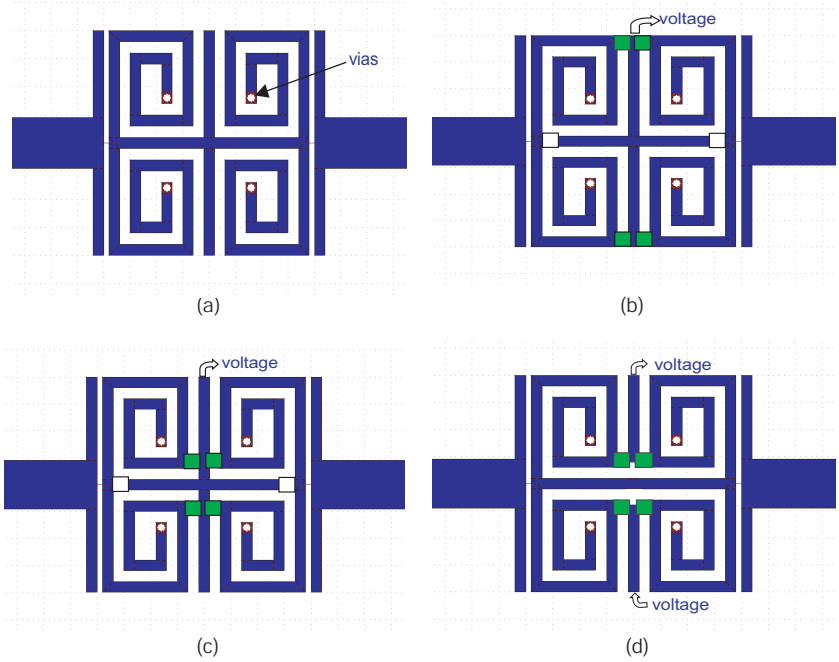


Figure 4: Different modifications of ForeS realized with proper polarization of the diodes.

diodes placed instead of light green patches. Various forms of ForeS could be generated applying very simple biasing as indicated in Fig. 4. Both input ports are DC decoupled, while current return path is provided through the vias that simplifies the biasing circuit.

Different types of diodes could be used to reconfigure the ForeS: PIN diodes, varactor or Schottky barrier diodes. Orientation of each diode in a pair should be an opposite: left and right in respect to the central biasing line so, beam lead anti-parallel diode pairs or common cathode diode

arrangement could be used. White patches represent the gaps which serve to isolate a central biasing line from the ground in Figs. 3. (b) and (c). By switching on/off bias voltage, different resonant frequencies, insertion losses and spurious responses can be obtained.

Four modifications of Fores are simulated using metal patches instead of diodes and results are summarized in Table 2. Physical dimensions are kept the same as in Fores_1 (Table 1).

Table 2: Summarized results of modified ForeSs occupying the same area (9.0×9.0 mm).

ForeS type	***End-Opened Cross***	***End-Connected Cross***	***Center-Connected Cross***	***Center-Disconnected Cross***
f_{R1} [GHz]	1.53	1.876	2.0	1.59
s_{21}^1 [dB]	-1.726	-1.86	-1.87	-1.58
s_{11}^1 [dB]	-14.85	-14.17	-14.3	-15.75
B [MHz]	58.5	95.5	108.3	69
Q_L	26.15	19.64	18.47	20.12
Q_0	45.67	33.26	31.2	42.25
f_{R2} [GHz]	3.72	7.75	6.0	6
S_{21}^2 [dB]	-1.428	-1.12	1.22	-1.17
f_{R3} [GHz]	6.0	-	6.8	6.775
s_{21}^3 [dB]	-1.652	-	-1.02	-1.0

It is seen that structure called "End-Opened Cross" offers the greatest decrease in resonant frequency, but disadvantage is that the second resonance appears at 3.7 GHz. The structure called "Center-Disconnected Cross" demonstrated the best overall characteristics as: low resonant frequency, low insertion loss at the resonance, small reflection (-15.75 dB), a good Q-factor, while the second harmonic appears at 6.0 GHz. The best out of band spurious response is observed in ForeS called "End-Connected Cross" (basic ForeS) because the second harmonic appears at 7.75 GHz. It is noticed that simulated structures have different resonant frequencies in range 1.53 GHz to 2.0 GHz, which covers almost 24% of the tuning range.

Table 3: Comparison of end-coupled resonators: two types of ForeS and ZOR.

Resonator	***End-Connected Cross***	***Center-Disconnected Cross***	***ZOR***
L_C [mm]	9.0	8.4	24.4
g_{EC} [mm]	0.5	0.6	0.2
W_C/L_{stub} [mm]	9.0	8.4	9.9
g_S [mm]	0.1	0.1	-
W_S [mm]	0.4	0.4	-
f_R [GHz]	1.986	1.985	2.013
s_{21}^1 [dB]	-4.75	-4.93	-7.3
B [MHz]	45	35.75	21
Q_L	44	55.5	95.85
Q_0	54	67.2	106.26

4. COMPARISON WITH ZOR

Two structures are chosen for further optimization: End-Connected Cross and Center-Disconnected Cross. During optimization, end-coupled gap, cell length and width are changed to match characteristics of ZOR as close as possible. The results are shown in Table 3. In comparison to ZOR, both optimized ForeSs demonstrate the considerable size reduction, especially great reduction is observed in resonator length (L_C). Considerably lower attenuation at resonance is noticed (-4.75 dB and -4.93 dB in respect to -7.3 dB at ZOR). Also, ZOR s_{21}-characteristic has spurious responses at 3.695 GHz (-9.7 dB), and 4.325 GH (-6.0 dB). The second resonance of End-Connected Cross appears at 8.0 GHz, which means spurious free region up to the 4th harmonic.

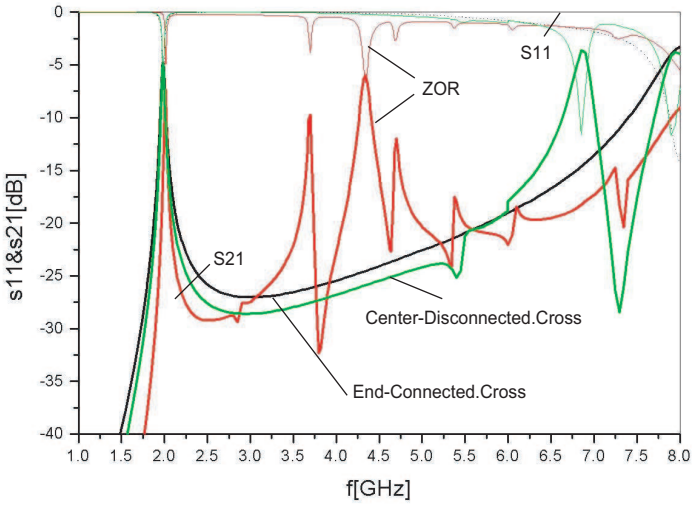


Figure 5: Simulated s_{11} and s_{21} for two types of ForeSs and for ZOR.

5. CONCLUSION

In this paper a novel super compact LH resonator (ForeS) with small insertion loss is proposed. It consists of four spirals, which are both mutually coupled and end-coupled with microstrip line. It is shown that keeping overall dimensions of unit cell constant as well as the gap between the cell and microstrip line, resonant frequency can be changed from 1.87 GHZ to 3.18 GHz only by changing the spiral width and gap. Insertion loss at resonance remains almost unchanged (-1.8 dB to -2.0 dB), while 3-dB-bandwidth changes from 95 MHz to 290 MHz. Four different versions of ForeS are investigated to find the best ones to compare with zeroth-order resonator. Two chosen modifications of ForeS demonstrated considerable low insertion loss, better out-of-band suppression of unwanted harmonics and tremendous size reduction in comparison to ZOR.

It is also shown that the ForeS can be modified electronically by simple diode arrangement and simple biasing which gives additional freedom in changing resonant frequency around 24% (from 1.53 GHz to 2.0 GHz), Q-factor and out-of-band response.

ACKNOWLEDGMENT

This work is supported by Eureka program (METATEC-METAmaterialCbased TEchnology for broadband wireless Communications and RF identification E! 3853 project).

REFERENCES

1. 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, No. 18, 4184–4187, May 2000.
2. Marqués, R., J. Martel, F. Mesa, and F. Medina, "Left handed media simulation and transmission of EM waves in subwavelength SRR-loaded metallic waveguides," *Phys. Rev. Lett.*, Vol. 89, 183901-03, 2002.
3. Eleftheriades, G. V., O. Siddiqui, and A. K. Iyer, "Transmission line models for negative refractive index media and associated implementation without excess resonators," *IEEE Microwave Wireless Compon. Lett.*, Vol. 13, 51–53, Feb. 2003.
4. Lai, A., C. Caloz, and T. Itoh, "Composite right/left-handed transmission line metamaterials," *IEEE Microwave Magazine*, Vol. 5, No. 3, 34–50, September 2004.
5. Sanada, A., C. Caloz, and T. Itoh, "Zeroth order resonance in composite righ/left-handed transmission line resonators," *Proc. Asia-Pacific Microwave Conference*, Seoul, Korea, Vol. 3, 1588–1592, 2003.

A CRLH Microstrip Delay Line for High-speed Electronic Circuits

S. Sebak, L. Zhu, V. K. Devabhaktuni, and C. Wang

Department of ECE, Concordia University
1455 de Maisonneuve West, Montreal, H3G 1M8, Canada

Abstract— In this paper, we present a composite right/left-handed microstrip delay line, which simultaneously exhibits negative refractive index and negative group-delay. The proposed structure shows advantages of compact dimensions, lower loss, and longer delay, when compared to traditional right-handed and other left-handed structures. Based on a simulation study, a design methodology that helps the users to tune the group-delay of a given circuit is developed. Delay minimization of an RF power amplifier circuit is presented as a potential application.

1. INTRODUCTION

Research on radio-frequency (RF) delay lines for high-speed/wide-band applications is becoming important in the context of modern electronic-warfare systems, satellite-communication systems, and signal-processing systems, in which memory storage of signals is a necessity [1–3]. Delay lines based on microstrip transmission lines show lower loss and wider band-width, but fail to achieve longer delay, owing to high wave-velocity of RF/microwave signals traveling through such lines.

About thirty years ago, *Veselago* predicted that electromagnetic (EM) plane waves, in a medium with negative permittivity (ε) and negative permeability (μ), propagate in a direction opposite to that of the energy-flow [4]. Recent studies have shown that left-handed (LH) structures exhibit both negative and positive group-delays in the negative refractive index (NRI) region. Group-delay is related to group-velocity (v) and length of the microstrip line (L), or to frequency derivative of the transmission phase (φ), as

$$\tau = \frac{L}{v} = -\frac{\partial \varphi}{\partial \omega}. \tag{1}$$

In this paper, we present a composite right/left-handed (CRLH) delay line utilizing interdigital capacitors, which offers longer delay compared to traditional right-handed (RH) and other LH microstrip delay lines. The proposed structure is compact, and its group-delay can be tuned by adjusting physical parameters of the interdigital capacitors. Based on a simulation study, a design methodology has been developed to help users easily tune the group-delay of the line/structure. One of the useful characteristics of the proposed structure (i. e., relatively longer negative group-delay) is exploited for reducing group-delay in an electronic circuit.

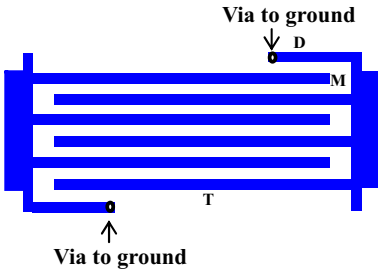


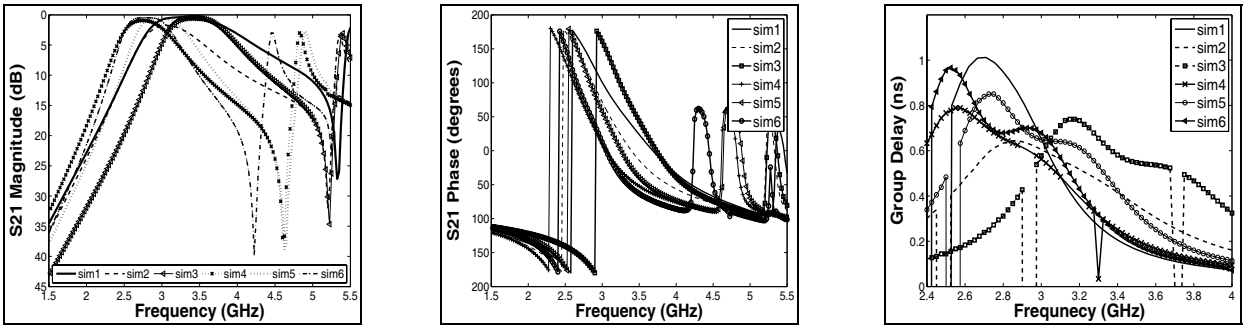
Figure 1: CRLH structure [6] with $D = 0.5\,\text{mm}$, $T = 5\,\text{mm}$, $M = 0.2\,\text{mm}$, $N = 8$, and $\varepsilon = 12.9$.

2. PROPOSED CRLH DELAY LINE

Any left-handed transmission line (LH-TL) is a CRLH structure [5], i. e., LH at low-frequencies and RH at high-frequencies, and exhibits band-pass characteristics. The CRLH structure [6] is shown in Fig. 1. By adjusting the physical dimensions including length of outer arms (D), length of fingers (T), width of fingers (M), and number of fingers (N), any given group-delay specification can be achieved. Using Zeland's *IE3D* software tool, we performed EM simulations for different geometries as shown in Table 1. Results of the simulations are shown in Figs. 2(a)–(c). A close inspection of these results in the $1.5 \sim 5.5\,\text{GHz}$ range has resulted in the following observations.

Table 1: Different combinations of physical parameters for which EM data is collected. (M, T and D are in mm).

	Sim. 1	***Sim. 2***	***Sim. 3***	***Sim. 4***	***Sim. 5***	***Sim. 6***
M	0.4	0.2	0.2	0.2	0.2	0.2
N	6	6	8	8	8	8
T	4.5	4.5	4.5	5	5	5
D	1.5	1.5	1.5	1.5	0.9	0.49



(a) (b) (c)

Figure 2: Comparison of (a) S_{21} magnitude, (b) S_{21} phase, and (c) group-delay for different geometries shown in Table 1.

Case 1 (simulations 1 and 2): When M is decreased from 0.4 mm to 0.2 mm, group-delay decreases while both centre frequency and bandwidth increase considerably. It is to be noted that group-delay $\propto M$.

Case 2 (simulations 2 and 3): When M is kept constant and N is increased from 6 to 8, group-delay increases significantly. Once again, it is to be noted that group-delay $\propto N$.

Case 3 (simulations 3 and 4): When M and N are kept constant, and T is increased from 4.5 mm to 5 mm, both group-delay and pass-band increase marginally.

Case 4 (simulations 4 and 5): When D is decreased from 1.5 mm to 0.9 mm, both magnitude and phase of S_{21} change marginally. As such, D can be used to adjust/fine-tune impedance matching.

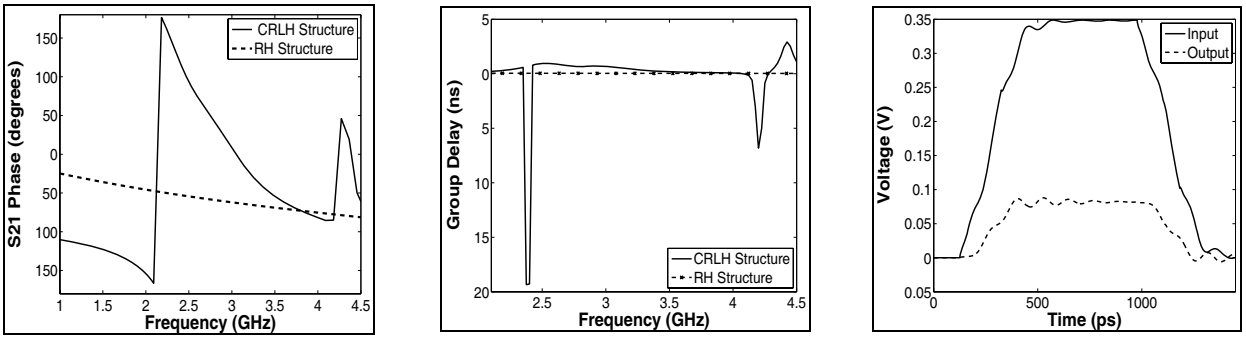
Case 5 (simulations 5 and 6): Keeping all other design parameters constant, D alone is adjusted to 0.49 mm resulting in a perfect matching. This adjustment does not affect group-delay.

3. COMPARISON OF VARIOUS STRUCTURES

Compared to traditional RH delay lines, the proposed CRLH delay line is relatively more compact, while offering a much longer group-delay as can be inferred from the sharper slope of $\varphi(S_{21})$ throughout the frequency range of interest (see Fig. 3(a)). Simulations using Zeland's *MDSPICE* software show that both negative and positive group-velocities can be achieved within the NRI band. As can be seen in Fig. 3(b), positive group-delay of the RH delay line is almost zero (≈ 0.03 ns) over the entire frequency range, whereas that of the CRLH delay line is 0.9 ns (≈ 27 times longer). In the frequency ranges $2.2 \sim 2.4$ GHz and $4.1 \sim 4.2$ GHz, negative group-delay is observed to be -19.4 ns and -6.5 ns respectively for the proposed structure, both of which are longer than other LH structures, e. g., [7]. The time-domain response of the proposed structure in Fig. 3(c) shows that the output waveform leads the input waveform indicating a negative "time-delay" (approx. 0.5 ns).

4. APPLICATION EXAMPLE

In essence, the proposed CRLH structure provides the user with certain "degrees of freedom" (i. e., physical parameters) to achieve a given group-delay. In the design/optimization of high-speed/high-frequency circuits, controlling the group-delay of output waveforms is critical [8]. As an application of this work, we advocate cascading of the proposed CRLH delay line with a given electronic circuit, without considerably altering the circuit responses other than group-delay. As can be seen in the



(a) (b) (c)

Figure 3: Comparison of (a) S_{21} phase and (b) group-delay for the CRLH and the traditional right-handed structures, and (c) schematic showing input/output voltage waveforms for the CRLH structure.

example that follows, the design methodology developed based on the simulation study becomes useful.

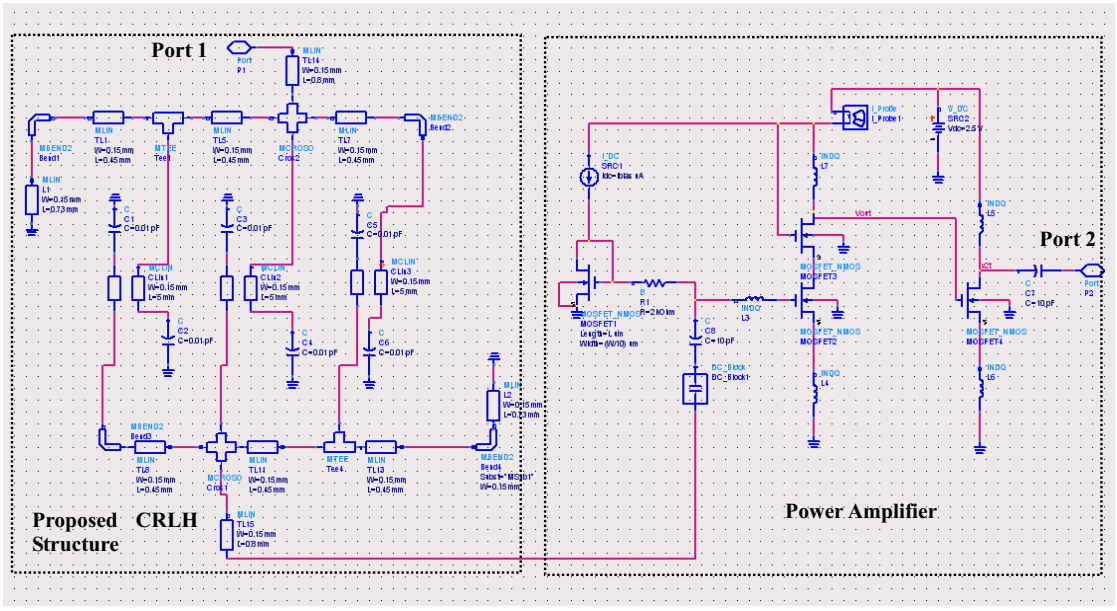


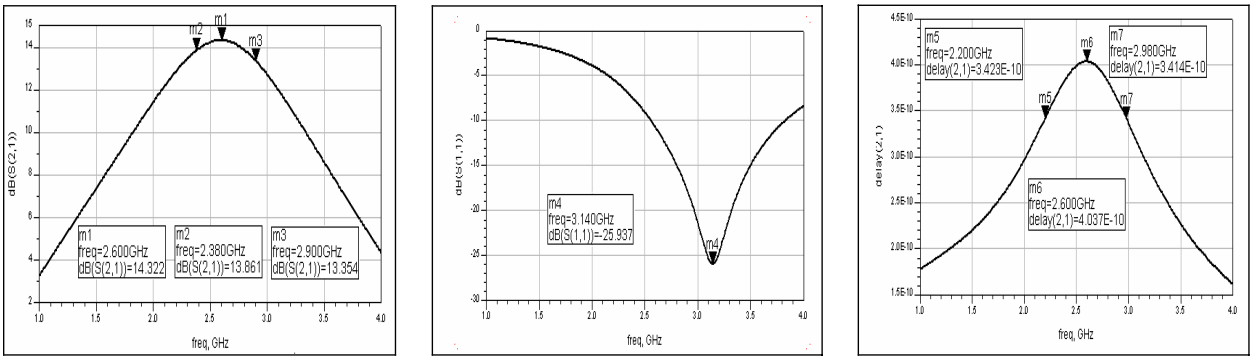
Figure 4: Circuit schematic showing the proposed CRLH structure cascaded with the given power amplifier circuit.

In this application example, a power amplifier (PA) circuit operating in the $1 \sim 4$ GHz range is considered. As can be seen in Fig. 5(a) and Fig. 5(b), the gain of the amplifier is 14.3 dB at the centre frequency 2.6 GHz, and its return loss is greater than -15 dB in the $2.6 \sim 3.5$ GHz range. Frequency-domain simulations of the PA circuit in Agilent's *Advanced Design System* (*ADS*) indicate a positive group-delay of 0.3–0.4 ns in the $2.2 \sim 3.0$ GHz range (see Fig. 5(c)). The objective here is to reduce this delay using the proposed CRLH delay line.

We start the process of investigating group-delay by studying the behavior of the CRLH structure exclusively in the $1 \sim 4$ GHz range. Conceptually, the (modified) objective is to achieve a negative group-delay by tweaking its physical parameters. Initial values of the parameters are set to be $M = 0.4$ mm, $N = 6$, $T = 5$ mm, and $D = 1.5$ mm. First, M is decreased to 0.15 mm in order to achieve a wider pass-band, similar to that of the PA circuit. Second, N is increased to 8 so as to adjust the centre frequency of the CRLH structure to 2.75 GHz (which is closer to the center frequency of the PA), while maintaining an average group-delay of -0.6 ns in the $2.2 \sim 3.0$ GHz range. Third, a "fine-adjustment" is performed by increasing T to 5.14 mm such that group-delay of the CRLH delay line increases to -0.4 ns and its centre frequency reaches 2.6 GHz. Finally, parameter D is adjusted to 0.73 mm in order to lower the return loss. As a result of this systematic/step-by-step

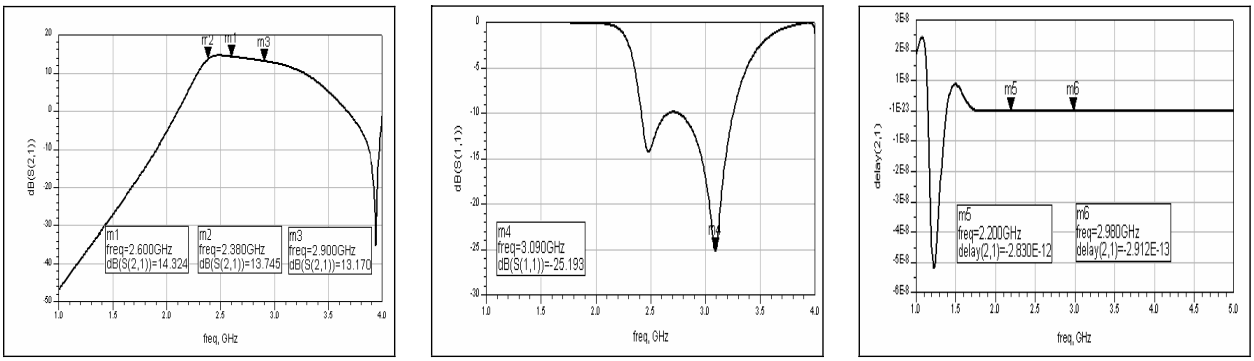
design process, the optimized CRLH delay line is seen to have an insertion loss < 0.1 dB in the $2.2 \sim 3.0$ GHz range and a return loss > -20 dB in the $2.5 \sim 2.8$ GHz range. To summarize, the most important result here is that the optimized CRLH delay line shows an average group-delay of -0.39 ns in the frequency range of interest.

The optimized CRLH structure is cascaded with the PA circuit as shown in Fig. 4. Simulation results of the overall circuit in *ADS* are presented in Figs. 6(a)–(c). A comparison of Figs. 6(a)–(c) with Figs. 5(a)–(c) shows that cascading the CRLH delay line has not affected the amplifier response considerably. For instance, gain of the overall circuit is greater than 13 dB in the frequency range $2.4 \sim 2.9$ GHz and its return loss is lower than -25 dB at 3.1 GHz. The group-delay of the overall circuit is reduced to < 0.3 ps as can be seen in Fig. 6(c), and this reduction is significant considering the original group-delay of the given PA circuit.



(a) (b) (c)

Figure 5: Power amplifier simulations showing (a) magnitude of S_{21}, (b) magnitude of S_{11}, and (c) group-delay.



(a) (b) (c)

Figure 6: Combined EM-circuit responses showing (a) magnitude of S_{21}, (b) magnitude of S_{11}, and (c) group-delay.

5. CONCLUSIONS

In this paper, a new CRLH delay line has been presented. EM simulations of the proposed line indicate not only NRI, but also negative and positive group-delays much longer than traditional RH and LH structures. The CRLH delay line can be cascaded with a PA circuit to minimize its group-delay. It has been shown that such an approach does not alter/affect the original PA circuit characteristics such as gain and return loss. Dispersive properties of LH materials make LH structures (e. g., the proposed delay line) suitable for efficiently controlling dispersive effects in electronic circuits and dispersion management in high-speed interconnects. This work can be useful for high-speed/wide-band system optimization.

ACKNOWLEDGMENT

The authors acknowledge useful discussions with Dr. R. Raut of the Department of Electrical and Computer Engineering, Concordia University, Montreal, Canada.

REFERENCES

1. Toughlian, E. N. and H. Zmuda, "A photonic variable RF delay line for phased array antennas," *J. Lightwave Tech.*, Vol. 8, 1824–1828, 1990.
2. Tantawi, S. G., R. D. Ruth, A. E. Vlieks, and M. Zolotorev, "Active high-power RF pulse compression using optically switched resonant delay lines," *IEEE Trans. Microwave Theory Tech.*, Vol. 45, 1486–1492, 1997.
3. Su, H. T., Y. Wang, F. Huang, and F. Lancaster, "Characterizing a double-spiraled meander superconducting microstrip delay line using a resonator technique," *Proc. IEEE MTT-S Int. Microwave Symp.*, 135–138, Fort Worth, TX, June 2004.
4. Veselago, V. G., "The electrodynamics of substance with simultaneously negative values of ε and μ," *Soviet Physics Uspekhi*, Vol. 10, 509–514, 1968.
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 Propagat.*, Vol. 52, 1159–1166, 2004.
6. Zhu, L., V. K. Devabhaktuni, and C. Wang, "CAD of left-handed transmission line bandpass filters," *Proc. Progress in Electromagnetics Research Symp.*, Accepted, Beijing, China, April 2007.
7. Siddiqui, O., M. Mojahedi, and G. Eleftheriades, "Periodically loaded transmission line with effective refractive index and negative group velocity," *IEEE Trans. Antennas Propagat.*, Vol. 51, 2619–2625, 2003.
8. Kim, S., N. Gopal, and L. Pillage, "Time-domain macromodels for VLSI interconnect analysis," *IEEE Trans. Antennas Propagat.*, Vol. 13, 1257–1270, 1994.

Design and Application of a VNA-based Polarimetric SAR for Deep Soil Moisture Estimation

Wenji Zhang and Xiaojuan Zhang
Institute of Electronics, Chinese Academy of Sciences, Beijing 100080, China

Abstract— A vector network analyzer (VNA) based polarimetric VHF/UHF synthetic aperture radar (SAR) for deep soil moisture estimation is proposed. It can penetrate into the soil for about 0.25–3 m, even in the presence of vegetation. The VHF/UHF SAR data can be used for retrieval of deep soil moisture. Due to the large penetration depth of SAR in the VHF/UHF band, the soil is modeled as a multilayer structure with rough interfaces. The backscattering of electromagnetic waves from multilayer soil is studied.

1. INTRODUCTION

Synthetic aperture radar (SAR) is a coherent active microwave imaging technology. It is widely used in the field of active microwave remote sensing for mapping the properties of the earth surface. The characterization and classification of land and forest using polarimetric SAR data has been extensively invested since 1950's.

In the past decades in order to achieve high range and azimuth resolution most of the SAR developed are working in the S, C, X, Ku bands. In the high frequencies the microwave has very limited penetration ability. In order to penetrate into deep soil and dense foliage an increasing amount of studies has been carried out on the design and application of low frequency SAR, such as HF/VHF/UHF or even lower [1]. Low frequency radar is proved to have a very powerful ability for underground and obscured object detection. It can be used to discover target buried under deep ground and tank or airplane under dense foliage.

In this paper a vector network analyzer (VNA) based dual polarized VHF/UHF SAR for deep soil moisture estimation is proposed. In the low frequency the radar signal can penetrates into the soil for about several meters, depending on the soil type and moisture. Thus the soil is modeled as a multilayer rough surface and the backscattering of multilayer soil is studied.

2. VHF/UHF VNA-BASED SAR SYSTEM

Many physical and geometric parameters of the ground contribute to the grey value of a SAR image pixel. Accurate inversion of soil properties requires SAR data taken at different frequencies, polarization, incident angles. At present no such air or space borne low frequency SAR data can be used to inverse the properties of deep soil. In order to study the scattering of soil and vegetation in the VHF/UHF band we are developing a vector network analyzer based SAR [2]. Fig. 1 shows the configuration of our VNA-based SAR system.

The VNA-based SAR operated at frequencies from 100 to 500 MHz in a stepped frequency continuous wave mode. The system consisted of a vector network analyzer, a microwave amplifier, a pair of dual polarized log-period antenna, an antenna positioner, a 25 meters vertical pole mounted on a truck, and a control program. The vector network analyzer (HP8753ES) was used as transmitter and receiver, which generated the transmitting signal from 100 to 500 MHz and recorded the receiving signal both in amplitude and phase. In the frequencies between 100–500 MHz the backscattering from soil is relatively weak and a microwave amplifier was used to achieve a transmitting power of 20 dBm. The log-period antenna is a dual polarized broadband antenna, which operates between 50–800 MHz and has a gain of 8 dB. The synthetic aperture was achieved by the movement of the truck in the horizontal direction and the moving of antennas up and down along the vertical pole controlled by an antenna positioner. The scanning aperture was 20 m in the horizontal and 15 m in the vertical direction. The control program was specially designed to record and store the data uploaded from the vector network analyzer and control the antenna positioner. A metallic sphere and trihedral was used to calibrate the radar system.

3. MULTILAYER SOIL SCATTERING

The penetration depth of electromagnetic wave is defined as the transmission distance when the power of transmitting signal attenuated to 1/e of the incident wave. In the S, C, X, Ku bands

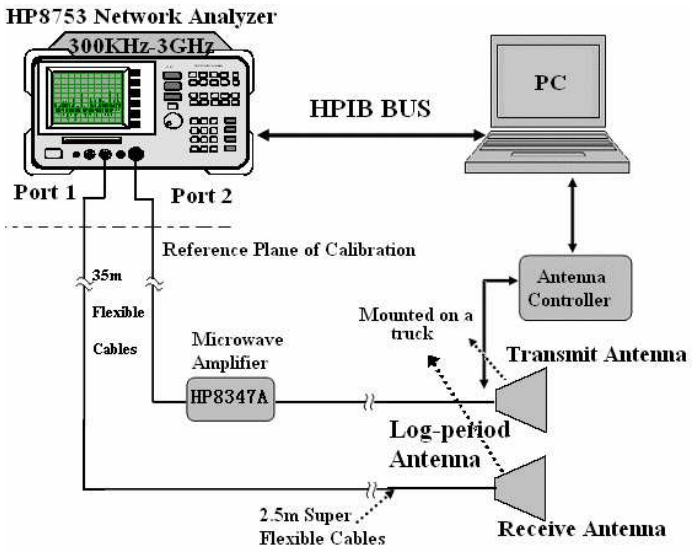


Figure 1: Configuration of the VNA-based SAR.

microwave can only penetrate into soil for no more than 10 cm. The main scattering comes from the up-layer soil. So the scattering model, such as SPM, KA, IEM, is always modeled as a two-layer rough surface with an air-soil interface [4]. In the VHF/UHF band the scattering from subsurface and the multiple scattering between up and subsurface has to be taken into account. In order to describe the scattering mechanism of soil in the VHF/UHF band more accurately, we consider a simplest case. In this case the soil is modeled as a three-layer structure with an air-soil interface and a smooth interface between different types of soil, as shown in Fig. 2. The air-soil rough interface is denoted by a random process $z = f(x, y)$ and the interface between different soils is assumed to be smooth.

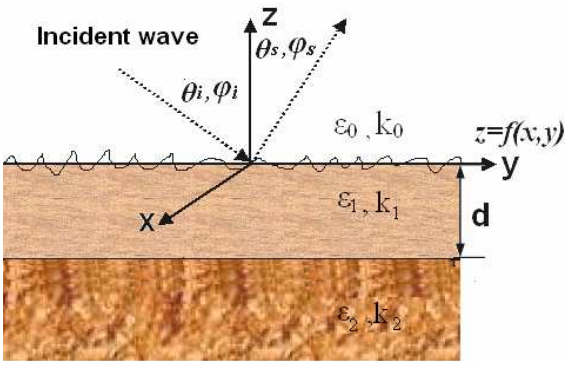


Figure 2: Three-layer soil structure.

Assume the incident plane wave is

$$\vec{E}_i = \vec{E}_0 e^{ik_{ix}x - ik_{iz}z} \tag{1}$$

where $k_{ix} = k_i \sin\theta_i$, $k_{iz} = k_i \cos\theta_i$, k_0, k_1, k_2 are respectively wave number in each layer. The co-polarized backscattering coefficient can be written as [3]:

$$\sigma_{hh} = 16k^4 \cos^4\theta_i \left|\alpha_{Lhh}\right|^2 W(2k\sin\theta_i, 0); \quad \sigma_{vv} = 16k^4 \cos^4\theta_i \left|\alpha_{Lvv}\right|^2 W(2k\sin\theta_i, 0) \tag{2}$$

for HH polarization:

$$\alpha_{Lhh} = \alpha_{shh}\left(\frac{1 + R_{dhh}}{1 + R_{uhh}R_{dhh}}\right)^2 \tag{3}$$

where

$$\alpha_{shh} = -\frac{\cos\theta_i - \sqrt{\varepsilon_1 - \sin^2\theta_i}}{\cos\theta_i + \sqrt{\varepsilon_1 - \sin^2\theta_i}} \tag{4}$$

R_{uhh} is the Fresnel reflection coefficient from the boundary $z = f(x, y)$:

$$R_{uhh} = \frac{\cos\theta_i - \sqrt{\varepsilon_1 - \sin^2\theta_i}}{\cos\theta_i + \sqrt{\varepsilon_1 - \sin^2\theta_i}} \tag{5}$$

R_{dhh} is the lower boundary ($z = -d$) reflection coefficient:

$$R_{dhh} = r_{dhh}e^{i\varphi} \tag{6}$$

where

$$i\varphi = 2k_1 d\,(i\beta - \gamma)\,; \quad \beta = \mathrm{Re}\left(\sqrt{\varepsilon_1 - \sin^2\theta_i}\right); \quad \gamma = \mathrm{Im}\left(\sqrt{\varepsilon_1 - \sin^2\theta_i}\right) \tag{7}$$

r_{dhh} is the Fresnel reflection coefficient from boundary $z = -d$:

$$r_{dhh} = \frac{\sqrt{\varepsilon_1}\cos\theta_t - \sqrt{\varepsilon_2 - \sin^2\theta_i}}{\sqrt{\varepsilon_1}\cos\theta_t + \sqrt{\varepsilon_2 - \sin^2\theta_i}} \tag{8}$$

By Snell's law

$$\theta_t = \arcsin(\sin\theta_i/\sqrt{\varepsilon_1}) \tag{9}$$

for VV polarization:

$$\alpha_{Lvv} = \alpha_{svv}\frac{1 + R_{dvv}^2 + 2R_{dvv}\dfrac{\sin^2\theta_i - \cos^2\theta_t}{\sin^2\theta_i + \cos^2\theta_t}}{(1 + R_{uvv}R_{dvv})^2} \tag{10}$$

where

$$\alpha_{svv} = (1 - \varepsilon_1)\frac{\sin^2\theta_i + \ cos^2\theta_t}{(\sqrt{\varepsilon_1}\cos\theta_i + \cos\theta_t)^2} \tag{11}$$

R_{uvv} is the Fresnel coefficient from the boundary $z = f(x, y)$:

$$R_{uvv} = \frac{\varepsilon_1\cos\theta_i - \sqrt{\varepsilon_1 - \sin^2\theta_i}}{\varepsilon_1\cos\theta_i + \sqrt{\varepsilon_1 - \sin^2\theta_i}} \tag{12}$$

R_{dvv} is the lower boundary reflection coefficient:

$$R_{dvv} = r_{dvv}e^{i\varphi} \tag{13}$$

where r_{dvv} is the Fresnel reflection coefficient from boundary $z = -d$:

$$r_{dvv} = \frac{\varepsilon_2\cos\theta_t - \sqrt{\varepsilon_1(\varepsilon_2 - \sin^2\theta_i)}}{\varepsilon_2\cos\theta_t + \sqrt{\varepsilon_1(\varepsilon_2 - \sin^2\theta_i)}} \tag{14}$$

To validate this model we assume $d \to \infty$ or $\varepsilon_1 = \varepsilon_2$ it should turn to the SPM, Fig. 3 shows that they have a perfect match. Using the deduced scattering model we get some simulation results. Fig. 4 shows the backscattering of soil at different frequencies with the second layer thickness d=2 m. Fig. 5 shows the effect of second layer thickness on the backscattering coefficient. Fig. 6 shows co-polarization ratio at different incident angles with soil dielectric constant (soil moisture) of the second layer.

4. SUMMARY

A vector network analyzer (VNA) based dual polarized VHF/UHF SAR for deep soil moisture estimation is presented in this paper. It provides a convenient method for measuring the scattering of soil under laboratory conditions and provides data for the inversion of deep soil moisture. Combination with the multilayer rough surface scattering model the VNA-based SAR data can be used to inverse deep soil moisture. In order to eliminate soil roughness influence co-polarization ratio data is used to inverse deep soil moisture and depth. Field experiments are scheduled for next spring and available results will be presented.

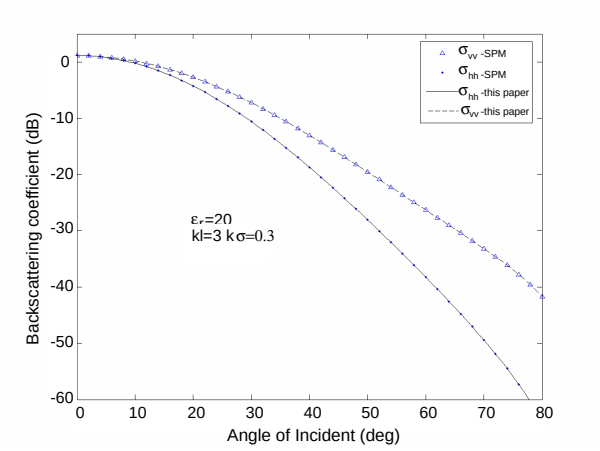


Figure 3: Validation of this model.

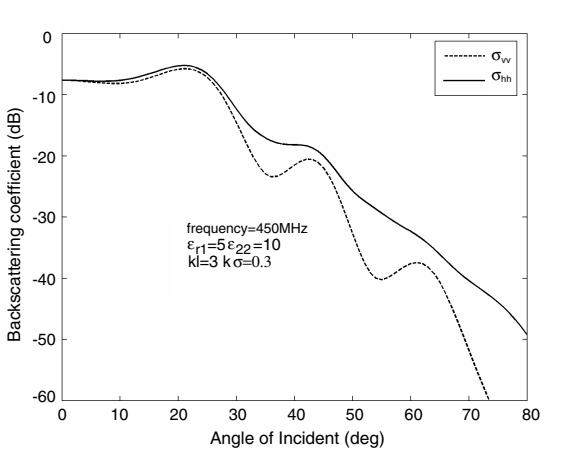


Figure 4: Backscattering coefficient at different incident angles ($d = 2\,\mathrm{m}$).

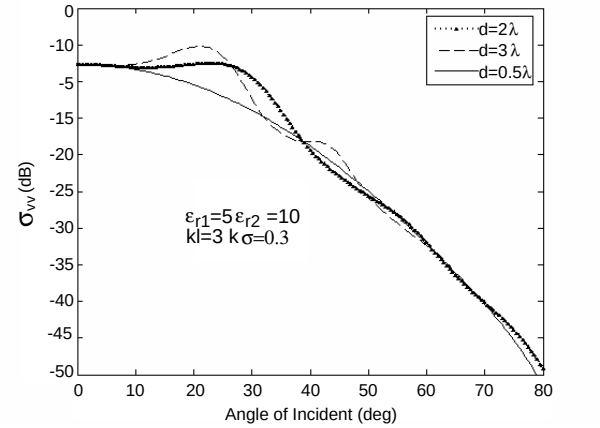


Figure 5: Influence of second layer thickness on σ_{vv}.

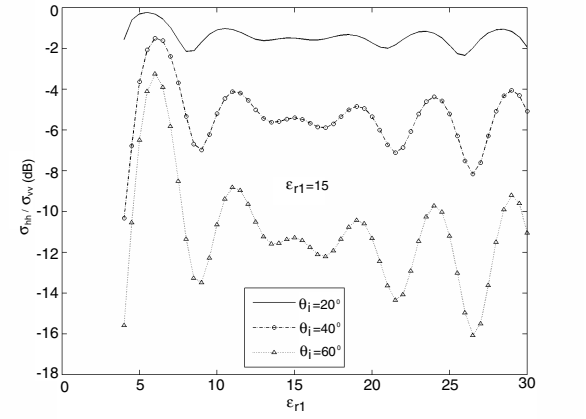


Figure 6: Influence of second layer dielectric on σ_{hh}/σ_{vv}.

ACKNOWLEDGMENT

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

REFERENCES

1. Moller, Rodriguez, Moghaddam, and Hoffman, "A Dual-low-frequency radar for sub-canopy and deep soil moisture measurements," *Big Sky Conference*, March 2003.
2. Zhou, Z. S., V. Boerner, and M. Sato, "Development of a ground-based polarimetric broadband sar system for non-invasive ground-truth validation in vegetation monitoring," *IEEE Trans. Geoscience and Remote Sensing*, Vol. 42, No. 9, 1803–1810, 2004.
3. Fuks, I. M., "Radar contrast polarization dependence on subsurface sensing," *Proc IGARSS*, Vol. 98, 1455–1459, 1998.
4. Ulaby, F. T., R. K. Moore, and A. K. Fung, *Microwave Remote Sensing*, Vol. 2, Chaper, 11, Artech house, 1990.

Inversion of Permittivities for Layered Rough Surfaces

Xiaorong Lu[1,2], **Xiaojuan Zhang**[2], **and Guangyou Fang**[2]

[1]Graduate University of Chinese Academy of Science, Beijing 100049, China
[2]Key Lab of High Power Microwave and EM Radiation, IECAS, Beijing 100049, China

Abstract— Nelder-Mead simplex optimization algorithm is used to retrieve the permittivities of surface and subsurface of bared soil, based on a dual-frequency multi-polarization radar system. Although the algorithm is not the typical Newton-type algorithm which is often used, it can not only get quite good results, but also take relatively small amount of time to calculate. So if we need a real-time algorithm, this algorithm is a good choice.

1. INTRODUCTION

Since microwave is used to monitor the earth, it has been applied to estimate moisture of soil. Not only water precipitation, evaporation, and circulation can be obtained by monitoring moisture change in large scale, but also the absolute value of moisture by transmitting, receiving and analyzing the electromagnetic signals. However, former work has been focused on retrieving the surface soil moisture. And now we are on the subsurface soil moisture by using a dual-frequency multi-polarization radar system. Its frequency bands are UHF/VHF. Since they are lower than most of the former used bands, we can acquire the deeper soil moisture [1]. And the next three parts are focusing on how to retrieve deep soil moisture, including relationship between permittivity and moisture, forward scattering model, inversion algorithm.

2. RELATIONSHIP BETWEEN PERMITTIVITY AND MOISTURE

In the 1970's, Ulaby get the relationship between moisture and permittivity of soil in the 1.4–18 GHz range, based on measured data [2]. Subsequently they get the dielectric properties in the 0.3–1.3 GHz range [3]. Figure 1 shows the curves in UHF band. The frequency is assumed to be 450 MHz, and the percentages of sand and clay are 31.51% and 43.43% respectively. It is obvious that the real part of dielectric constant is more sensitive to volumetric moisture than the imaginary part. And when the permittivity is gotten, moisture then can be retrieved. So in the forth part, only the inversion of permittivity and depth is discussed.

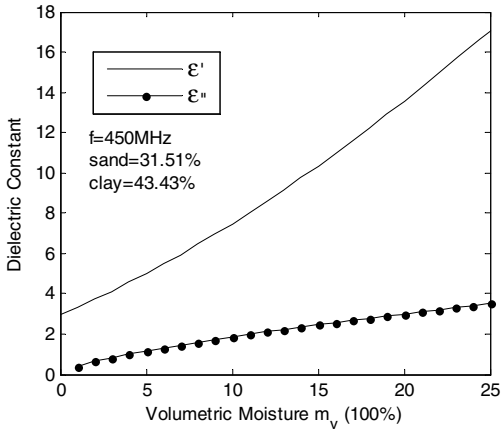


Figure 1: Relationship between moisture and permittivity of soil in range 0.3–1.3 GHz.

3. FORWARD MODEL

The models of scattering from rough surface can be classified as academic and experiential models. The academic scattering model includes SPM (Small Perturbation Model) [4], Kirchhoff approximation [5], etc. As SPM is the same with most of farmland, it is also used here. First order

solution of SPM is as follows. When permeability equals to 1, the backscattering coefficients can be calculated through following equation.

$$\sigma_{pq}^{r} = 8k^4\sigma_1^2\cos^4\theta\left|\alpha_{pq}\right|^2 W\left(2k\sin\theta, 0\right) \tag{1}$$

where

$$\alpha_{hh} = \frac{(1-\varepsilon)}{\left[\cos\theta + \sqrt{\varepsilon - \sin^2\theta}\right]^2}, \quad \alpha_{vv} = (1-\varepsilon)\frac{\sin^2\theta - \varepsilon\left(1+\sin^2\theta\right)}{\left[\varepsilon\cos\theta + \sqrt{\varepsilon - \sin^2\theta}\right]^2}$$
$$\alpha_{hv} = \alpha_{vh} = 0$$

and $W\left(2k\sin\theta, 0\right)$ is the Fourier transformation of surface correlation coefficient.

SPM can be extended to layered rough surface [5]. The improved model is as follows.

$$\sigma_{pq}^{r} = 8k^4\sigma^2\cos^4\theta\left|\alpha_{pq}\right|^2 W\left(2k\sin\theta, 0\right) \tag{2}$$

where

$$\alpha_{Lhh} = \alpha_{shh}\left(\frac{1+\tilde{R}_{1hh}}{1+R_{hh}\cdot\tilde{R}_{1hh}}\right)^2, \qquad \alpha_{Lvv} = \alpha_{svv}\frac{1+\tilde{R}_{1vv}^2 + 2\tilde{R}_{1vv}\frac{\sin^2\theta-\cos^2\theta'}{\sin^2\theta+\cos^2\theta'}}{(1+R_{vv}\tilde{R}_{1vv})^2}$$
$$R_{hh} = \frac{\cos\theta - \sqrt{\varepsilon-\sin^2\theta}}{\cos\theta + \sqrt{\varepsilon-\sin^2\theta}}, \qquad R_{vv} = \frac{\varepsilon\cos\theta - \sqrt{\varepsilon-\sin^2\theta}}{\varepsilon\cos\theta + \sqrt{\varepsilon-\sin^2\theta}}$$
$$\tilde{R}_{1hh} = \frac{\sqrt{\varepsilon}\cos\theta' - \sqrt{\varepsilon_1-\sin^2\theta}}{\sqrt{\varepsilon}\cos\theta' + \sqrt{\varepsilon_1-\sin^2\theta}}e^{i\varphi}, \qquad \tilde{R}_{1vv} = \frac{\varepsilon\cos\theta' - \sqrt{\varepsilon(\varepsilon_1-\sin^2\theta)}}{\varepsilon\cos\theta' + \sqrt{\varepsilon(\varepsilon_1-\sin^2\theta)}}e^{i\varphi}$$
$$i\varphi = 2kH(i\beta - \gamma), \qquad \beta = Re\sqrt{\varepsilon-\sin^2\theta}, \gamma = \mathrm{Im}\sqrt{\varepsilon-\sin^2\theta}$$

θ^t is the refraction angle, and it obeys *Snell*'s law $\sin\theta' = \sin\theta/\sqrt{\varepsilon}$.

4. INVERSION ALGORITHM

In our work Nelder-Mead simplex optimization algorithm, not the Newton-type algorithms often used, is employed. It can be applied to optimize unconstrained multi-variable non-linear function. For the single roughed surface, the forward model is explained by Equation (1), and we can erase the impact of surface parameters easily through getting the rate of two kinds of co-polarization. So we make use of the following cost function.

$$f = \left|\sqrt{\frac{\sigma_{hh}}{\sigma_{vv}}} - \left|\frac{\left(\varepsilon_r\cos\theta + \sqrt{\varepsilon_r - \sin^2\theta}\right)^2}{\left(\cos\theta + \sqrt{\varepsilon_r - \sin^2\theta}\right)^2\left(\varepsilon_r\sin^2\theta + \varepsilon_r - \sin^2\theta\right)}\right|\right| \tag{3}$$

Since the backscattering coefficients is less sensitive to variation of imaginary part of dielectric constant than real part [6]. And according to part 2, we only need to retrieve the real part of permittivity. Simulated medium dielectric constant ε_r is assumed to be 10.354–2.4614i, and radar frequency, incident angle are 450 MHz and 25° respectively. Figures 2(a) and (b) show the inversion results. From Figure 2(a) we can see that when the imaginary part is given, relative error of reconstructed real part hardly changes with the variation of initial estimated real part. And when real part is decided, variation of imaginary part influences inversion results, as Figure 2(b) shows. However, this affection is not strong.

When the forward model is for layered rough surface, there are three variables are needed to be decided [6]. They are dielectric constants of two mediums, and distance between layers of different mediums. The cost function is as follows.

$$L = \sum_{i=1}^{n}\left\{\frac{[x(i) - x_{ap}(i)]}{x_{ap}}\right\}^2 + \sum_{i=1}^{m}\left\{\frac{[f(i) - d(i)]}{d(i)}\right\}^2 \tag{4}$$

Two simulated medium dielectric constants ε, ε_1 and distance d are assumed to be 10.354–2.4614i, 12.237–2.7928i and 0.5 m respectively. Frequencies are 450 MHz and 300 MHz, and the

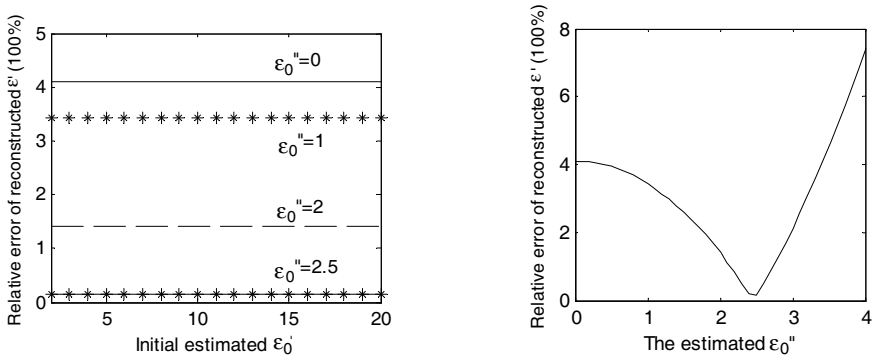


Figure 2: (a) Reconstructed real part to different initial estimated imaginary part, (b) Reconstructed real part to different initial estimated real part.

incident angle is 25°. Figures 3(a), (b), and (c) show the inversion results. When initial assumption of $\varepsilon_1^{'}$ and d are 12 and 0.48 m, relative error of reconstructed real part of ε is changing with initial estimated ε' in Figure 3(a). When initial assumption of ε' and d are 10 and 0.48 m, relative error of reconstructed real part of ε_1 is changing with initial estimated $\varepsilon_1^{'}$ in Figure 3(b). When initial assumption of ε' and $\varepsilon_1^{'}$ are 10 and 12, relative error of reconstructed d is changing with initial estimated d in Figure 3(c). However, the inversion results show that the relative errors are related to the initial assumption. And the forward model is more sensitive to ε' than $\varepsilon_1^{'}$ and d. So a more rational forward model should be established.

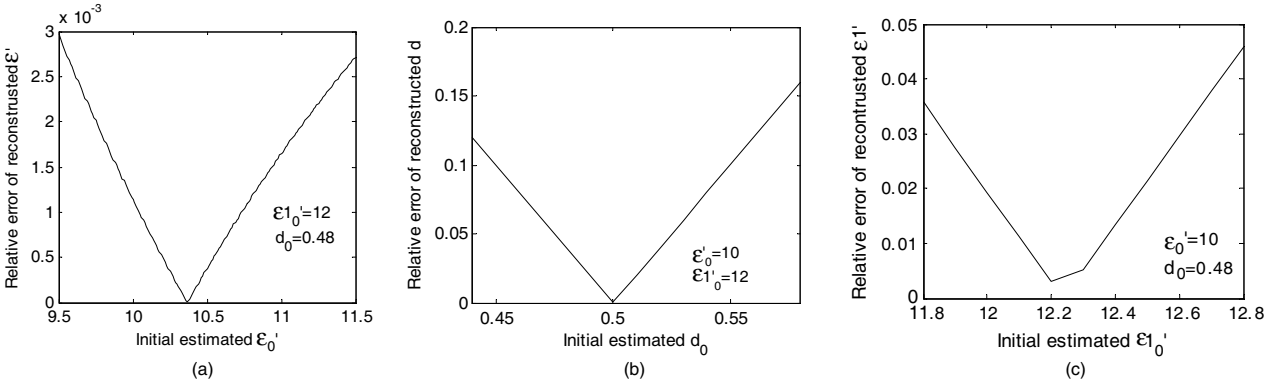


Figure 3: Reconstructed results for two-layered rough surface with different initial estimation.

5. SUMMARY

Based on dual-frequency multi-polarization radar system, volumetric moisture for surface and surface of soil can be retrieved. This is very important to monitor the variance of soil, and estimate the absolute moisture value. A conclusion that only the real part of dielectric constant is needed to inverse is obtained. In order to inverse permittivities, Nelder-Mead simplex optimization is used. This kind of method requires relatively small amount of calculating. But when it comes to tri-variable, the results are more relaying on initial assumption, which requires improving the algorithm. At the same time, better forward scattering model for layered rough surface should be established, for the current model cannot reflect actual physical process perfectly.

REFERENCES

1. Zhang, X., W. Zhang, X. Lu, G. Kang, and G. Fang, "Dual-frequency synthetic aperture radar for deep soil moisture estimation," *Journal of Electronics (accepted).*
2. Dobson, M. C., F. T. Ulaby, M. Hallikainen, and M. El-rayes,"Microwave dielectric behavior of wet soil-part II: four-component dielectric mixing models," *IEEE. Trans. Geosci. Remote Sensing*, GE-23, 35–26, 1982.
3. Peplinski, N. R., F. T. Ulaby, and M. C. Dobson, "Dielectric properties of soils in the 0.3–1.3-GHz range," *IEEE. Trans. Geosci. Remote Sensing*, Vol. 33, No. 3, 803–807, 1995.

4. Cosfriff, R. L., W. H. Peake, and R. C. Taylor, *Terrain Handbook II*, Ant. Lab. Eng. Expt. Station, Ohio State Univ., Columbus, Ohio., 1959.
5. Fuks, I. M., "Radar contrast polarization dependence on subsurface sensing," *IEEE Proc. IGARSS'98*, WA, Vol. 3, 1455–1459.
6. Tabatabateenejad, A. and M. Moghaddam, "Inversion of subsurface properties of a layered medium with rough boundaries," *Proc. X X VIII th General Assembly of International Union of Radio Science*, New Delhi, India, October, 2005.

Soil Moisture Algorithm Validation with Ground Based Networks

T. J. Jackson, **M. H. Cosh**, **R. Bindlish**, **J. Du**, **and X. Zhan**
U.S. Department of Agriculture, Agricultural Research Service, USA

Abstract— Soil moisture products from satellite sensors have to be validated because the retrieval algorithms utilize formulations, parameters and ancillary data that have not been thoroughly developed and verified. In addition, due to the complexity in scaling points to large footprints it will always be impossible to translate first principle physics/electromagnetics to these levels. Validation programs for the Advanced Microwave Scanning Radiometer (AMSR) instruments on the NASA EOS Aqua and the JAXA ADEOS-II platforms are currently ongoing. In order to address the validation issues noted, as part of the AMSR validation activity, several networks of *insitu* soil moisture sensors were established. There are now close to five years of observations available. The soil moisture algorithms used by NASA and JAXA for their respective standard products are being compared to the network observations, along with an alternative algorithm. Preliminary results indicate that the algorithms can perform within acceptable error bounds. The issues addressed here are common to both current and future satellite missions.

1. INTRODUCTION

Validation of satellite-based soil moisture algorithms and products is particularly challenging due to the disparity of scales. The key issue is spatial scale; conventional measurements of soil moisture are made at a point whereas satellite sensors provide an integrated area/volume value over a large area. Spatial variations that must be considered occur at a number of scales including the point scale, over geographic units defined by land cover, soils, and topography, and as the result of rainfall events and climate. Land remote sensing has focused on the meter to kilometer scale. For microwave remote sensing we must consider the kilometer to 40 km scale, which presents new challenges. This spatial scale issue is common to current satellite instruments (Special Sensor Microwave Imager-SSMI, Advanced Microwave Scanning Radiometer-AMSR) as well as future missions such as the Soil Moisture Ocean Salinity (SMOS) mission (Kerr et al. 2002) and the Global Climate Observing Mission (GCOM).

Soil moisture products from satellite sensors have to be validated because the retrieval algorithms utilize formulations, parameters and ancillary data that have not been thoroughly developed and verified. In addition, due to the complexity in scaling points to large footprints it will always be impossible to translate first principle physics/electromagnetics to these levels.

After acknowledging that validation is a necessity we are faced with how to approach this difficult problem. The spatial and temporal variability of soil moisture, the large footprint size, and the significance of other factors (particularly vegetation) pose problems.

Soil moisture validation as discussed here focuses on providing data for the evaluation of surface soil moisture products. The criterion for validation is the comparison of conventionally estimated surface soil moisture and the estimates based on the selected soil moisture retrieval algorithm. Ground based networks, large scale field experiments, and model products can all play a role. Here we focus on ground based networks.

Historically, validation of a soil moisture product was not a part of passive microwave remote sensing satellite missions until the AMSR instruments were launched. Validation programs for the AMSR instruments on the NASA EOS Aqua and the JAXA ADEOS-II platforms are currently ongoing and include many of the elements that are necessary. There is much to be learned and gained by using these experiences in future soil moisture missions.

2. U.S. DEPARTMENT OF AGRICULTURE (USDA) WATERSHED NETWORKS

Ground based soil moisture sampling networks have been in existence for many years; however, these were designed for other applications and did not provide data compatible with satellite validation. The density of these networks was typically sparse, the temporal frequency was poor, and the depth of measurement was too deep. At the outset of the AMSR validation program we initiated a project to correct some of these problems. Networks of dedicated soil moisture observing sites were established at USDA Agricultural Research Service (ARS) watersheds.

To implement these networks, additional surface soil moisture and temperature sensors (0–5 cm depth) were installed at and around existing instrument locations in four USDA ARS research

watersheds. For consistency, the same soil moisture/temperature instrument was used at all sites and watersheds. By utilizing the existing infrastructure, primarily rain gage networks, it was possible to establish these networks at a minimal expense.

The general site conditions are summarized in Table 1. Each is located in a different climate region. These networks provide estimates of the average soil moisture over watersheds and surrounding areas that approximate the size of the passive microwave satellite footprint. This is performed on a continuous basis (reporting every 15–30 minutes), partially in real time. A public database of the watershed data for all sites is being developed that will be made available via the Internet.

Table 1: USDA ARS watershed networks.

Watershed	Type	No. Sensors	Size (km^2)
Walnut Gulch, Arizona	Semi-Arid Range	21	148
Reynolds Creek, Idaho	Mountainous Range and Forest	16	238
Little River, Georgia	Humid Agriculture and Forest	19	334
Little Washita, Oklahoma	Sub-humid Grazingland	19	610

For each network an attempt has been made to verify the measurements provided by the individual site sensors though field and laboratory experiments using gravimetric sampling [1, 2]. In addition, the scaling of the network to the satellite scale has been established for several sites [3]. Figure 1 shows a comparison between the Walnut Gulch watershed network average soil moisture (19 sensor sites) and the average of 66 sites during a field experiment in 2004. This result demonstrates that the network can be used reliably to represent the large scale (footprint) average soil moisture for validation.

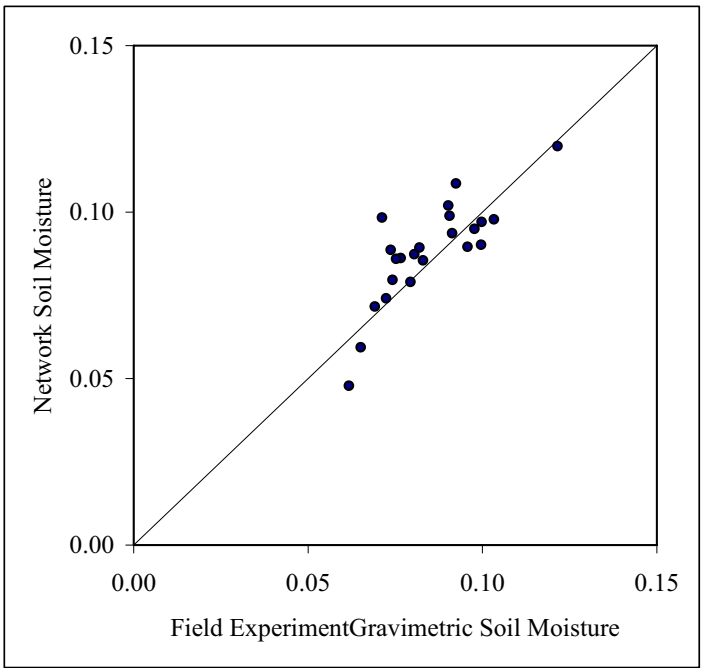


Figure 1: Comparison of the Walnut Gulch, Arizona watershed network average soil moisture and extensive field experiments conducted in 2004.

3. SATELLITE PRODUCT COMPARISONS

We have begun the process of comparing the ground based watershed average soil moisture provided by the networks to the soil moisture algorithm products. This has taken quite some time because we were waiting to have a significant length of record (to include seasonal and annual cycles), final approval of the AMSR brightness temperature products, and revisions of the soil moisture algo-

rithms. Both NASA and JAXA are moving to produce updated datasets, which should be complete within a few months. Therefore, the results described here should be considered preliminary.

Comparisons include the soil moisture product provided by NASA [6], a retrieval using an algorithm similar to that used by JAXA, and another algorithm described in Jackson [4]. Each algorithm has a similar basis but utilizes a somewhat different approach.

We have been able to conduct some preliminary analyses using data from the Walnut Gulch, Arizona (WG) watershed. Figure 2 shows the network average versus the retrieval from June 18, 2002–Dec. 31, 2005. This is a rather long period of record and some clear patterns of algorithm performance are apparent in the plot. Features of note include the limited range of response in both the NASA and JAXA products as compared to the network observations and a positive bias in the NASA estimates. We do caution that these are preliminary results that need to be reconsidered once the final agency products are available and all watershed sites are available.

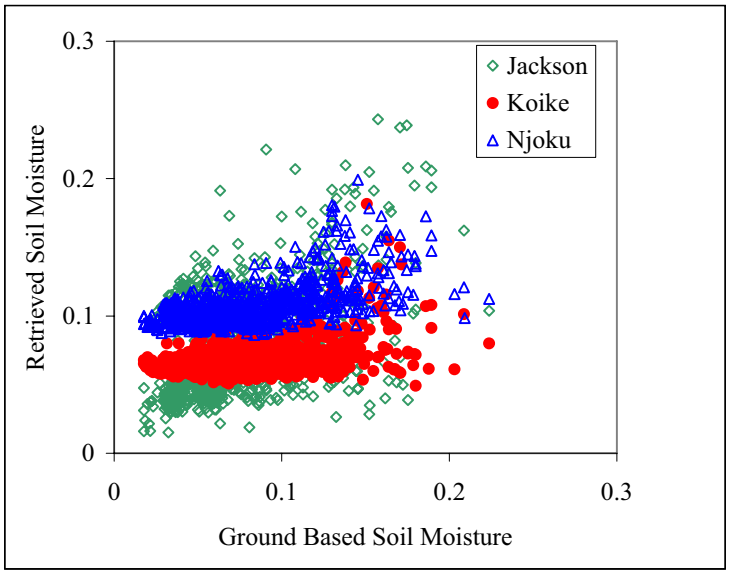


Figure 2: Preliminary validation of soil moisture algorithms using the Walnut Gulch watershed network (2002–2005).

4. SUMMARY

Validation of satellite based soil moisture products is necessary to insure the quality of the information. Retrieval algorithms have inherent limitations resulting from simplifications required for implementations. Soil moisture is particularly to validate difficult due to the mismatch in observing scales of conventional and satellite observing systems. This issue of spatial scale is common to both current and future satellite missions. An integrated approach using *insitu* networks, field campaigns and comparison to other satellite products is necessary but must be built from ground based observations. Preliminary results presented here indicate that retrievals must be carefully interpreted and analyzed before being used in any application.

ACKNOWLEDGMENT

This work is supported by the NASA Aqua Validation Program and JAXA ADEOS-II Program.

REFERENCES

1. Cosh, M. H., T. J. Jackson, R. Bindlish, J. Famiglietti, and D. Ryu, "A comparison of an impedance probe for estimation of surface soil water content over large region," *Journal of Hydrology*, Vol. 311, 49–58, 2005.
2. Cosh, M. H., T. J. Jackson, P. J. Starks, and G. Heathman, "Temporal stability of surface soil moisture in the Little Washita River Watershed and its applications in satellite soil moisture product validation," *Journal of Hydrology*, Vol. 323, 168–177, 2006.
3. Cosh, M. H., T. J. Jackson, S. Moran, and R. Bindlish, "Temporal persistence and stability of surface soil moisture in a semi-arid watershed," *Remote Sensing of Environment*, in press.

4. Jackson, T. J., "Measuring surface soil moisture using passive microwave remote sensing," *Hydrological Processes*, Vol. 7, 139–152, 1993.
5. Kerr, Y., P. Waldteufel, J. P. Wigneron, J. Font, and M. Berger, "Soil moisture retrieval from space: The Soil Moisture Ocean Salinity (SMOS) mission," *IEEE Trans. Geosci. Remote Sens.*, Vol. 39, 1729–1735, 2001.
6. Njoku, E. G., T. J. Jackson, V. Lakshmi, T. K. Chan, and S. V. Nghiem, "Soil moisture retrieval from AMSR-E," *IEEE Trans. Geosci. Remote Sens.*, Vol. 41, 215–229, 2003.

Validation of Diurnal Cycle and Intra-seasonal Variability of TRMM Satellite Rainfall

Yimin Ji

College of Sciences, George Mason University, Fairfax, VA 22030-4444, USA

Abstract— The version 6 TRMM rainfall products showed significant improvements on satellite rain retrievals as compared to rain gauge observations. The differences between TRMM TMI rain products and gauge observations are less than 10% over Florida, Texas, Kwajalein, and Darwin ground validation sites for the 8-year period from January 1998 to December 2005. The rain differences between Version 6 TMI and PR data are also within 10% over these GV sites except the Darwin area. The diurnal cycle from 8-years TRMM TMI and PR rain data showed peaks in the early morning over ocean and in the early afternoon over land. These results are accurate as compared to the gauge statistics over TRMM ground validation sites and other ground observations. TRMM TMI rain products indicate that variances of diurnal cycle are less than 5% of intra-seasonal variances over tropical oceans. Parallel spectrum analyses on TMI and Global Precipitation Climatology Project (GPCP) rain products indicate a near-perfect match of intra-seasonal spectral distributions over tropical oceans.

1. INTRODUCTION

The Tropical Rainfall Measuring Mission (TRMM) has improved our understanding of rainfall properties and variations in the tropics since its launch in December 1997 [1]. The most innovative instrument on TRMM is the first space borne Precipitation Radar (PR). The measurements of PR have provided invaluable information on the intensity and distribution of the rain, as well as the 3-D structure of storm. The TRMM Microwave Imager (TMI) is a passive microwave sensor based on the design of the highly successful Special Sensor Microwave/Imager (SSM/I) which has been flying on Defense Meteorological Satellites since 1987. However, the higher resolution of TMI, as well as the additional 10.7 GHz frequency made TMI a better instrument than its predecessors. The coincident PR and TMI measurements, and the coincident TRMM satellite and ground validation measurements, have supplied valuable information to improve our rainfall retrieval algorithms as well as our understanding of rainfall structures. Both PR and TMI rain retrieval algorithms have been advanced significantly since the launch of TRMM through multiple TRMM data reprocessing. The TRMM data have been reprocessed six times so far. The differences between gauge and TRMM satellite rain products, and the differences between PR and TMI products had been improved during each of these reprocessings, reflecting a improved understanding of the mechanisms and structures of tropical rainfall. TRMM has provided not only single sensor rain products, but also high quality multi-satellite merged rain data sets.

The local overpass time of TRMM satellite drifts each day completing a daily cycle in 46 days. This unique design provides a great opportunity of studying diurnal cycle for various TRMM science parameters, such as rainfall, surface temperature, and fire products. However, due to the poor sampling of single satellite data, aliasing of a strong diurnal cycle may cause a false intra-seasonal oscillation and seasonal and intra-seasonal variability may also affect diurnal cycle studies. While an ultimate way to eliminate such aliasing is to increase temporal and spatial sampling resolutions, as proposed by Global Precipitation Measurement (GPM) mission, an improved description of the aliasing effect is necessary and can be obtained from the TRMM observations. This paper uses version 6 TRMM TMI, PR, and merged rainfall data in the past eight years to compare the magnitudes of diurnal cycle and intra-seasonal variability over both ocean and land. The TRMM ground validation (GV) data from a number of GV sites in the same period are used to validate the results.

2. DIURNAL CYCLE OF TRMM RAINFALL

The spatial distribution of global rainfall for different local time from 8-year TRMM TMI observations indicate weak diurnal cycle over ocean and a strong diurnal cycle over land as well as warm pool (Fig. 1, left columns). The peaks of rainfall occur during early morning over oceans (19–24 UTC over warm pool, 07–12 UTC over Atlantic Ocean) and during early afternoon over continental land (UTC 13–18 in Africa, UTC 19–24 in North and South Americas, UTC 06–12 in South Asia).

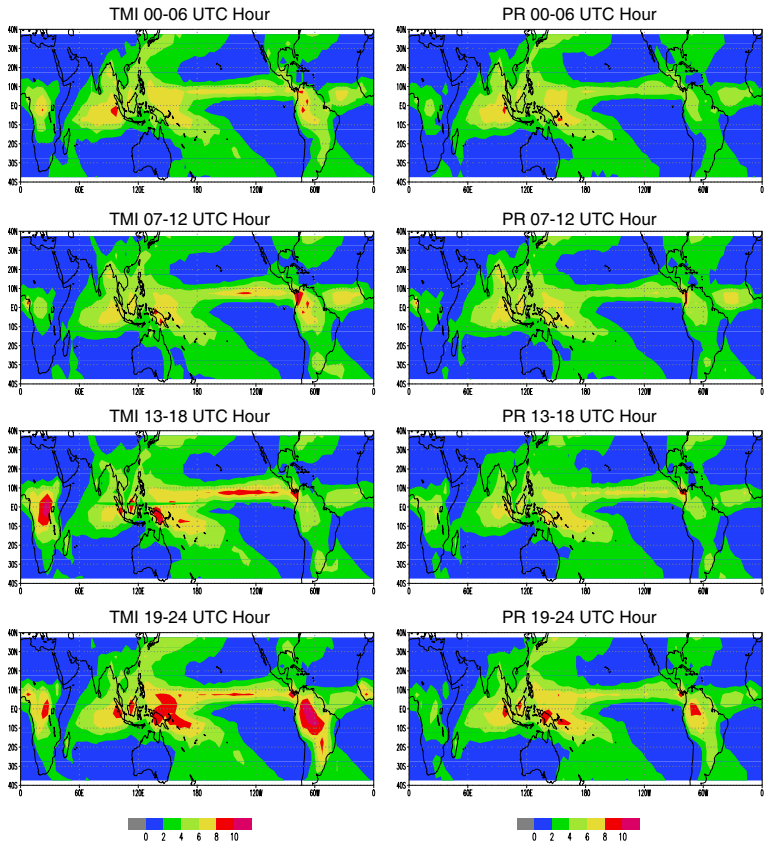


Figure 1: Spatial distributions of global rainfall for different local time from TRMM TMI product (left panels) and PR product (right panels).

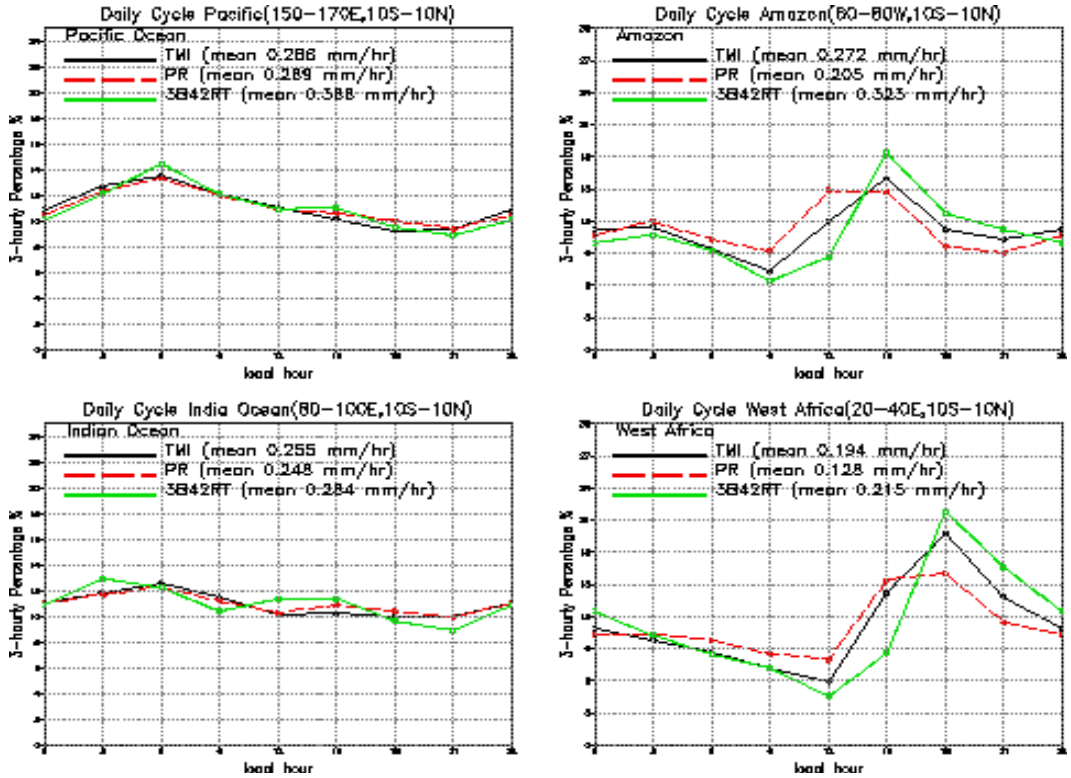


Figure 2: Diurnal cycles of TRMM TMI (solid line), TRMM PR (dashed line), and TRMM Merged products (dotted line) over Pacific (upper left), Indian ocean (lower left), Amazon area (upper right), and West Africa (lower right).

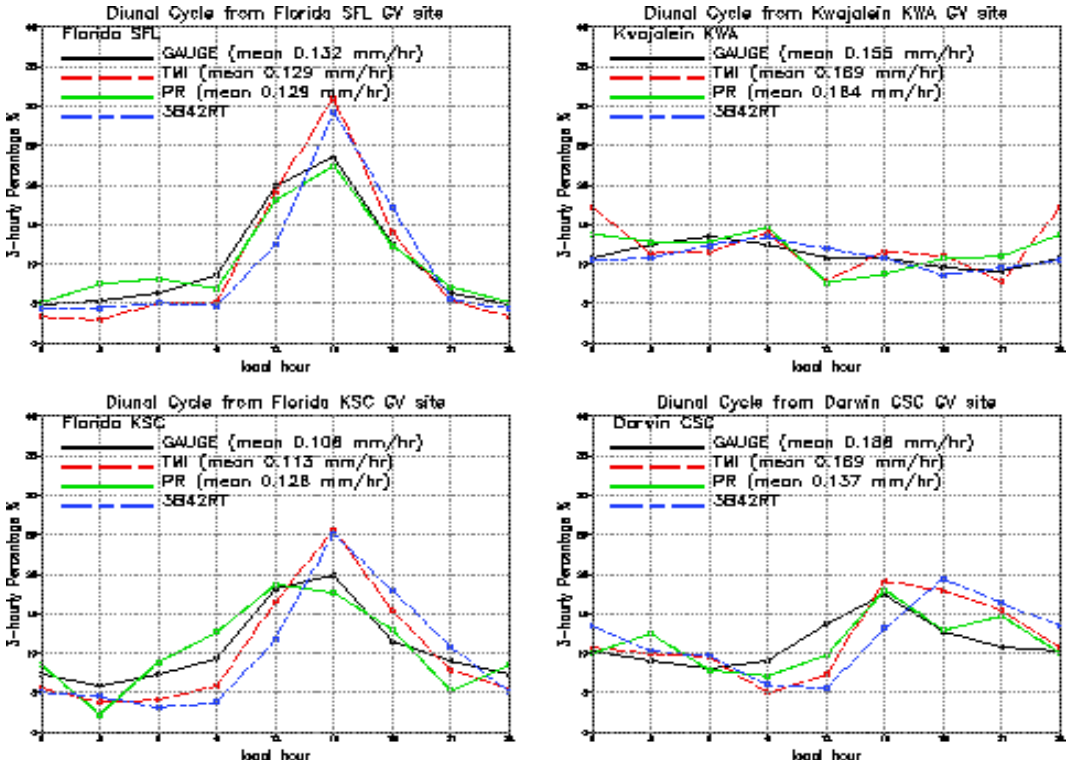


Figure 3: Diurnal cycles of TRMM gauge (solid line), TRMM TMI (dashed line), TRMM PR (dotted line), and TRMM Merged products (broken line) over SFL (upper left), KSC (lower left), KWA (upper right), and CSC (lower right) GV sites.

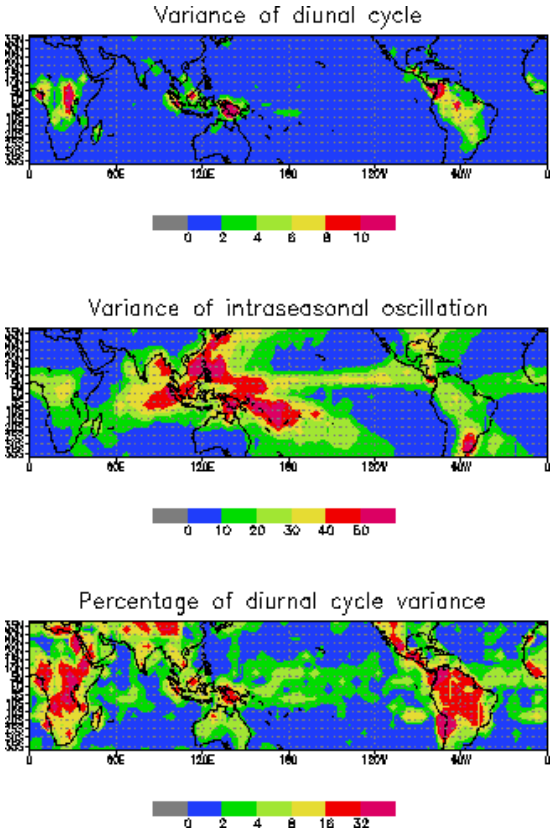


Figure 4: Variances of diurnal cycle (upper), variances of intraseasonal variability (middle), and the percentage of variances of diurnal cycle as compared to that of intraseasonal oscillation (lower).

The results of TRMM PR products (Fig. 1, right columns) are similar. However, PR data show an overall lighter rain.

Figure 2 compares the diurnal cycles of TRMM TMI, TRMM PR, and TRMM Merged products over four typical regions. The TMI and PR data cover a period from January 1998 to December 2005. The merged data start from January 2002. In Pacific and Indian oceans (left columns), diurnal cycles are quite flat. The rainfall has a peak around 6 am local time, which is similar to the finding of Cronin and McPhaden [2]. The results of the three products are very similar. Over land (right columns), the rainfall peaks around 3 pm local time in Amazon region and 6 pm local time in Western Africa region. The early afternoon rain peak in Amazon region was found in other studies (e. g., [3]). Over selected ocean areas the TMI and PR show less than 10% differences of mean rainfall.

Inter-comparisons of mean rainfall and diurnal cycles between TRMM satellite and ground validation (GV) gauge products over four different GV sites are shown in Fig. 3.

Over Florida SFL GV site (Fig. 3 upper left), the 8-year means of TMI, and PR estimates are about 3% lower than the gauge mean rain. These values are also similar to the local climate mean. In the smaller GV area (KSC GV site, lower left), the difference of mean rain is less than 10% between gauge and TMI estimates, but about 20% between gauge and PR estimates. The diurnal cycles from all the three satellite products are very similar to the gauge observations with rain peaks around 3 pm local time. The amplitudes of oscillation between PR and gauge data are very close, while TMI and merged data showed larger amplitudes of diurnal oscillation. Over Kwajalein ocean site (Fig. 3 upper right), the 8-year means of TMI, and PR estimates are about 10% higher than the gauge mean. All products show weak diurnal cycles with peaks around 9 am local time. However, the diurnal cycles of satellite products don't march the gauge data very well due to the small size of gauge area (about 0.5°). For a larger area (Fig. 2, upper left) which covers the KWA site, the diurnal cycles from all satellite products are very similar gauge results. In Darwin CSC (lower right), TMI mean rain is about 10% lower than the GV rain while the PR mean rain is about 30% lower than the GV rain. The peak of diurnal cycle is around 3 pm.

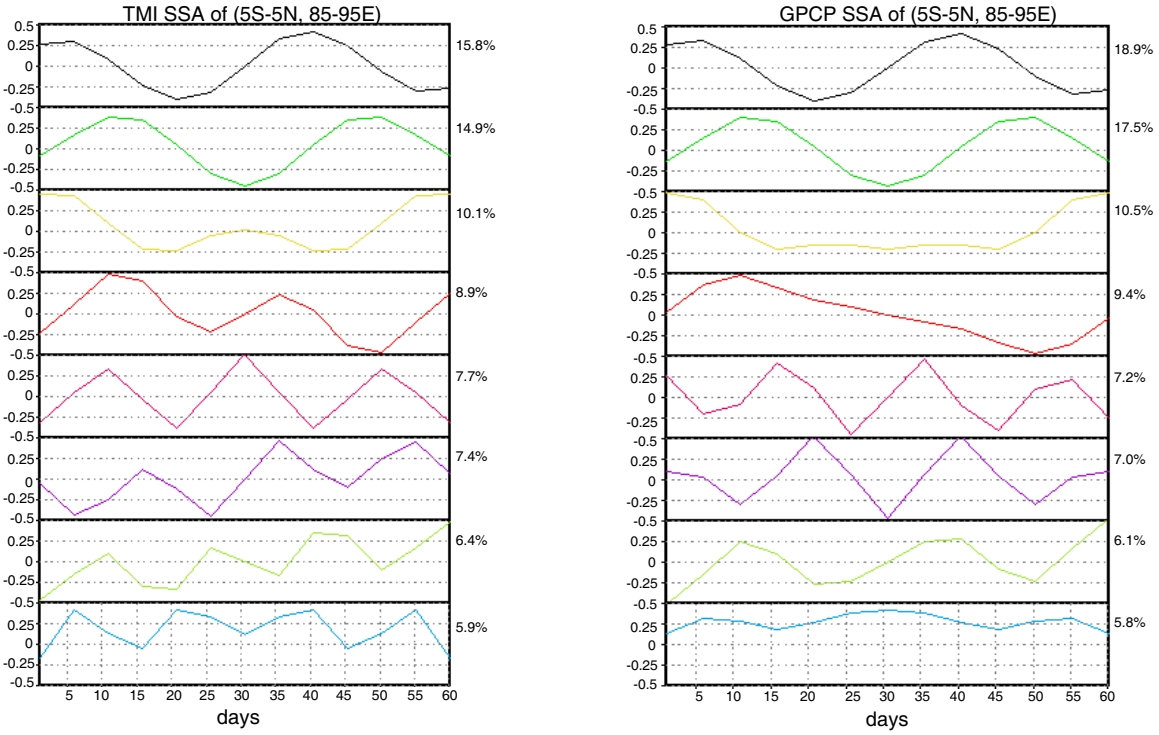


Figure 5: First eight spectra of intraseasonal oscillations from TMI rainfall in Indian Ocean (left), and from GPCP products in Indian Ocean (right).

3. INTRA-SEASONAL VARIABILITY

Aliasing of diurnal cycle may cause a false intraseasonal oscillation if sampling methods are inappropriate. Such effects may be removed by increasing grid size or time interval. The aliasing effect of diurnal cycle on intraseasonal variability is small over ocean since the variance of diurnal cycles is less than 5% of the intraseasonal variance (Fig. 4, upper panel). However, the results may differ over land since the variance of diurnal cycle is comparable to that of intraseasonal variability over land (Fig. 4 middle panel).

The intraseasonal spectrum from TRMM observation (Fig. 5 left) showed validity as compared to the geostationary satellite based observation (Fig. 5 right) over equatorial ocean region. The dominant spectrum is the 30-60 day Madden-Julian Oscillation.

4. SUMMARY

The differences of mean rainfall between gauge and TMI are typically within 10%. The diurnal cycles derived from satellite products are generally agree well among themselves and with the gauge observations. The results show early morning rain peak over oceans and afternoon rain peaks over lands.

REFERENCES

1. Kummerow, C., J. Simpson, O. Thiele, W. Barnes, A. T. C. Chang, E. Stocker, R. F. Adler, A. Hou, R. Kakar, F. Wentz, P. Ashcroft, T. Kozu, Y. Hong, K. Okamoto, T. Iguchi, H. Kuroiwa, E. Im, Z. Haddad, G. Huffman, B. Ferrier, W. S. Olson, E. Zipser, E. A. Smith, T. T. Wilheit, G. North, T. Krishnamurti, and K. Nakamura, "The status of the tropical rainfall measuring mission (TRMM) after two years in orbit," *Journal of Applied Meteorology*, Vol. 39, No. 12, 1965–1982, 2000.
2. Cronin, M. F. and M. J. McPhaden, "Diurnal cycle of rainfall and surface salinity in the western Pacific warm pool," *Geophys. Res. Let.*, Vol. 26, 3465–3468, 1999.
3. Machado, L. A. T., H. Laurent, and A. A. Lima, "Diurnal march of the convection observed during TRMM-WETAMCC/LBA," *J. Geophys. Res.*, Vol. 107, 8064, 2002.

The Optimal Design of the LVDS Bus with High EMS

K. Radkovsky[1], **M. Kaska**[1], **Z. Motycka**[1]
P. Drexler[2], **T. Jirku**[2], **Z. Szabo**[2], **E. Kadlecova**[2], **and P. Fiala**[2]

[1]TES, Prazska 597, 674 01 Trebic Czech Republic
[2]Department of Theoretical and Experimental Electrical Engineering
Kolejni 2906/4, 61200 Brno, Czech Republic

Abstract— The article deals with results which were obtained during the process of designing the LVDS (low voltage differential signal) bus. This bus is very resistant to interference from the environment and to spurious emission of electromagnetic waves. According to the concrete requirements we prepared a design of the LVDS by means of numerical methods. The design was realized and its parameters verified experimentally. The basic parameter of bus is its impedance $Z = 100\,\Omega$, up to frequency $f_1 = 100\,\mathrm{MHz}$ of the first harmonic of rectangle signal.

1. INTRODUCTION

The LVDS bus is used in many applications, especially in those requiring fast and reliable data transmission. The advantages of this bus are commonly known. In special cases, we need to use the LVDS transmission technology with great immunity to electromagnetic interference (EMI). The main problem of the design of this type of bus is how to find the optimal impedance matching with respect to production technology. Fig. 1 shows the obvious LVDS bus configuration which interconnects a set of functional blocks, for example plug-in cards.

The high electromagnetic susceptibility (EMS) of the LVDS bus is required by relevant rules, valid for the system application in nuclear power plants. With respect to electromagnetic compatibility (EMS) regulations the specific design of the high-EMS LVDS bus was implemented with the help of numerical analysis, as shown in Fig. 2. The total length of the bus is $l = 265\,\mathrm{mm}$. The design exploits a multi-layer configuration, including the ground layer ensuring the shielding.

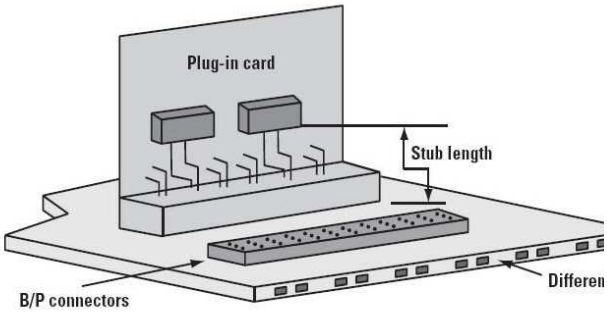


Figure 1: An example of the classical LVDS bus configuration.

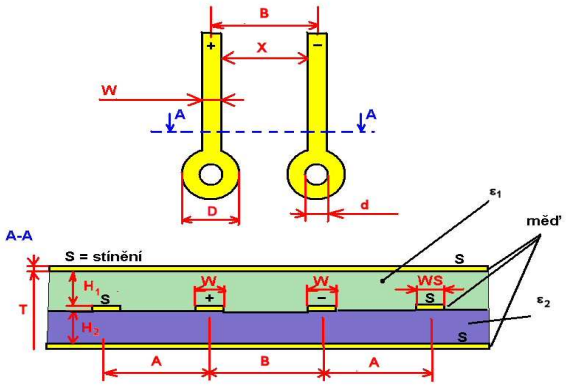


Figure 2: Details of the designed LVDS bus.

The bus design exploits the multi-layer PCB technology. Due to the used multi-layer technology, some restrictions have to be considered, for example the thickness of microstrip lines or dielectric layers. Fig. 2 shows basic dimensions of the bus design — the microstrip line spacing (A, B, X), the layer spacing (H_1, H_2), the thickness of the microstrip and the ground layer (T), the width of the microstrip line (W) and the width of the reference microstrip line (W_s). The thickness of the microstrip and the ground layer is $H = 35\,\mu\mathrm{m}$. The electrode S in Fig. 2 is the reference electrode placed on both sides of the PCB. The symbols + and − stand for the pair of the differential transmission line. The basic scheme of the setup of the LVDS system is in Fig. 3. The setup includes duplex data transmission. The right part of Fig. 3 shows the interconnection of the transmitter-receiver pairs via the bus. Each pair of the microstrip lines is terminated with the impedance $Z = 100 + j0\,\Omega$. Fig. 4 shows the example of the signal spectrum of the transmitter produced by TES.

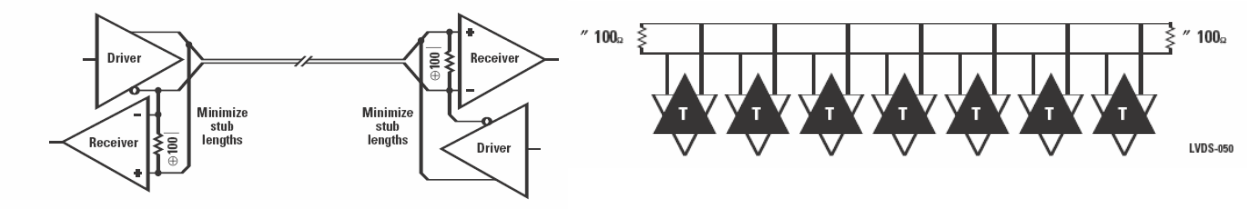


Figure 3: The interconnection of the LVDS transmitter-receiver system.

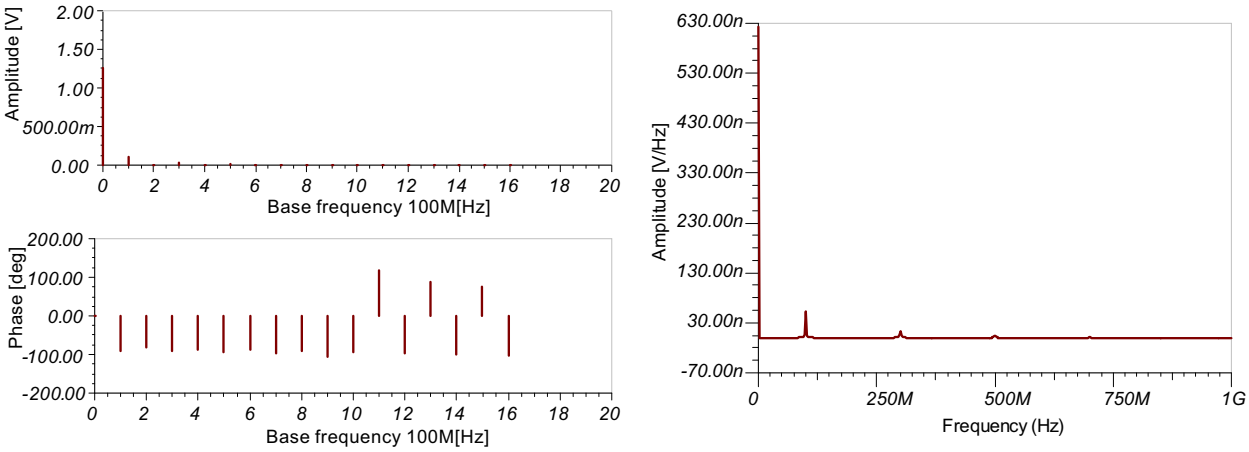


Figure 4: Spectrum of the TES LVDS transmitter.

Measurement of the LVDS transmitter was performed. The model of the transmitter was made on the basis of obtained results. The amplitudes of all the following significant spectrum frequencies $f_1 = 100\,\mathrm{MHz}$, $f_2 = 130\,\mathrm{MHz}$, $f_3 = 270\,\mathrm{MHz}$, $f_4 = 500\,\mathrm{MHz}$, $f_5 = 680\,\mathrm{MHz}$ are $U_1 = 1.300\,\mathrm{V}$, $U_2 = 0.310\,\mathrm{V}$, $U_3 = 0.031\,\mathrm{V}$, $U_4 = 0.002\,\mathrm{V}$, $U_5 = 0.001\,\mathrm{V}$. The finite element method (FEM) model of the LVDS bus was supplied by the transmitter model. The input impedance of the differential pair of transmission lines was calculated with the help of the superposition principle. The bus differential pair was terminated with the pre-defined impedance. However, it's rather problematic to talk about the impedance in the case of non-harmonic signal. The optimal design with constrained conditions was solved with target function with penalties

$$\mathbf{\Phi}^o_{V,G,X} = f^o_W + f^o_G + f^o_X \left[1 + \sum_{i=1}^{N_d} P^o_X(\chi_i) + \sum_{j=1}^{N_S} P^o_G(G_j)\right], O = 1, \ldots N_o \tag{1}$$

where f_W, f_G, f_X are Kreisselmeier- Steinhauser object, state and shape approximation functions respectively

$$f_W = \frac{1}{\zeta} \ln\left(\sum_{j=1}^{7} e_{\ln}^{(\zeta_j V_j)}\right) \quad \forall \zeta, \zeta_j \in \langle 1, 20\rangle, \tag{2}$$

$$f_G = \frac{1}{\eta} \ln\left(\sum_{\ell=1}^{N_S} e_{\ln}^{(\eta_\ell G_\ell)}\right) \quad \forall \eta, \eta_\ell \in \langle 1, 20\rangle, \tag{3}$$

$$f_X = \frac{1}{\xi} \ln\left(\sum_{k=1}^{N_d} e_{\ln}^{(\xi_k X_k)}\right) \quad \forall \xi, \xi_k \in \langle 1, 20\rangle. \tag{4}$$

where V, G, X are the object function, the state approximated function, the design approximated function respectively. This normed approximation function has very good properties and may be used in a high frequency FEM model and for global optimization. The design process has fast convergence to optimal solution. This method converges faster than the standard gradient optimization method.

2. MATHEMATICAL MODEL OF THE LVDS BUS

It is possible to carry out an analysis of an MG model as a numerical solution with the help of the finite element method (FEM). The electromagnetic part of the model is based on the solution of full Maxwell's equations.

$$\nabla \times \boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial t}, \nabla \times \boldsymbol{H} = \sigma \boldsymbol{E} + \frac{\partial \boldsymbol{D}}{\partial t} + \boldsymbol{J}, \nabla \cdot \boldsymbol{D} = \rho, \nabla \cdot \boldsymbol{B} = 0, \quad \text{in region } \Omega, \tag{5}$$

where $\boldsymbol{E}$ and $\boldsymbol{H}$ are the electrical field intensity vector and the magnetic field intensity vector, $\boldsymbol{D}$ a $\boldsymbol{B}$ are the electrical field density vector and the magnetic flux density vector, $\boldsymbol{J}_s$ is the current density vector of the sources, ρ is the density of free electrical charge and γ is the conductivity of material, Ω is thedefinition area of the model. The relationships between the electric and the magnetic field intensities and densities are given by material relationships

$$\boldsymbol{D} = \varepsilon \boldsymbol{E}, \quad \boldsymbol{B} = \mu \boldsymbol{H}. \tag{6}$$

The permittivity ε, the permeability μ and the conductivity γ in HFM are generally tensors with main axes in the direction of the Cartesian coordinates x, y, z. When all the field vectors perform rotation with the same angular frequency ω, it is possible to rewrite the first Maxwell equations as

$$\nabla \times \underline{\boldsymbol{E}} = -j\omega\mu \underline{\boldsymbol{H}}, \nabla \times \underline{\boldsymbol{H}} = (\sigma + j\omega\varepsilon)\underline{\boldsymbol{E}} + \underline{\boldsymbol{J}}_s, \quad \text{in region} \quad \Omega, \tag{7}$$

where $\underline{\boldsymbol{E}}$, $\underline{\boldsymbol{H}}$, $\underline{\boldsymbol{J}}_s$ are field complex vectors. Taking into account the boundary conditions given in (1) and after rearranging (7) we get

$$(j\omega)^2 \varepsilon \underline{\boldsymbol{E}} + \sigma \underline{\boldsymbol{E}} + \nabla \times \mu^{-1} \nabla \times \underline{\boldsymbol{E}} = -j\omega \underline{\boldsymbol{J}}_s. \tag{8}$$

We apply Galerkin's method with vector approximation functions $\boldsymbol{W_i}$. We use the vector form of the Green theorem on the double rotation element [3]. After discretisation we get the expression

$$-k_0[M]\{\boldsymbol{E}\} + jk_0[C]\{\boldsymbol{E}\} + [K]\{\boldsymbol{E}\} + \{\boldsymbol{F}\}, \tag{9}$$

where $\{\boldsymbol{E}\}$ is the column matrix of the electric intensity complex vectors. The matrixes $[K]$, $[C]$ and $[M]$ are in the form that is given in manual [4] and vector $\{\boldsymbol{F}\}$ is evaluated from expression

$$\{\boldsymbol{F}\} = -jk_0 Z_0 \int_{\Omega} [\boldsymbol{W_i}]\{\boldsymbol{J_s}\} d\Omega + jk_0 Z_0 \int_{\Gamma_0 + \Gamma_1} [\boldsymbol{W_i}]\{\boldsymbol{n} \times \boldsymbol{H}\} d\Gamma. \tag{10}$$

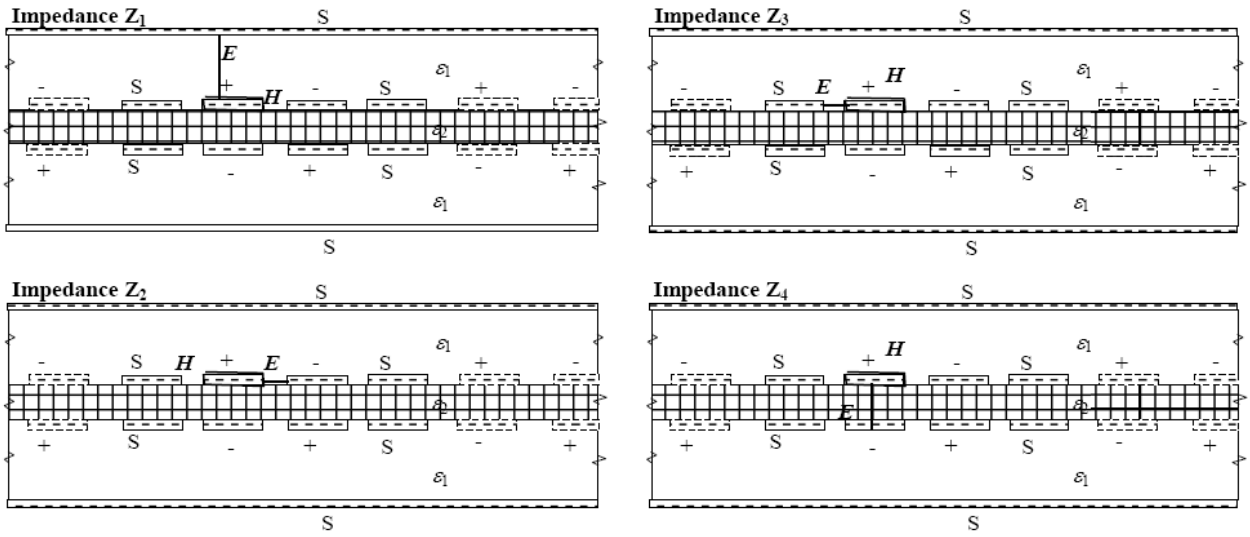


Figure 5: The analysis of LVDS bus impedance.

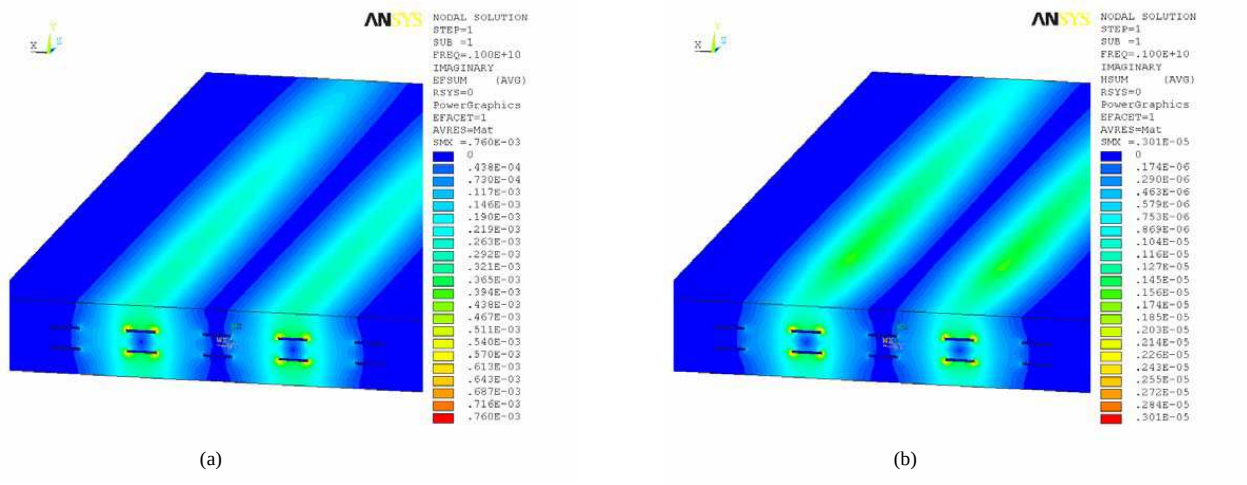


Figure 6: (a) Distribution of electric field intensity vector module E, (b) magnetic field intensity vectormodule H.

The vector approximation functions $\boldsymbol{W}$ are given in manual [4], k_0 is wave number for vacuum, Z_0 is the impedance of the free space. The set of Eq. (9) is independent of time and it gives $\underline{\boldsymbol{E}}$. For transient vector $\boldsymbol{E}$ we can write

$$\boldsymbol{E} = \mathrm{Re}\left\{\underline{\boldsymbol{E}} e^{j\omega t}\right\}. \tag{11}$$

The input impedance of the microstrip line was evaluated with the help of the IMPD program. Mutual impedances and impedances of wires referring to the common reference potential were investigated as shown in Fig. 5. In accordance with traces marked as E and H in Fig. 6, the distribution of the electric and the magnetic field intensities were analyzed. Next, the module of corresponding impedance was calculated (8).

$$U = \int_{E_l} \boldsymbol{E} \cdot d\ell, I = \int_{H_l} \boldsymbol{H} \cdot d\ell, Z = \frac{U}{I}. \tag{12}$$

For an optimal bus design, the relation (9) should be kept.

$$Z_1 = Z_2 = Z_3 = Z_4 = Z. \tag{13}$$

A geometric model utilizing the HF 120 element was built according to the mathematical model [1].

Table 1: Differential pair mutual impedances.

	LVDS1		**LVDS2**		**LVDS3**		**LVDS4**		**LVDS5**	
f [MHz]	**Z** [Ω]	φ[°]	**Z** [Ω]	φ[°]	**Z** [Ω]	φ[°]	**Z** [Ω]	φ[°]	**Z** [Ω]	φ[°]
1	100	0	100	0	100	0	100	0	100	0
10	98	−1	98	−1	98	−1	98	−1	98	−2
50	86	2	86	2	86	3	87	2	86	1
100	76	−6	88	2	86	5	74	23	78	4

3. CONCLUSION

The analysis was performed for the required frequencies and signal levels. A short APDL program was used to find the geometrical optimum of the LVDS bus dimension. Fig. 6 shows the distribution of modules of the electric field intensity vector $\boldsymbol{E}$ and the magnetic field intensity vector $\boldsymbol{H}$ of the selected bus pair. This evaluation serves to verify the designed characteristic impedance Z. A

wrong bus design causes the presence of standing waves on the transmission lines. The characteristic impedance values were experimentally verified for some frequencies. The spectrum analyzer Agilent N1996A was used in the experiment.

ACKNOWLEDGMENT

The paper was prepared within the framework of the research plan No. MSM 0021630513 of the Ministry of Education, Youth and Sports of the Czech Republic.

REFERENCES

1. ANSYS manual, Vol. 8, No. 1, USA.
2. Fiala, P., "Coupled electromagnetic model of smoothing choke," *Proceedings of the 5th AMTEE International Conference*, C05, Pilsen, Czech Republic, 2001.

A Passive Optical Location with Limited Range

P. Fiala, **T. Jirku**, **R. Kubasek**, **P. Drexler**, **and P. Konas**
Department of Theoretical and Experimental Electrical Engineering, Brno University of Technology
Kolejni 2906/4, 612 00 Brno, Czech Republic

Abstract— We know active and passive methods of a location. This article deals only with the passive location of dynamic targets (see [1–7]). The passive optics location is suitable just for the tracking of targets with mean velocity which is limited by the hardware basis. The aim of this work is to recognize plasma, particles etc. It is possible to propose such kind of evaluation methods which improve the capture probability markedly. The suggested method deals with the short-distance evaluation of targets. We suppose the application of three independent principles of how to recognize an object in a scanned picture. These principles use similar stochastic functions in order to evaluate an object location by means of simple mathematical operations. The methods are based on direct evaluation of picture sequence with the help of the histogram and frequency spectrum. We are trying to find out the probability of appearance of moving object in pictures. If the probability reaches a setting value, we will get a signal. The processing of dynamic pictures and their filtration represent a significant part of work [8–13]. Static objects, background (trees, buildings) must be filtered off before searching the objects. This filtration is also done by means of the probability function. The probability distribution of an object position is gained from a sequence of more pictures.

1. INTRODUCTION

At the DTEEE FEEC in Brno we are engaged in experimental ascertaining of water properties, elements and their moving (slow and quick) in an electric field. One of problems is how to recognize identical elements, their clusters and objects in dynamic state. It is necessary to realize this observation through automatic or semiautomatic image recognition according to the lifetime or speed of these effects. Thus, it is possible to record the trajectories of demanded objects.

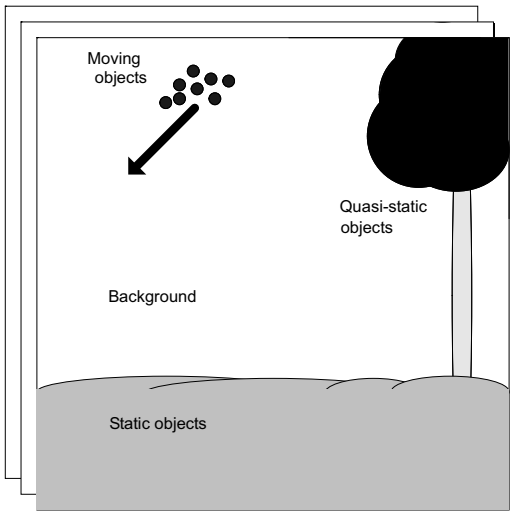


Figure 1: Example of an image recorded by a passive optical system.

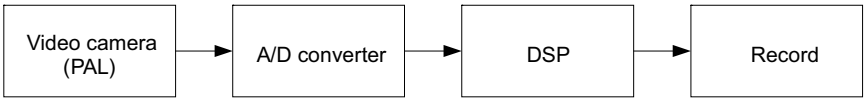


Figure 2: Block diagram of a passive optical locator with limited range.

A concept of passive optical location with limited range was designed. The basis of this concept is an optical system which can locate and deliver (record) pictures with sufficient resolution o_s and

sufficient rate o_r. For the first experiments, delivered components such as F-OS-226CA-PAL or TVP5150AM are used. The basic concept is shown in Fig. 2. A significant problem is the scanned picture processing and quick evaluation — with probability p_f — of whether the monitored object is the desired target object. More approaches are possible and they are described in sources [1–12]. The basic methods of image processing are described in source [13].

2. MATHEMATICAL MODEL

The main idea is based on three different methods of dynamic image analysis. The first one is the method of differential images, the second one is the histogram monitoring method and the last method is based on an analysis of image frequency and phase spectrum. These methods are bounded by probability functions.

The main goal of the experimental part of this work is the retrieval of an applicable probability function (Herodotus, Bernoulli, Bayes, Laplace, Jeffreys, Cox, Shannon) for the monitored target deliver recognition. The probability of a target delivering is given by:

$$p_f = p(A) \cdot p(B) \cdot p(C) \quad \text{in } \Omega, \tag{1}$$

where Ω is the monitored image area, $p(A)$ is the monitored object move probability given by the first method, $p(B)$ is the monitored object move probability given by the second method and $p(C)$ is the probability given by the third one.

The basic part of the differential images method model may be described the following way. We have $u_0,\ X\colon \Omega \to R$, with $\Omega \subset R^2$ then for the coeficients ω picewise smooth image approximation holds:

$$\omega_1 = \frac{u_{t+1} - u_t}{\Delta X_{t,t+1}},\ \omega_2 = \frac{u_{t+2} - u_{t+1}}{\Delta X_{t+2,t+1}},\ u \cong \omega_1 X_t + \omega_2 X_{t+1} \quad \text{in } \Omega, \tag{2}$$

the probability of the monitored object delivering is

$$p(A) = \left((u_{t+1} - u_t)/u_t\right) \cdot \left((u_{t+1} - u_{t+1})/u_{t+1}\right). \tag{3}$$

The method is completed with a method of segmentation [12]. The choice of an applicable number of segments choice makes the image processing time shorter. It is possible to describe the model as

$$\inf_{u,\Gamma} \left\{ F(u,\Gamma) = \int_{\Omega} |u - u_0|^2\, dx + k_1 \int_{\Omega - \Gamma} |\nabla u|^2\, dx + k_2 \int_{\Gamma} d\Gamma \right\}, \tag{4}$$

where Γ is the border of area Ω, k_1, k_2 are weight function. We obtain the approximated function solution u in the desired segment if we find the minimum.

The histogram monitoring method [14] is based on the statistics image processing method; there, similarly as in the Monte-Carlo method, the evaluation of an image is done depending on the image color structure. We can describe this method. We have $u_0,\ X\colon \Omega \to R$, with $\Omega \subset R^2$ then for weight k_3

$$p(B) = k_3 \frac{u_{t+1}(X) - u_t(X)}{u_t(X)} \quad \text{in } \Omega, \tag{5}$$

The frequency and phase spectrum method is a variant close to the histogram monitoring method. The difference of this method is in using the continuous image information spectrum and the added information is in the phase spectrum space. This information can make the whole image processing more accurate. It is possible to describe the analysis model for $u_0,\ X\colon \Omega \to R$, with $\Omega \subset R^2$ from the Fourier transform:

$$U_F(m,n) = \sum_{X_1=0}^{M-1} \sum_{X_2=0}^{N-1} u_0(X_1, X_2)\, e^{-j\frac{2\pi m X_1}{M}}\, e^{-j\frac{2\pi n X_2}{N}}, \tag{6}$$

where M, N are the information about the image, m, n are the counts of coefficients in the discrete series. Then the amplitude spectrum is given by

$$U_{AF}(m,n) = |U_F(m,n)| \tag{7}$$

and the phase spectrum

$$U_{\varphi F}(m,n) = \arctan\left(\frac{\text{imag}\,(U_F(m,n))}{\text{real}\,(U_F(m,n))}\right). \tag{8}$$

This approach is not suitable for fast image processing, because in this case each pixel is solved. A preferable image spectrum processing method uses the discrete cosine transform

$$U_c(m,n) = \sum_{X_1=0}^{M-1}\sum_{X_2=0}^{N-1} u_0\,(X_1,X_2)\cos\left(\frac{\pi(2X_1+1)m}{2M}\right)\cos\left(\frac{\pi(2X_2+1)n}{2N}\right). \tag{9}$$

The transformed image has similar properties as the amplitude spectrum of the Fourier transform for the low-frequency character of image, but it is processed in a faster way. The dependent dynamic object presence probability is with weight k_4

$$p(C) = k_4\frac{u_{t+1}(X) - u_t(X)}{u_t(X)} \quad \text{in } \Omega. \tag{10}$$

3. EXPERIMENTS

The experiment was prepared on the basis of the Fig. 2 diagram. In Fig. 3, the experimental connection of the optical locator with limited range basic parts is shown. Basic measurement and dynamic image analysis are performed on these components. Next, the device for the plasma element monitoring will be prepared, as mentioned above. The locator system will be tested on

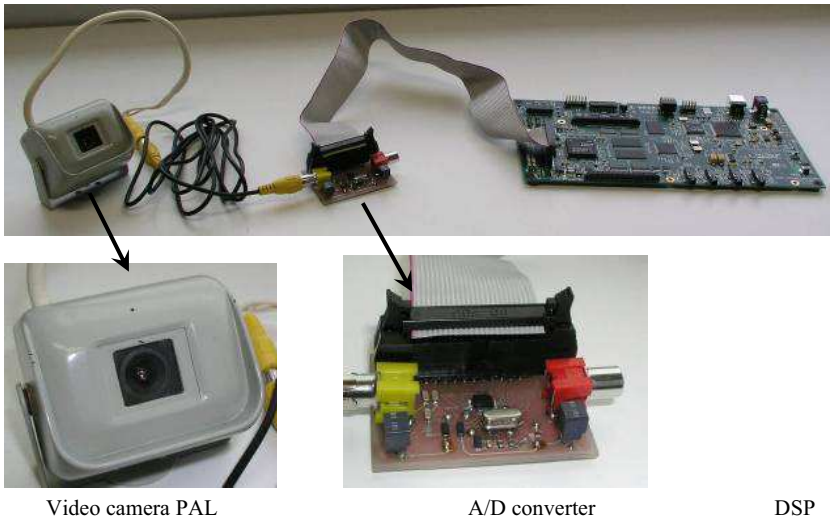


Figure 3: Basic parts of the passive optical locator with limited range.

bird moves. The real image is in Fig. 4 and the image recorded by the video-camera is in Fig. 5.

Figure 4: Example of an image scene.

Figure 5: Example of a recorded image scene, $o_s = 640 \times 480$ pixels.

ACKNOWLEDGMENT

The research described in the paper was financially supported by the FRVŠ, research plan No MSM 0021630513 ELCOM, No MSM 0021630516 and GAAV grant No. B208130603.

REFERENCES

1. Herman, S. and P. Moulin, "A particle filtering approach to FM-band passive radar tracking and automatic target recognition," *Proceedings of IEEE Aerospace Conference*, Boston, 2002.
2. Internet source: http://www.ifp.uiuc.edu/∽smherman/darpa/.
3. Ehrman, L. M. and A. D. Lanterman, "Automated target recognition using passive radar and coordinated flight models," *Proceedings of SPIE International Conference*, Orlando, Florida, 2003.
4. Lanterman, A. D. and D. C. Munson, "Deconvolution techniques for passive radar imaging," *Proceedings of SPIE International Conference*, Orlando, Florida, 2002.
5. Wu, Y. and D. C. Munson, Jr., "Multistatic passive radar imaging using the smoothed pseudo Wigner-Ville distribution," *Proceedings of IEEE Image Processing Conference*, 604–607, 2001.
6. Lind, F. D. and J. D. Sahr, "Passive radar and the low frequency array," *Proceedings of General Assembly of the Inetrnational Union of Radio Science*, Maastricht, Netherlands, 2002.
7. Neyt, X., J. Raout, M. Kubica, V. Kubica, S. Roques, M. Acheroy, and J. G. Verly, "Feasibility of STAP for passive GSM-based radar," *Proceedings of the IEEE Radar Conference*, Verona, NY, 546–551, April 2006.
8. Töreyin, B. U., A. E. Cetin, A. Aksaya, and M. B. Akhan, "Moving object detection in wavelet compressed video," *Signal Processing: Image Communication*, Vol. 20, 225–264, 2005.
9. Morimoto, C., D. DeMenthon, L. Davis, and R. Chellappa, "Detection of independently moving objects in passive video," *Proceedings of Intelligent Vehicles Symposium*, Detroit, MI, 270–275, 1995.
10. Nguyen, H. T., M. Worring, and A. Dev, "Detection of moving objects in video using a robust motion similarity measure," *IEEE Transactions on Image Processing*, Vol. 9, 137–141, 2000.
11. Vese, L. A., "Multiphase detection and image segmentation," *UCLA C.A.M. Report 02-36*, June 2002.
12. Chang, K, *Handbook of Optical Components and Engeneering*, Wiley-Interscience, New Jersey, 2003.
13. Gonzales, C. G., R. E. Woods, and S. L. Eddins, *Digital Image Processing Using Matlab*, Pearson Prentice Hall, New Jersey, 2004.

Model of a Reactor Chamber with Microwave Heating

Petr Drexler, **Tomáš Jirku**, **Zoltán Szabó**, **and Pavel Fiala**
Department of Theoretical and Experimental Electrical Engineering
Brno University of Technology, Kolejni 4, 612 00 Brno, Czech Republic

Abstract— This contribution brings the results obtained by numerical modeling of a reactor chamber design. The reactor chamber is intended for the preparation of oil. The chamber is designed for the operating frequency $f = 2.4\,\mathrm{GHz}$ with regulated microwave power. The power regulation allows the optimization of heat distribution in non-homogenous material with respect to chemical and physical material changes — phase changes. The paper presents a numerical model of a coupled problem. The coupled problem includes the radio frequency model and thermal field with non-linear thermal material. It was solved using sophisticated combined numerical method — the finite element method (FEM) and the finite volume method (FVM). The used material respects its own phase change from the liquid condition to the gas phase. The apparatus serves as an experimental prototype of the commercial one, that is currently being used by the ARS — Altman Recovery System Company.

1. INTRODUCTION

The reactor exploiting active porous substance was designed for oil preparation. The reactor is fed with industrially produced mixture of oil and water. The desired reaction proceeds in the ceramic porous material. To achieve the desired reaction condition, it is necessary to heat the material and, simultaneously, remove the products of the reaction. After the reaction of water, further heating is undesirable with respect to side reactions.

Considering the above mentioned requirements, microwave heating was chosen. The microwave heating effect is selective for the reaction of water. The designed reactor operates at a frequency of $f = 2.4\,\mathrm{GHz}$, which allows selective heating in the active porous material of the chamber. The basic scheme of the reactor is shown in Figure 1.

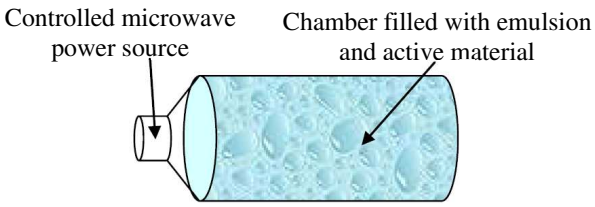


Figure 1: The basic scheme of the reactor.

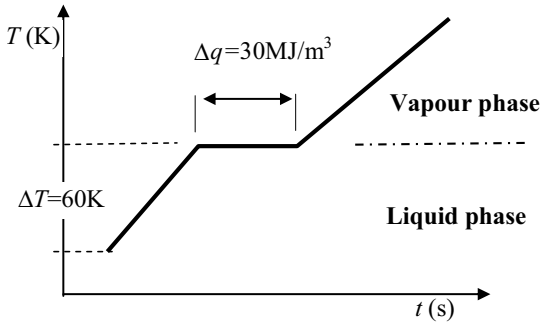


Figure 2: The phase-change characteristics of water.

2. MATHEMATICAL MODEL

It is possible to carry out an analysis of an MG model as a numerical solution by means of the Finite element method (FEM). The electromagnetic part of the model is based on the solution of full Maxwell's equations

$$\nabla \times \boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial t}, \quad \nabla \times H = \sigma \boldsymbol{E} + \frac{\partial \boldsymbol{D}}{\partial t} + \boldsymbol{J}, \quad \nabla \cdot \boldsymbol{D} = \rho, \quad \nabla \cdot \boldsymbol{B} = 0 \quad \text{in } \Omega \tag{1}$$

where $\boldsymbol{E}$ and $\boldsymbol{H}$ is the electrical filed intensity vector and the magnetic filed intensity vector, $\boldsymbol{D}$ and $\boldsymbol{B}$ are the electrical field density vector and the magnetic flux density vector, $\boldsymbol{J}_S$ is the current density vector of the sources, ρ is the density of free electrical charge, γ is the conductivity of the material and Ω is definition area of the model. The relationships between the electrical and the magnetic field intensities and densities are given by material relationships

$$\boldsymbol{D} = \varepsilon \boldsymbol{E}, \qquad \boldsymbol{B} = \mu \boldsymbol{H}. \tag{2}$$

The permittivity ε, the permeability μ and the conductivity γ in HFM are generally tensors with main axes in the direction of the Cartesian coordinates x, y, z. When all the field vectors performs rotation with the same angular frequency ω, it is possible to rewrite the first Maxwell equation

$$\nabla \times \underline{\boldsymbol{E}} = -j\omega\mu\underline{\boldsymbol{H}}, \quad \nabla \times \underline{\boldsymbol{H}} = (\sigma + j\omega\varepsilon)\underline{\boldsymbol{E}} + \underline{\boldsymbol{J}}_S \quad \text{in } \Omega, \tag{3}$$

where $\underline{\boldsymbol{E}}$, $\underline{\boldsymbol{H}}$, $\underline{\boldsymbol{J}}_S$ are field complex vectors. Taking into account boundary conditions given in (1) and after rearranging (3) we get

$$(j\omega)^2\varepsilon\underline{\boldsymbol{E}} + \sigma\underline{\boldsymbol{E}} + \nabla \times \mu^{-1}\nabla \times \underline{\boldsymbol{E}} = -j\omega\underline{\boldsymbol{J}}_S. \tag{4}$$

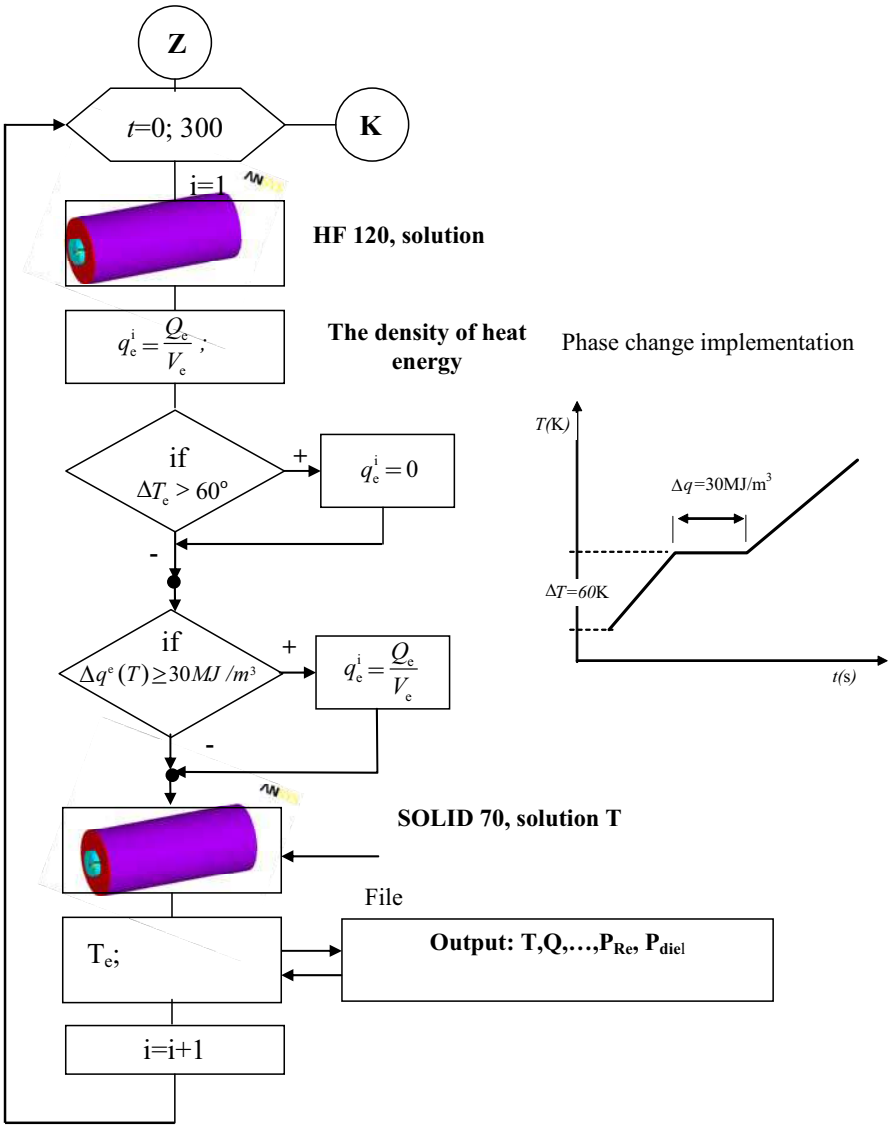


Figure 3: The simplified iteration algorithm of the model evaluation.

We apply the Gallerkin method with vector approximation functions $\boldsymbol{W}_i$ and use the vector form of the Green theorem on the double rotation element [3]. After discretation we get the expression

$$-k_0[M]\{\boldsymbol{E}\} + jk_0[C]\{\boldsymbol{E}\} + [K]\{\boldsymbol{E}\} = \{\boldsymbol{F}\}, \tag{5}$$

where $\{\boldsymbol{E}\}$ is the column matrix of the electrical intensity complex vectors. The matrixes $[K]$, $[C]$ and $[M]$ are in the form that is given in manual [4] and vector $\{\boldsymbol{F}\}$ is evaluated from the expression

$$\{\boldsymbol{F}\} = -jk_0Z_0 \int_{\Omega} [\boldsymbol{W}_i]\{\boldsymbol{J}_S\}d\Omega + jk_0Z_0 \int_{\Gamma_0+\Gamma_1} [\boldsymbol{W}_i]\{\boldsymbol{n} \times \boldsymbol{H}\}d\Gamma. \tag{6}$$

The vector approximation functions $\boldsymbol{W}$ are given in manual [4]; k_0 is the wave number for vacuum, Z_0 is the impedance of free space. The set of equation (5) is independent of time and gives $\underline{\boldsymbol{E}}$. For the transient vector $\boldsymbol{E}$ we can write

$$\boldsymbol{E} = \mathrm{Re}\left\{\underline{\boldsymbol{E}} \mathrm{e}^{\mathrm{j}\omega t}\right\}. \tag{7}$$

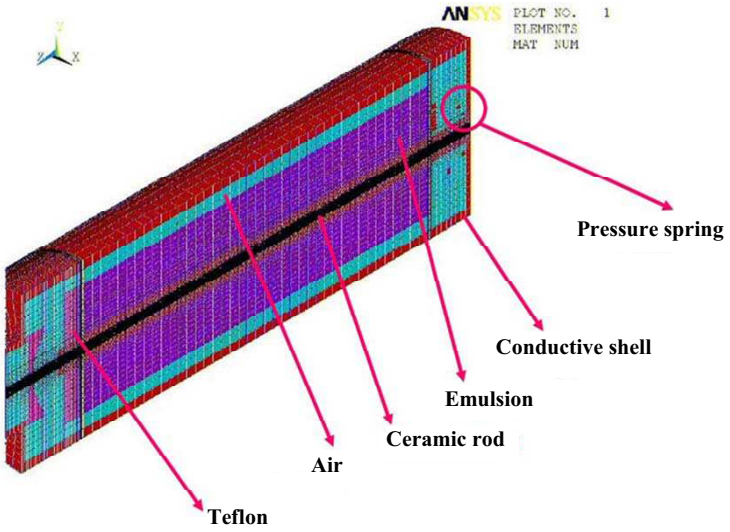


Figure 4: The geometric model of the HF reactor chamber.

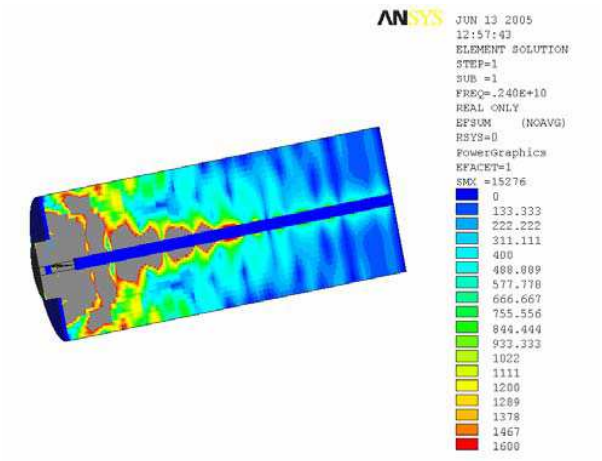


Figure 5: The distribution of the electric field intensity vector module.

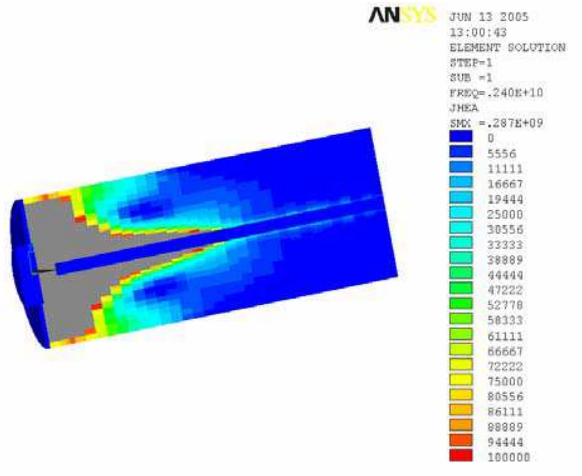


Figure 6: The distribution of the current density vector module.

The results were obtained by the solution of the non-linear thermal model with phase change of the medium. The phase change occurs via the phase conversion of water to steam. Figure 2 shows the phase-change time characteristic of water. The thermal model is based on the first thermodynamic law

$$q + \rho\, c\, \boldsymbol{v} \cdot \mathrm{div} T - \mathrm{div}(k\ \mathrm{grad} T) = \rho c \left(\frac{\partial T}{\partial t}\right), \tag{8}$$

where q is the specific heat, ρ is the specific weight, c is the specific heat capacity, T is the temperature, t is the time, k is the thermal conductivity coefficient, v is the medium flow velocity. If we consider the Snell's principle, the model can be simplified as

$$q - \mathrm{div}(k\ \mathrm{grad} T) = \rho c \left(\frac{\partial T}{\partial t}\right) \tag{9}$$

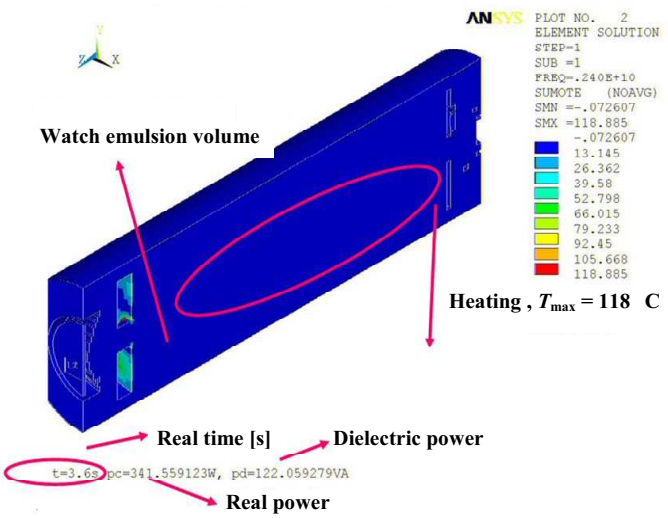


Figure 7: Temperature distribution in the reactor chamber at a specific moment.

The solution was obtained with the help of the ANSYS solver. The iteration algorithm (FEM/FVM) was realized using the APDL language as the main program. The simplified description of the algorithm is shown in Figure 3.

3. FEM/FVM MODEL

A geometric model using HF119, HF120 and SOLID70 [1] in ANSYS software was built — Figure 4. A solution of the coupled field model was performed using the APDL program. According to the microwave model solution the specific heat is evaluated. The non-linear thermal model including the phase change solves the temperature distribution. The analysis was performed for the time interval $t \in <0,\ 300>$ s and the analysis results were experimentally verified. The simulated results were found to correspond to measured values. A middle electrode was used in the model. The purpose of the middle electrode was to ensure the homogenous distribution of electromagnetic power and, subsequently, to increase the reaction efficiency.

4. CONCLUSION

Different variants of the reactor design were analyzed. The aim was to evaluate the process and time of desiccation. Figures 5, 6 and 7 show the distribution of the electric field intensity vector module, the current density vector module and the temperature in the reactor chamber. Based on the results of the analysis, the concept of the industrial reactor was designed. The reactor is under tests currently. This problem was solved using the novel combined FEM/FVM method, which is very similar to the exact physical reality.

ACKNOWLEDGMENT

The research described in the paper was financially supported by FRVS grant, by research plan No. MSM 0021630516 and grant GAAV No. B208130603.

REFERENCES

1. *Manuals ANSYS*, Vol. 8.1, USA.
2. Fiala, P., "Coupled electromagnetic model of smoothing choke," *AMTEE'01 Fifth International Conference*, Plzeň, str. C05, ISBN 80-7082-756-4, 10.9–12.9, 2001.

Design, Numerical Analysis and Test of HF Absorber

E. Kadlecová, P. Fiala, M. Zeman, M. Steinbauer, and Z. Szabó
Department of Theoretical and Experimental Electrical Engineering
Brno University of Technology, Kolejní 2906/4, 612 00 Brno, Czech Republic

Abstract— The paper presents numerical analysis of the microwave absorber design. It is possible to use modified concept of the absorber. There were used novel numerical methods (finite element method (FEM)) for thin layers modelling with sandwich non-isotropic electromagnetic materials. Modified absorber has appropriate properties and it is possible to use it in non-reflecting chamber construction. The non-reflecting chamber will be used for open space testing of relativistic microwave pulse generator; $P_{\max} = 500\,\mathrm{MW}$, $t_{\mathrm{p}} = 10 - 100\,\mathrm{ns}$. An experimental testing of the proposed pyramidal absorbers was done in the laboratories at University of Defence Brno, Czech Republic.

1. INTRODUCTION

It is possible to measure the power supplied by open space impulse by use of the calorimetric converter. The sensor is connected to the measuring device by help of the coaxial line of length approximately $l = 10\,\mathrm{m}$. The closed space has to be realised as a deaden room, see Fig. 1.

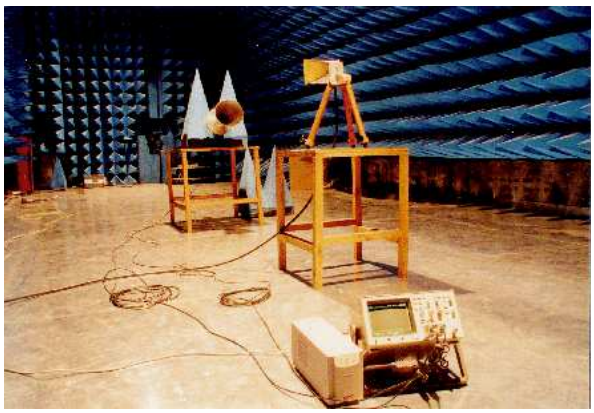


Figure 1: Measurement in semi-anechoic room.

Figure 2: The shielded laboratory.

The calorimetric sensor has disc design. The carbon with changed crystal lattice is used as one of the thin layers. Combined calorimetric sensors was designed for microwave vircator with output power $P_{\max} = 500\,\mathrm{MW}$, length of pulse $t_{\mathrm{p}} \in < 10,\ 100 > \mathrm{ns}$.

Usually, two types of the absorber materials (non-reflecting) are used. The polyurethane foam impregnated by graphite is used in high frequency range (more than 1 GHz) — the foam employs heat losses in material. The ferrite absorber is used in low frequency range — the ferrite absorber employs magnetic losses. The mentioned materials are used to prevent reflection of the electromagnetic waves inside the laboratory. The idea is to rebuild the laboratory in Fig. 2 to the shielded non-reflecting chamber for measurements and experiments with power microwave pulse generators $-P_{\max} = 500\,\mathrm{MW}$, length of pulse $t_{\mathrm{p}} \in < 1,\ 100 > \mathrm{ns}$.

2. SHIELDING OF THE WALLS

It is possible to define an electromagnetic shielding by help of shielding coefficient τ. It represents the ratio of the electrical field E_t (magnetic field H_t) in the given place of the chamber to the electrical field E_i (magnetic field H_i) incoming at the absorber.

$$\tau = \frac{E_t}{E_i}, \qquad \tau = \frac{H_t}{H_i} \tag{1}$$

The metal plate is used to prevent the electromagnetic wave go trough the wall. The damping effect of the metal plate consists of three damping — reflection, absorption and repeatable reflection.

2.1. Effectiveness of the Reflection

Shielding by means of reflection $\boldsymbol{R}$ is caused by partial reflection of the electromagnetic wave on the impedance boundary. The impedance boundary consists of the air — characteristic impedance Z_0 and metal plate — characteristic impedance Z_M. The characteristic impedances and the amount of the damping from [1] are given

$$Z_0 = \sqrt{\frac{\mu_0}{\varepsilon_0}} \qquad Z_M = \sqrt{\frac{j\omega\mu}{\sigma}} \tag{2}$$

$$R = 20\log\left|\frac{Z_0 + Z_M}{2Z_M} - \frac{Z_0 + Z_M}{2Z_0}\right| = R_1 + R_2 \tag{3}$$

2.2. Effectiveness of the Absorption

When electromagnetic wave travels trough metal plate it is absorbed and it causes heat losses. Absorption coefficient $\boldsymbol{A}$ of the metal plate is given [1]

$$A = 20\log\left|e^{\gamma t}\right| = 20\log e^{\frac{t}{\delta}}[\mathrm{dB}], \tag{4}$$

where δ is depth of penetration into the metal plate of thickness t. The depth of penetration is given

$$\delta = \sqrt{\frac{2}{\omega\mu\sigma}}[\mathrm{m}], \tag{5}$$

where ω is frequency, σ is electrical conductivity, μ is the permeability of vacuum. After manipulation of (4) we get the absorption coefficient

$$A = 8.69 \cdot \frac{t}{\delta}[\mathrm{dB}]. \tag{6}$$

2.3. Pyramidal Absorbers

Nowadays, the longitudinal non-homogenous loss environment is made from pyramid or cylinder Fig. 3. It uses dielectric losses and is made from polystyrene or polyurethane that is impregnated by graphite. Linearly increasing cross-section of the pyramid serves as an impedance transformer. The transformer transforms open space impedance to the very low impedance of the absorber. However, the damping maximum takes place in the back side of the absorber. It is necessary to have pyramid with height of $\lambda/4$ on the lowest working frequency.

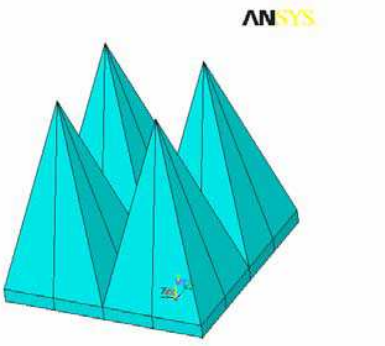


Figure 3: Geometrical model of pyramidal absorber.

3. MATHEMATICAL MODEL

It is possible to carry out analysis of a MG model as a numerical solution by help of Finite element method (FEM). The electromagnetic part of the model is based on the solution of full Maxwell's equations.

$$rot\ \boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial t}, \quad rot\ \boldsymbol{H} = \sigma\boldsymbol{E} + \frac{\partial \boldsymbol{D}}{\partial t} + \boldsymbol{J}_S, \quad div\ \boldsymbol{D} = \rho, \quad div\ \boldsymbol{B} = 0 \quad \text{in region } \Omega \tag{7}$$

where $\boldsymbol{E}$ and $\boldsymbol{H}$ is the electrical filed intensity vector and the magnetic filed intensity vector, $\boldsymbol{D}$ and $\boldsymbol{B}$ are the electrical field density vector and the magnetic flux density vector, $\boldsymbol{J}_S$ is the current density vector of the sources, ρ is the density of free electrical charge and γ is the conductivity of the material, Ω is definition area of the model. The relationships between electrical and magnetic field intensities and densities are given by material relationships

$$\boldsymbol{D} = \varepsilon \boldsymbol{E}, \qquad \boldsymbol{B} = \mu \boldsymbol{H}. \tag{8}$$

The permittivity ε, the permeability μ and the conductivity γ in HFM are generally tensors with main axes in the direction of the Cartesian co-ordinates x, y, z. When all the field vectors performs rotation with the same angular frequency ω, it is possible to rewrite the first Maxwell equations

$$\nabla \times \underline{\boldsymbol{E}} = -j\omega\mu \underline{\boldsymbol{H}}, \quad \nabla \times \underline{\boldsymbol{H}} = (\sigma + j\omega\varepsilon)\underline{\boldsymbol{E}} + \underline{\boldsymbol{J}}_S \quad \text{in region } \Omega \tag{9}$$

where $\underline{\boldsymbol{E}}$, $\underline{\boldsymbol{H}}$, $\underline{\boldsymbol{J}}_S$ are field complex vectors. Taking into account boundary conditions given in (7) and after rearranging (9) we get

$$(j\omega)^2 \varepsilon \underline{\boldsymbol{E}} + \sigma \underline{\boldsymbol{E}} + \nabla \times \mu^{-1} \nabla \times \underline{\boldsymbol{E}} = -j\omega \underline{\boldsymbol{J}}_S. \tag{10}$$

We apply Galerkin's method with vector approximation functions $\boldsymbol{W}_i$. We use vector form of the Green theorem on the double rotation element [3]. After discretisation we get the expression

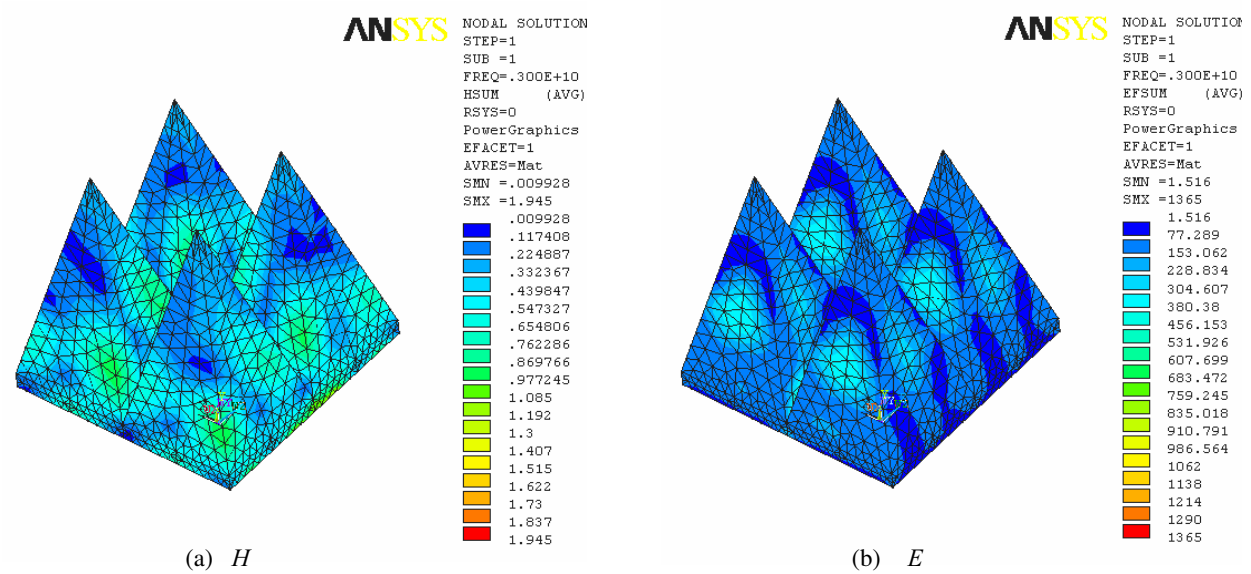
$$-k_0[M]\{\boldsymbol{E}\} + jk_0[C]\{\boldsymbol{E}\} + [K]\{\boldsymbol{E}\} = \{\boldsymbol{F}\}, \tag{11}$$

where $\{\boldsymbol{E}\}$ is column matrix of electrical intensity complex vectors. The matrixes [K], [C] and [M] are in the form that is given in manual [4] and vector $\{\boldsymbol{F}\}$ is evaluated from expression

$$\{\boldsymbol{F}\} = -jk_0 Z_0 \int_{\Omega} [\boldsymbol{W}_i]\{\boldsymbol{J}_S\} d\Omega + jk_0 Z_0 \int_{\Gamma_0 + \Gamma_1} [\boldsymbol{W}_i]\{\boldsymbol{n} \times \boldsymbol{H}\} d\Gamma. \tag{12}$$

Vector approximation functions $\boldsymbol{W}$ are given in manual [4], k_0 is wave number for vacuum, Z_0 is impedance of the free space. The set of equation (5) is independent of time and it gives $\underline{\boldsymbol{E}}$. For transient vector $\boldsymbol{E}$ we can write

$$\boldsymbol{E} = \mathrm{Re}\left\{\underline{\boldsymbol{E}} \mathrm{e}^{\mathrm{j}\omega t}\right\}. \tag{13}$$



(a) H (b) E

Figure 4: Evaluation of modules of magnetic intensity H and electric intensity E.

4. GEOMETRICAL MODEL

The geometrical model was created with standard tools in FEM system ANSYS with help of automated mesh and nodes generators. After that the mathematical model was formed. The applied element is HF120. The geometrical model of HF absorber is shown in Fig. 3. The initial frequency was taken $f = 1,\ 2,\ 3,\ 4,\ 5\,\mathrm{GHz}$. The concept of the absorber uses special properties of the pyrographite and ferrite dust. These are placed on the pyramid walls with appropriate densities. By means of this concept it is possible to achieve appropriate damping of the incoming electromagnetic wave.

(a) (b)

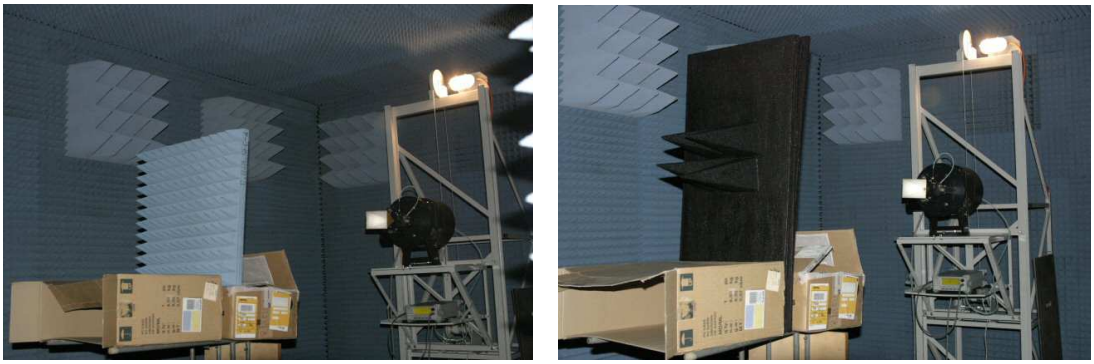
Figure 5: Experiments with HF absorber-attenuation tests in UO laboratories, (a) $l = 2\,\mathrm{m}$ distance between absorbers and measuring antenna, (b) testing absorbers.

(a) (b)

Figure 6: Experiments with HF absorber-attenuation tests in UO laboratories, (a) computer controlled microwave generator, (b) $l = 10\,\mathrm{cm}$ distance between absorbers and measuring antenna.

5. RESULTS OF THE ANALYSIS

In the Fig. 4, there is shown module of vector functions of magnetic intensity H and electric intensity E. There are results for initial TE plane wave with frequency $f = 3\,\mathrm{GHz}$, $E_\mathrm{x} = 100\,\mathrm{V/m}$. We can see the quality of electromagnetic absorber in one numerical solution in Fig. 4. One can see good absorption of specific power. The adaptive numerical model of the absorber was created in the ANSYS program with APDL. It is possible to change the absorber geometry and to perform the optimization according to given parameters. The results of the model was used in construction of the prototype. The prototype was measured according to (1) to (6) in the laboratories at UD Brno. There were used generator and horn antenna with frequency $f = 10\,\mathrm{GHz}$. There were tested HF absorber with $l = 100\,\mathrm{mm}$ and $2000\,\mathrm{mm}$ inside of receiving antenna, see Figs. 5–7. Measured attenuation was from 16.09 to 20.8 dB. It depends on number of used absorption plates. The testing room and apparatus were calibrated with standard absorbers and result was compared with experimental model. The design of HF absorber was made with special low-cost material.



(a) (b)

Figure 7: Experiments with HF absorber-attenuation tests in UO laboratories, (a) testing measurement (calibration) with classical HF absorber, (b) testing experimental absorbers.

6. CONCLUSION

Presented work is a part of the project conducted in co-operation with VTUPV Vyskov, PROTOTYPA a.s. and TESLA Vrsovice Praha. The shielded non-reflecting laboratory for microwave pulse generators is being built. In order to perform the tests of Vircator pulsed generator it is necessary to have a laboratory with shielding up to the frequency of RTG and the laboratory has to eliminate the reflections of the EMG wave in the rage $f > 10\,\mathrm{MHz}$. It is necessary to have absorption coefficient of the EMG wave $A_{\min} > 16.5\,\mathrm{dB}$ for $f = 1\,\mathrm{GHz}$. The HF absorbers were discussed and the analysis of the basic properties was performed for chosen pyramidal absorber. The absorption coefficient resulting from numerical analysis is compared with experimental measurement.

ACKNOWLEDGMENT

The research described in the paper were financially supported by FRVŠ by research plan No. MSM 0021630513 ELCOM and grant GAAV No. B208130603.

REFERENCES

1. Svačina, J., *Elektromagnetická Kompatibilita (Principy a Metody)*, paperback VUT v Brně, ISBN 80-214-1873-7.
2. Fiala, P., "Koncepce a návrh kalorimetrického čidla ultrakrátkého elektromagnetického impulsu," research report, VUT FEKT Brno, 70pgs, 18.3.2003.
3. Stratton, J. A., *Teorie Elektromagnetického Pole*, SNTL Praha, 1961.
4. *ANSYS Theory Reference Manual*, Swanson Analysis System, Inc., Huston, USA.

Properties and Numerical Simulation of $CaCl_2.6H_2O$ Phase Change

I. Behunek, T. Bachorec, and P. Fiala

Department of Theoretical and Experimental Electrical Engineering
Faculty of Electrical Engineering and Communication
Brno University of Technology, Kolejni 2906/4, 612 00 Brno, Czech Republic

Abstract— This paper deals with the application of phase change materials (PCM) for thermal management of integrated circuits as a viable alternative to active forced convection cooling systems. The paper presents an analytical description of melting and freezing process, an analytical solution of $CaCl_2.6H_2O$ phase change in 1D and the results of a numerical model of phase change. The numerical model joins thermal field, air flow with respects to material phase change. This numerical model was obtained by means combined of the finite element method (FEM) and finite volume method (FVM) applied as main method in ANSYS software.

1. INTRODUCTION

Phase change materials can store large amounts of heat without undergoing significant temperature changes because of their high latent heat of fusion. We know different PCM — inorganic, organic, and combinations of amorphous and crystallic substances, clathrates etc. The stored energy is given by the calorimetric equation

$$Q = \rho V \Delta h_m + \int_{T_0}^{T_m} V \rho c \mathrm{d}T + \int_{T_m}^{T_e} V \rho c \mathrm{d}T \tag{1}$$

where ρ is density, V volume, c specific heat, Δ_{hm} enthalpy, Q heat, and T_m temperature. The PCM may be used for active or passive electronic cooling applications with high power at the package level.

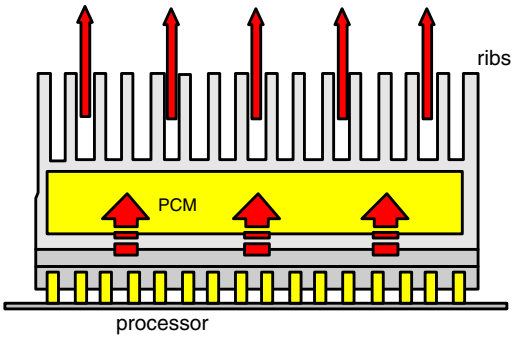


Figure 1: Example of processor cooler with phase change material.

2. PROPERTIES OF $CACL_2.6H_2O$

Big latent heat, good thermal conductivity and inflammability are the main advantages of inorganic materials. But they cause corrosion and suffer from loss the of water. Incongruent melting and supercooling are the biggest problem of their exploitation. During melting and freezing there are precipitations of other phases, which do not take part in next process of charging and discharging. Poor nucleation, slow rate of crystal growth or high rate of heat removal may be the reason for supercooling. Impurities have a strong influence on the cooling curves as well.

In the Fig. 2 there is exhibited the phase change of $CaCl_2.6H_2O$ during heating and cooling. Dashed lines would show theoretical behaviour if melting and freezing were at constant temperature T_m — case of pure crystallic substance. Impurity and methodology of measuring (probe always is only in small amount of hexahydrate) are the causes of variances. During solidification the supercooling occurred because of weak nucleation. Crystallization was initiated due to a solid particle

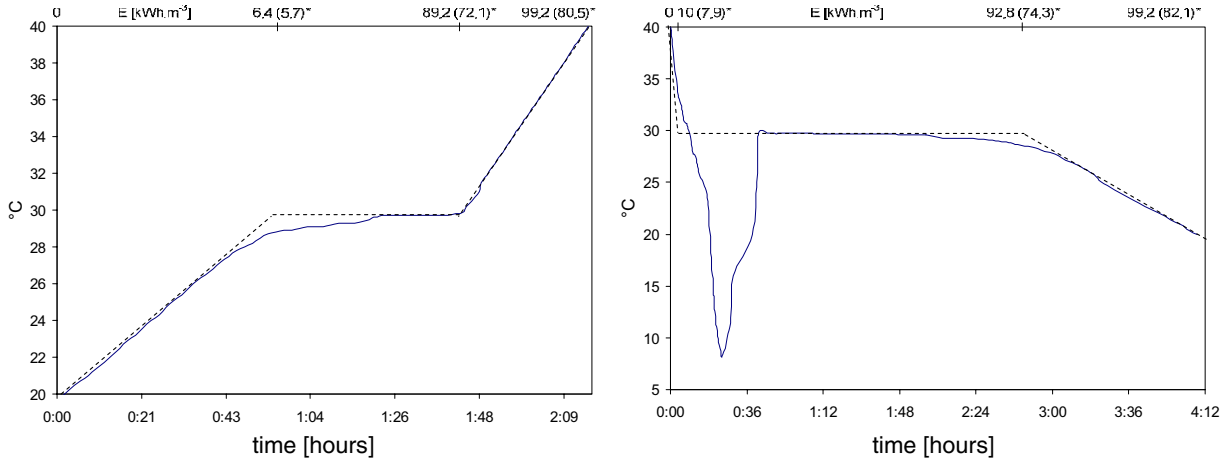


Figure 2: Phase change of $CaCl_2.6H_2O$ (right — melting, left — freezing).

of PCM which we put to the measured sample. Otherwise there would not be any crystallization. We can use plastics, mild steel or copper for enclosures. Aluminium or stainless steel are not suitable. Sometimes we can see temperature fluctuation above T_m during solidification. We found an explanation in the binary diagram.

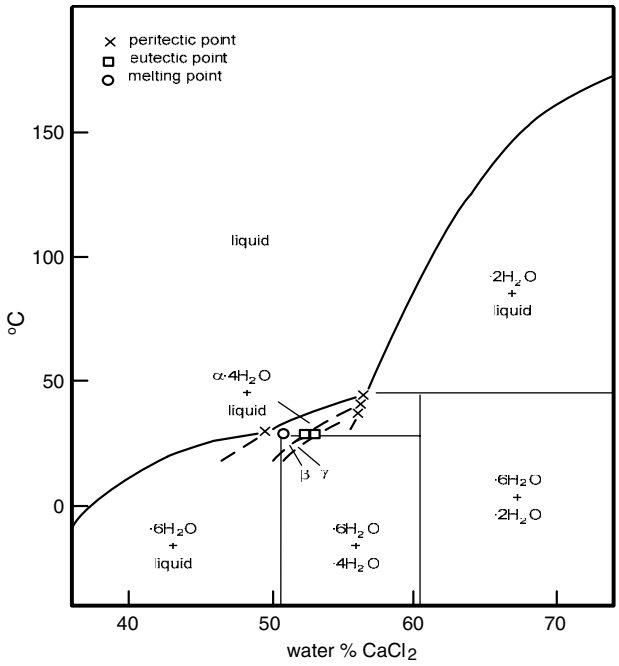


Figure 3: Binary diagram of $CaCl_2.6H_2O$.

Modification is possible to do by additives. From number of potential candidates $Ba(OH)_2$, $BaCO_3$, and $Sr(OH)_2$ were chosen. They seemed to be feasible. When we used $Ba(OH)_2$ and $Sr(OH)_2$ at 1% part by weight there was no supercooling. We could increase stability of the equilibrium condition with addition of KCl (2 wt%) and NaCl. NaCl is weak soluble in $CaCl_2.6H_2O$, therefore part by weight is only about 0.5%. Disadvantage is that melting point decreases about 3°C at 26–27°C. The melting point of pure $CaCl_2.6H_2O$ is 29.6°C. Due to availability on the market and price we chose for modification $BaCO_3$. We obtained the best results for 1.2 wt%. In Fig. 1 we can see that supercooling is 3–4°C but then crystallization started spontaneously and temperature increased at 28–29°C. It is obvious that nucleation is slower in comparison with pure crystallic matter. Supercooling is not considered as a disadvantage. If we use $CaCl_2.6H_2O$ in the heat accumulator we will be able to store energy at lower temperature (about 26°C) and suppress heat losses. Next disadvantage of $BaCO_3$ is that it carbonates because of atmospheric CO and CO_2. This means the loss of properties. In our case $BaCO_3$ will be isolated from surrounding environment.

3. ANALYTICAL DESCRIPTION AND SOLUTION OF PHASE CHANGE

The analytical solution is exact but we consider several simplified assumptions. The most important of them is that we can solve solidification of PCM only in one dimensional body. The 1D body could be semifinite plane, cylinder or sphere.

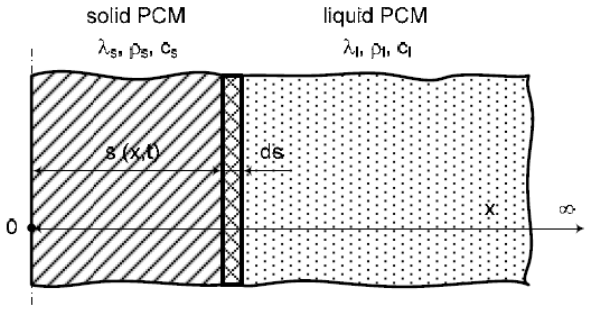


Figure 4: Solidification of semi-infinite plate of PCM.

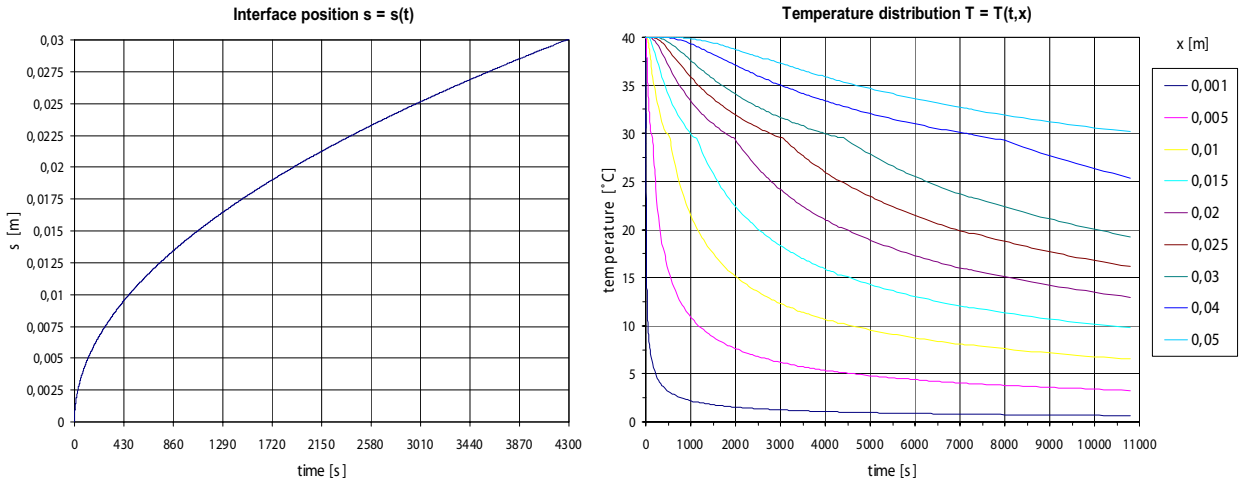


Figure 5: Interface position and temperature distribution in $CaCl_2.6H_2O$.

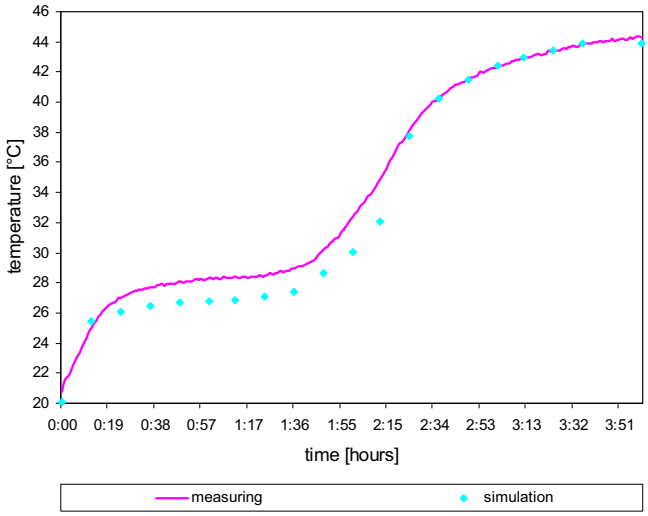


Figure 6: Phase change of $CaCl_2.6H_2O$ (the comparison between measuring and simulation).

We consider semi-infinite mass of liquid PCM at initial temperature T_0 which has been cooled

by a sudden drop of surface temperature $T_p = 0°\mathrm{C}$. This temperature is constant during the whole process of solidification. The simplified assumptions are

- body is semi-infinite plane, heat flux is one-dimensional in the x-axis, plane is an interface between the solid and the liquid, there is an ideal contact on the interface, temperature of surface is constant ($T_p = 0°\mathrm{C}$), crystallization of PCM is at constant temperature T_m, thermophysical properties of the solid and the liquid are different but they do not depend on temperature, there is no natural convection in the liquid.

Initial and boundary conditions:

- initial temperature T_0 for $x \geq 0$ at time 0 s
- temperature equals T_m on the interface between the solid and the liquid ($x = s$)

$$x = s \wedge t > 0 \quad \Rightarrow T_s = T_l = T_m = \mathrm{const} \tag{2}$$

- evolved latent heat during a motion of interface (volume element ds, area $1\,\mathrm{m}^2$, time 1 s)

$$\mathrm{d}Q_{\Delta h_m} = \Delta h_m \rho_l 1 \frac{ds}{dt} \tag{3}$$

- position of interface is function of time

$$s = s(t) = 2\varepsilon\sqrt{a_s t}, \tag{4}$$

this dependence is called the parabolic law of solidification.

- boundary condition for phase change

$$\lambda_s \left(\frac{\partial T_s}{\partial x}\right)_{x=s} = \lambda_l \left(\frac{\partial T_l}{\partial x}\right)_{x=s} + \Delta h_m \rho_l \frac{ds}{dt} \tag{5}$$

for

$$x \to \infty \wedge t > 0 \Rightarrow T_l = T_0 = \mathrm{const} \tag{6}$$

$$x = 0 \wedge t \geq 0 \Rightarrow T_p = T_s(x = 0) = 0°C \tag{7}$$

If we solve fourier relations of heat conduction with conditions above for the solid and the liquid, we get the following equations which allow calculating temperatures in solid, liquid PCM and the location of interface. The results are in the Fig. 5 which are related to Eqs. (8), (9), (10).

$$\frac{e^{-\varepsilon^2}}{\mathrm{erf}(\varepsilon)} - \frac{\lambda_l}{\lambda_s}\sqrt{\frac{a_s}{a_l}}\frac{T_0 - T_m}{T_m}\frac{e^{-\frac{a_s}{a_l}\varepsilon^2}}{\mathrm{erfc}\left(\varepsilon\sqrt{\frac{a_s}{a_l}}\right)} = \frac{\Delta h_m \rho_1 \varepsilon a_s \sqrt{\pi}}{\lambda_s T_m} \tag{8}$$

We get $\varepsilon = 0.3514$ for $CaCl_2.6H_2O$. Temperatures in the solid and liquid part of body are

$$T_s = \frac{T_m}{\mathrm{erf}(\varepsilon)}\mathrm{erf}\left(\frac{x}{2\sqrt{a_s t}}\right) \tag{9}$$

$$T_l = T_0 + \frac{T_0 - T_m}{\mathrm{erfc}\left(\varepsilon\sqrt{\frac{a_s}{a_l}}\right)}\mathrm{erf}\left(\frac{x}{2\sqrt{a_l t}}\right) \tag{10}$$

4. NUMERICAL SOLUTION

In Fig. 6 there is a result of numerical model. We can see time dependence of $CaCl_2.6H_2O$ temperature in the package. We can compare results obtained by numerical simulation with the measuring. Differences between simulation and measuring are caused due to inaccuracy of model with respect to reality. We used tabular values of pure $CaCl_2.6H_2O$ but in pipes there is modified hexahydrate with 1.2% of $BaCO_3$. We would need to know temperature dependence of thermal conductivity, specific heat, and density during phase change exactly.

5. CONCLUSION

We showed measured properties of $CaCl_2.6H_2O$ and did its modification in order to avoid supercooling. There is the demonstration in the article that we can apply theory of solidification of metals to PCM. We presented analytical description and solution phase change in 1D. In 3D we had to exploit numerical method. The numerical model was solved by means of combined FEM/FVM as main software in ANSYS program. If we compare the results of simulation with experimental measuring, we see quite good congruence.

ACKNOWLEDGMENT

The paper was prepared within the framework of the research plan No. MSM 0021630516 of the Ministry of Education, Youth and Sports of the Czech Republic.

REFERENCES

1. *Ansys User's Manual*, Svanson Analysys System, Inc., Huston, USA, 2003.
2. Běhunek, I. "Properties of inorganic PCM," *Honeywell EMI Conference and Competition 2005*, Brno, Ing. Zdeněk Novotný CSc., Ondráčkova 105, ISBN 80-214-2942-9, Brno, 2005.
3. Lane, G. A., *Solar Heat Storage: Latent Heat Materials, Volume II: Technology*, CRC Press, Inc., ISBN 0-8493-6586-4, Boca Raton, Florida, USA, 1986.

Turbulence Modeling of Air Flow in the Heat Accumulator Layer

I. Behunek and P. Fiala
Department of Theoretical and Experimental Electrical Engineering
Faculty of Electrical Engineering and Communication
Brno University of Technology, Kolejni 2906/4, 612 00 Brno, Czech Republic

Abstract— The article deals with one layer of a heat accumulator which is suitable for solar systems. Here is presented numerical coupled model with turbulent flow thermal field depends on time. There is a numerical model of the air turbulence, heat transfer, conduction and also phase change of $CaCl_2.6H_2O$ which is used to increase the density of stored energy. The numerical solution was done by the help of finite element method (FEM) in ANSYS software.

1. INTRODUCTION

Phase changes of materials are a perspective way of thermal energy storage. The application of PCM offers a lot of advantage. We can reach higher density of stored energy. Table 1 shows the calculation for classical materials and PCM. The initial temperature 20 °C and final temperature 50 °C at the end of heating are supposed. Next advantage is a possibility to store heat at low temperature. We don't need to have such good thermal insulation and solar collectors work with better efficiency so the demand for area of collectors decreases.

Table 1: The comparison of classical materials and PCM.

material	density of storage energy [kWh·m^{-3}]
water	34.5
gravel	23.0
paraffine wax	62.4
$CaCl_2.6H_2O$	117.4
$Na_2CO_3.10H_2O$	131.7
$Na_2HPO_4.12H_2O$	134.7

2. MATHEMATICAL AND NUMERICAL MODEL OF TURBULENCE

The model of air velocity distribution is derived for incompressible fluid as

$$\operatorname{div} \boldsymbol{v} = 0 \tag{1}$$

for a stabile state of flow stands

$$\operatorname{div} \rho\boldsymbol{v} = 0 \tag{2}$$

from the energy conservation law. We suppose a turbulent flow

$$\operatorname{rot} \boldsymbol{v} = 2\boldsymbol{\omega} \tag{3}$$

where $\boldsymbol{\omega}$ is the angular velocity of fluid. If we use the Stokes theorem, the Helmoholtz theorem for moving particle and continuity equation, we can formulate from the balance of forces the Navier-Stokes equation for the fluid element

$$\frac{\partial \boldsymbol{v}}{\partial t} + \boldsymbol{v} \cdot \operatorname{grad}\boldsymbol{v} = \boldsymbol{A} - \frac{1}{\rho}\operatorname{grad} p + \boldsymbol{v} \cdot \Delta\boldsymbol{v} \tag{4}$$

where $\boldsymbol{A}$ is an external acceleration and $\boldsymbol{v}$ kinematic viscosity. In the Equation (4) we can substitute pressure losses

$$\begin{aligned}\operatorname{grad} p \;=\; & -\left(K_x\rho v_x|\boldsymbol{v}| + \frac{f}{D_h}\rho v_x|\boldsymbol{v}| + C_x\mu v_x\right)\boldsymbol{u}_x - \left(K_y\rho v_y|\boldsymbol{v}| + \frac{f}{D_h}\rho v_y|\boldsymbol{v}| + C_y\mu v_y\right)\boldsymbol{u}_y \\ & -\left(K_z\rho v_z|\boldsymbol{v}| + \frac{f}{D_h}\rho v_z|\boldsymbol{v}| + C_z\mu v_z\right)\boldsymbol{u}_z \end{aligned} \tag{5}$$

where K are suppressed pressure losses, f resistance coefficient, D_h hydraulic diameter, C air permeability of system, μ dynamic viscosity and $\boldsymbol{u}$ unit vector of the Cartesian coordinate system. The resistance coefficient is obtained from the Boussinesq theorem

$$f = aRe^{-b} \tag{6}$$

where a, b are coefficient from [3].

The model of the velocity field is formulated from the condition of steady-state stability which is expressed

$$\int_{\Omega} \boldsymbol{f} \mathrm{d}\Omega + \int \boldsymbol{t} \mathrm{d}\Gamma = 0 \tag{7}$$

where $\boldsymbol{f}$ are specific forces in area $\boldsymbol{\Omega}$ and $\boldsymbol{t}$ are pressures, tensions and shear stresses on the interface of area $\boldsymbol{\Gamma}$. By means of the transformation into local coordinates we obtain a differential form for the static equilibrium

$$\boldsymbol{f} + \mathrm{div}^2 \boldsymbol{T}_v = 0 \tag{8}$$

where $\boldsymbol{T}_v$ is a tensor of the internal tension

$$\boldsymbol{T}_v = \begin{bmatrix} X_x & X_y & X_z \\ Y_x & Y_y & Y_z \\ Z_x & Z_y & Z_z \end{bmatrix} \tag{9}$$

where X, Y, Z are stress components which act on elements of area. It is possible to add a form of specific force from (1)–(3) to the condition of static equilibrium. The form of specific force is obtained by means of an external acceleration $\boldsymbol{A}$, by condition of pressure losses and shear stresses τ

$$\rho \left(\frac{\partial \boldsymbol{v}}{\partial t} + \boldsymbol{v} \cdot \mathrm{grad} \boldsymbol{v} \right) - \rho \boldsymbol{A} - \sum_{l=1}^{N_s} \boldsymbol{F}_1 + \mathrm{div}^2 \boldsymbol{T}_v = 0 \tag{10}$$

where $\boldsymbol{F}_1$ are discrete forces. The model which covers forces, viscosity, and pressure losses is

$$\rho \left(\frac{\partial \boldsymbol{v}}{\partial t} + \boldsymbol{v} \cdot \mathrm{grad} \boldsymbol{v} \right) - \rho \boldsymbol{A} - \sum_{l=1}^{N_s} \boldsymbol{F}_1 + \mathrm{grad}\, p - \boldsymbol{v} \cdot \Delta \boldsymbol{v} = 0 \tag{11}$$

We can prepare a discretization of Equation (20) by means of the approximation of velocity $\boldsymbol{v}$ and acceleration $\boldsymbol{a}$

$$\begin{aligned} \boldsymbol{v} &= \sum_{k=1}^{N_\varphi} \boldsymbol{v}_{vk} W_k(x,\, y,\, z), \qquad \forall (x,\, y,\, z) \subset \Omega, \\ \boldsymbol{a} &= \sum_{k=1}^{N_\varphi} \boldsymbol{a}_{vk} W_k(x,\, y,\, z), \qquad \forall (x,\, y,\, z) \subset \Omega, \end{aligned} \tag{12}$$

where $\boldsymbol{v}_v$, $\boldsymbol{a}_v$ are immediate node values, W is a base function, N_φ is a number of mesh nodes. If we apply the approximation (10) and Galerkin principle in (11) we get the semidiscrete solution

$$\begin{aligned} &\int_{\Omega} W_j \left[\rho \left(\sum_{i=1}^{N_v} W_i \frac{\partial \boldsymbol{v}_{vi}}{\partial t} + \sum_{i=l}^{N_v} W_i \boldsymbol{v}_{vi} \cdot \mathrm{grad} \boldsymbol{v}_{vi} \right) + \sum_{i=l}^{N_v} W_i \mathrm{grad}\, p_i - \rho \sum_{i=l}^{N_v} W_i \boldsymbol{A}_{vi} - \sum_{i=l}^{N_v} W_i \sum_{l=1}^{N_s} F_{li} \right] \mathrm{d}\Omega \\ &- \int_{\Omega} W_j \left[\sum_{i=1}^{N_v} W_i \mathrm{div} \boldsymbol{v} \cdot \mathrm{grad}\, \boldsymbol{v}_{vi} \right] \mathrm{d}\Omega - \int_{\Gamma} W_i \left[\sum_{i=1}^{N_v} W_i \boldsymbol{X}_i \right] \mathrm{d}\Gamma = 0, \quad j = 1,\, 2,\, \ldots,\, N_V \end{aligned} \tag{13}$$

where X are the known conditions on the interface of area. We substitute pressure losses in (13) for the function (5) and we obtain the model of air flow

$$\rho\int_{\Omega} W_j W_i \mathrm{d}\Omega \frac{\mathrm{d}\boldsymbol{v}_{vi}}{\mathrm{d}t} + \rho\int_{\Omega} W_j \boldsymbol{v}_{vi}\cdot\left(\frac{\mathrm{d}W_i}{\mathrm{d}x}\boldsymbol{u}_x + \ldots + \frac{\mathrm{d}W_i}{\mathrm{d}z}\boldsymbol{u}_z\right)\mathrm{d}\Omega\boldsymbol{v}_{vi}$$
$$+\int_{\Omega} W_j\cdot\left(\frac{\mathrm{d}W_i}{\mathrm{d}x}\boldsymbol{u}_x + \ldots + \frac{\mathrm{d}W_i}{\mathrm{d}z}\boldsymbol{u}_z\right)\mathrm{d}\Omega p_i - \rho\int_{\Omega} W_j \mathrm{d}\Omega\boldsymbol{A}_i - \int_{\Omega} W_j \mathrm{d}\Omega\boldsymbol{F}_{li}$$
$$-\int_{\Omega}\left(\frac{\mathrm{d}W_j}{\mathrm{d}x}+\frac{\mathrm{d}W_j}{\mathrm{d}y}+\frac{\mathrm{d}W_j}{\mathrm{d}z}\right)v_i\cdot\left(\frac{\mathrm{d}W_i}{\mathrm{d}x}\boldsymbol{u}_x+\frac{\mathrm{d}W_i}{\mathrm{d}y}\boldsymbol{u}_y+\frac{\mathrm{d}W_i}{\mathrm{d}z}\boldsymbol{u}_z\right)\mathrm{d}\Omega\boldsymbol{v}_{vi}$$
$$-\int_{\Gamma} W_j \mathrm{d}\Gamma\boldsymbol{X}_i = 0, \quad i,\, j = 1,\, 2,\, \ldots,\, N_v. \tag{14}$$

On the interface there are conditions

$$\boldsymbol{n}\cdot(v) = 0 \tag{15}$$

on the border Γ_{vr1} where $\boldsymbol{n}$ is a normal vector to direction of air flow

$$\boldsymbol{n}\cdot(p) = 0 \tag{16}$$

on the border Γ_{vr2} where $\Gamma_{vr1}\subset\Gamma_{vr2}$ is the interface between the solid body and the liquid. We can write

$$\left[C_{ij}^f\right]\left\{\frac{\mathrm{d}\boldsymbol{v}_{vi}}{\mathrm{d}t}\right\}+[K_{ij}^{sx}-K_{ij}^{Fx}]\{\boldsymbol{v}_{vi}\}+[K_{ij}^{cx}]p=[K_{ij}^{gx}]\{\boldsymbol{A}_i\}+[F_{ij}^{bx}]\{F_{li}\}+[F_{ij}^{sx}]\{\boldsymbol{X}_i\}\quad i,j=1,\,2,\,\ldots,\,N_v. \tag{17}$$

Matrixes C_{ij}^f, K_{ij}^{sx}, K_{ij}^{Fx}, K_{ij}^{cx}, K_{ij}^{gx}, F_{ij}^{bx}, K_{ij}^{Fx}, F_{ij}^{sx} are related to coefficients of equations (14). We can rewrite a form for an element of mesh

$$\left[C_e^f\right]\left\{\frac{\mathrm{d}\boldsymbol{v}_{vi}}{\mathrm{d}t}\right\}+\left[K_e^{sx}-K_e^{Fx}\right]\{\boldsymbol{v}_{vi}\}+[K_e^{cx}]\,p_i=[K_e^{gx}]\{\boldsymbol{A}_i\}+[F_e^{bx}]\{F_{li}\}+[F_e^{sx}]\{\boldsymbol{X}_i\}\quad e=1,\,2,\,\ldots,\,N_e. \tag{18}$$

For the solution we used the standard k-ε model, the standard k-ω model and the SST (Shear Stress Transport Model). The standard k-ε gives exact results and use two equations for turbulent kinetic energy and its dissipation. Model k-ω solves the equations for turbulent kinetic energy and its specific dissipation rate. This model gives better results in the nearness of wall but worse in the distance from wall. The SST model combines and switch between k-ε and k-ω model in order to get the best result (see [5, 6, 9]).

3. NUMERICAL MODEL OF HEAT ACCUMULATOR LAYER

There is geometric model of one layer of accumulator in the Fig. 1. It consists from 26 PVC pipes in the square configuration. Inside of pipes there are 9.36 liters of modified $CaCl_2.6H_2O$. The air flows through the layer and transfers heat into pipes. Progress of numerical solution had two parts. First we solved turbulence model and got heat transfer film coefficients. These results were the input of solution for second part when thermal model was calculated. Time dependence of temperature distribution in the layer is final result.

Initial and boundary conditions

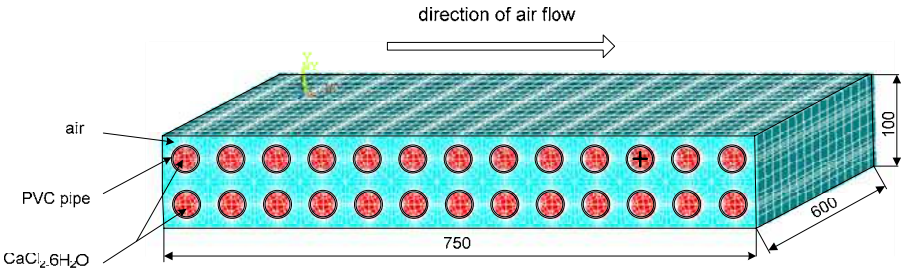


Figure 1: Geometric model of layer with mesh of elements.

- inlet temperature of the air is 50°C, inlet velocity of the air is $0.4\,\mathrm{m{\cdot}s^{-1}}$, outlet pressure is $101.3\,\mathrm{kPa} + 10$ Pa, initial temperature of the air, PVC and $CaCl_2.6H_2O$ is 20°C

There are distributions of velocity in Fig. 2 and next results for distribution of the turbulent kinetic energy, dissipation, temperature, and pressure in following pictures.

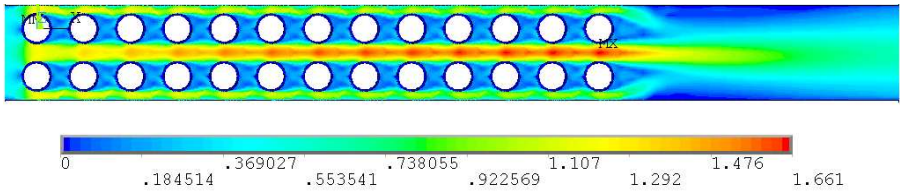


Figure 2: Velocity distribution of the air.

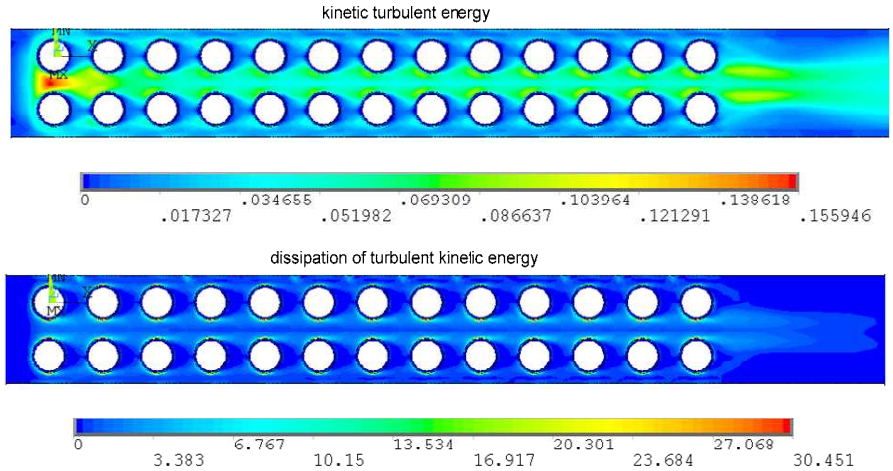


Figure 3: Distribution of kinetic energy, dissipation.

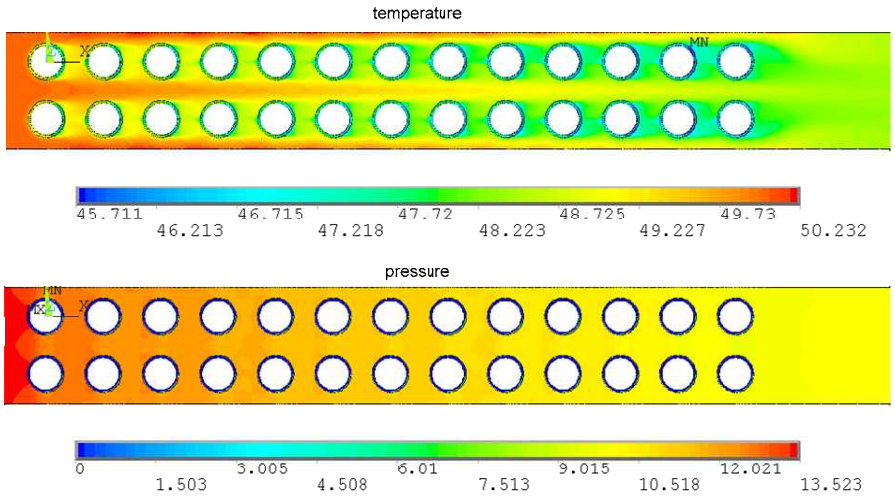


Figure 4: Distribution of temperature and pressure.

4. CONCLUSION

We presented a numerical model of one layer of heat accumulator. This heat accumulator exploits advantages and suppresses disadvantages of water and gravel accumulator. The numerical model was solved by means of FEM in ANSYS software. We solved coupled problem of air flow turbulence, heat transfer, conduction, convection, and also phase change. There are also problems with solution stability during iteration in ANSYS FLOTRAN. Finite element method (FLOTRAN) is not suitable

for turbulence modeling. The application of finite volume method (e.g., ANSYS CFX, FLUENT) for more complicated tasks is necessary.

The paper was prepared within the framework of the research plan No. MSM 0021630516 of the Ministry of Education, Youth and Sports of the Czech Republic.

REFERENCES

1. *Ansys User's Manual*, Svanson Analysys System, Inc., Huston, USA, 2003.
2. Specht, B., "Modellierung von beheizten laminaren und turbulenten Ströungen in Kanälen beliebigen Querschnitts," Technische Universität Carolo-Wilhelmina, Braunschweig, Germany, 2000.
3. Piszachich, W. S., "Nonlinear models of flow, diffusion and turbulence," Teubner Verlagsgesellschaft, Leipzig, Germany, 1985.
4. Wilcox, D. C., *Turbulence Modeling for CFD*, DCW Industries, Inc., La Canada, California, USA, 1994.
5. Zhang, Y., "Finite elemente zur Berechnung instationärer Strömungen mit bewegten Wandungen," Universität Stuttgart, Stuttgart, Germany, 1996.

Identifying of the Special Purpose Generator Pulses

P. Drexler and P. Fiala
Department of Theoretical and Experimental Electrical Engineering
Brno University of Technology, Kolejní 2906/4, 612 00 Brno, Czech Republic

Abstract— There are some suitable methods for the measurement of ultra-short solitary electromagnetic pulses (EMP) that are generated by high power microwave generators. The characteristics of EMPs are high power level ($P_{\max} = 250\,\mathrm{MW}$) and very short time duration ($t_\mathrm{p} \in <1,\ 60> \mathrm{ns}$). Special requirements for measurement methods have to be considered because of the specific EMPs properties. In the paper, two suitable methods for this application are presented. The first one, the calorimetric method, utilizes the thermal impacts of microwave absorption. The second method presented — the magneto-optic method — uses the Faraday's magneto-optic effect as a sensor principle. A combined calorimetric sensor was realized and there were made some experimental EMP measurements with good results. The sensor utilizing the magneto-optic method is still in development.

1. INTRODUCTION

In connection with the events of the last few years and with the increased number of terrorist activities, the problem of identifying and measuring the impact of electromagnetic weapons or other systems occurred. Among these weapons or systems there are also microwave sources which can reach extensive peak power of up to $P_{\max} = 250\,\mathrm{MW}$. Solitary, in some cases several times repeated, pulses lasting from $t_\mathrm{p} \in <1,\ 60> \mathrm{ns}$ cause the destruction of semiconductor junctions. The analysis of possible measuring methods convenient for the identification and measurement of the ultra-short solitary electromagnetic (EM) pulses is presented in this paper; some of the methods were selected and used for practical measurement.

2. METHODS

2.1. Method Based on Faraday's Induction Law

One group of methods is based on the application of Faraday's induction law where the pulse is located by sensor (coil with $N_\mathrm{s} = 1 \div 50$ turns). The signal induced in the coil is led to the recording device, generally an oscilloscope. Due to safety requirements, the distance between the sensor and the oscilloscope is $l = 50\,\mathrm{m}$. This parameter introduces quality decrease of the recorded information in the way of the signal amplitude reduction, change of the signal phase and the pulse prolongation.

The elimination of this limitation is in Version I, depicted in Figure 1, made by backward correction exploiting the Laplace transform. Pulses up to limit pulse length $T_{\max} = 1\,\mathrm{ns}$ were measured by this method and magnetic flux ϕ was evaluated [4]. Version II exploits the possibility of principal elimination of influence of the transmitting line between the sensor and the measuring device by an analogue U/f. converter. Available measuring devices can achieve measured pulses with the limit length $T_{\max} = 5\,\mathrm{ns}$.

The solution in Version III is similar to Version II; the difference is in the digital converter applied. By an available measuring devices application and fulfilled sampling theorem we can measure pulses with the limit length $T_{\max} = 20\,\mathrm{ns}$.

2.2. Method Based on Faraday's Magneto-optic Effect

Version IV in Figure 1 is based on Faraday's magnetooptic effect [4]. The connection between the sensor and the measuring device is implemented in the optical wavelength.

There are three basic types of the possible active sensors. The first type is a garnet with high Verdet constant, the second one is an optic fiber and the third one is based on magneto-optic properties of ferromagnetic mono/multi thin film. Other types of Version IV sensors are based on the magneto-optic Kerr's effects (MOKE), or surface MOKE (SMOKE) effect. By an available measuring devices application we can measure pulses with the limit length $T_{\max} = 0.1\,\mathrm{ns}$.

The named methods indicate the electromagnetic parts of the wave — electric or magnetic. They don't express the power conditions of the electromagnetic wave. For some of the measurement it is essential to evaluate power flow through the defined area.

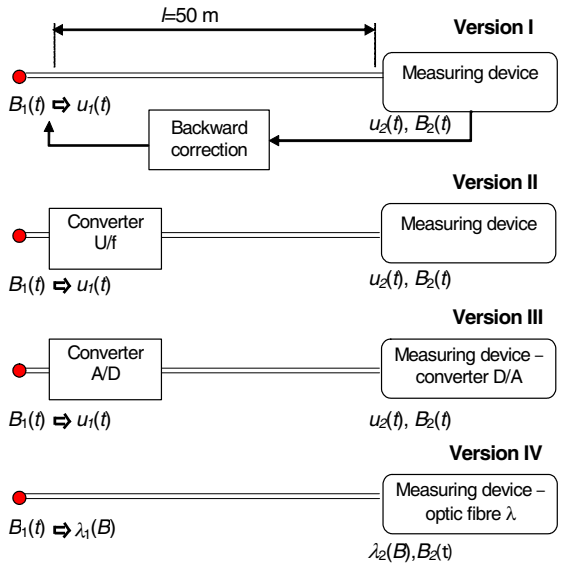


Figure 1: Principles of the methods based on Faraday's induction law and magnetooptic effect.

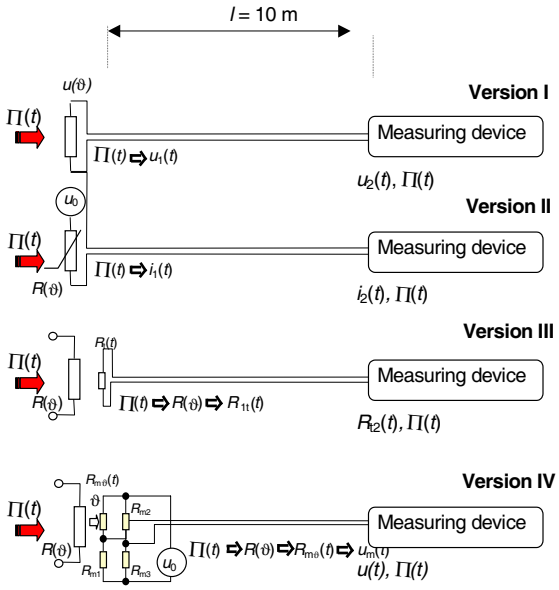


Figure 2: Calorimetric measurement of the solitary pulse power.

2.3. Calorimetric Methods

The group of calorimetric methods represents another type of converter to be introduced. We can measure power supplied by pulse (Poynting's vector when we use the calorimetric converter. The sensor is connected to the measuring device (oscilloscope) by an optic fiber of $l = 50\,\mathrm{m}$ length. Figure 2 depicts four versions of the method utilizing calorimetric measurement.

Version I discussed in [5, 6] has a sensor in the form of an ideal resistor and enables measurement of the maximum value of microwave power $P_{\max}$. The analyzed peak voltage corresponds to peak value of power $P_{\max}$. For available measuring devices we can measure pulses with the limit length $T_{\max} = 50\,\mathrm{ps}$.

Version II scans the change of resistance of the sensor, created by an evaporated thin layer, in dependence on the pulse energy. For available measuring devices we can reach the accuracy of 30% up to impulse limit length $T_{\max} = 0.1\,\mathrm{ns}$.

Version III is based on the measurement of the temperature change of the thermistor placed in contact with the layer. Under the same conditions as for the previous version we can reach the accuracy improvement of an order of magnitude.

Version IV is the bridge connection of version III. Several thermistors are attached in series to the evaporated layer; then, three resistors create a DC bridge of Weston type with the thermistors. The change of resistance in the thermistor arm is evaluated. The voltage in measuring the bridge diagonal is consequently integrated. Thus, the value of the pulse energy is obtained (and recorded by the measuring device). For available measuring devices we can measure pulses with the limit length $T_{\max} = 0.03\,\mathrm{ns}$ with accuracy to 10%.

3. USED CALORIMETRIC METHOD

The advantage of the calorimetric method is the capability of physically correct high power measurement. However, it doesn't provide the information about pulse waveform.

The calorimetric sensor was of the disc design. The carbon with changed crystal lattice was designed for use as one of the thin layer types.

3.1. Mathematical Model

It is possible to carry out an analysis of an MG model as a numerical solution by means of the Finite element method (FEM). The electromagnetic part of the model is based on the solution of full Maxwell's equations

$$\nabla \times \boldsymbol{E} = \frac{\partial \boldsymbol{B}}{\partial t},\ \nabla \times \boldsymbol{H} = \sigma \boldsymbol{E} + \frac{\partial \boldsymbol{D}}{\partial t} + \boldsymbol{J},\ \nabla \cdot \boldsymbol{D} = \rho,\ \nabla \cdot \boldsymbol{B} = 0 \text{ in } \Omega, \tag{1}$$

where $\boldsymbol{E}$ and $\boldsymbol{H}$ are the electrical field intensity vector and the magnetic field intensity vector, $\boldsymbol{D}$ and $\boldsymbol{B}$ are the electrical field density vector and the magnetic flux density vector, $\boldsymbol{J}_s$ is the current density vector of the sources, ρ is the density of free electrical charge, γ is the conductivity of the material and Ω is the definition area of the model. The relationships between electrical and magnetic field intensities and densities are given by material relationships

$$\boldsymbol{D} = \varepsilon \boldsymbol{E}, \qquad \boldsymbol{B} = \mu \boldsymbol{H}. \tag{2}$$

The permittivity ε, the permeability μ and the conductivity γ in HFM are generally tensors with main axes in the direction of the Cartesian coordinates x, y, z. When all the field vectors perform rotation with the same angular frequency ω, it is possible to rewrite the first Maxwell equations

$$\nabla \times \underline{\boldsymbol{E}} = -j\omega\mu \underline{\boldsymbol{H}}, \ \nabla \times \underline{\boldsymbol{H}} = (\sigma + j\omega\varepsilon)\underline{\boldsymbol{E}} + \underline{\boldsymbol{J}_s} \text{ in } \Omega. \tag{3}$$

where $\underline{\boldsymbol{E}}$, $\underline{\boldsymbol{H}}$, $\underline{\boldsymbol{J}_s}$ are field complex vectors. Taking into account boundary conditions given in (1) and after rearranging (3) we get

$$(j\omega)^2 \varepsilon \underline{\boldsymbol{E}} + \sigma \underline{\boldsymbol{E}} + \nabla \times \mu^{-1} \nabla \times \underline{\boldsymbol{E}} = -j\omega \underline{\boldsymbol{J}}_s. \tag{4}$$

We apply Galerkin's method with vector approximation functions $\boldsymbol{W}_i$ and use vector form of the Green theorem on the double rotation element [3]. After discretisation we get the expression

$$-k_0[M]\{\boldsymbol{E}\} + jk_0[C]\{\boldsymbol{E}\} + [K]\{\boldsymbol{E}\} = \{\boldsymbol{F}\}, \tag{5}$$

where $\{\boldsymbol{E}\}$ is column matrix of electrical intensity complex vectors. The matrixes $[K]$, $[C]$ and $[M]$ are in the form that is given in manual [4] and vector $\{\boldsymbol{F}\}$ is evaluated from the expression

$$\{\boldsymbol{F}\} = -jk_0 Z_0 \int_{\Omega} [\boldsymbol{W}_i]\{\boldsymbol{J}_s\} d\Omega + jk_0 Z_0 \int_{\Gamma_0+\Gamma_1} [\boldsymbol{W}_i]\{\boldsymbol{n} \times \boldsymbol{H}\} d\Gamma. \tag{6}$$

Vector approximation functions $\boldsymbol{W}$ are given in manual [4]; k_0 is the wave number for vacuum, Z_0 is the impedance of free space. The set of equations (5) is independent of time and gives $\underline{\boldsymbol{E}}$. For transient vector $\boldsymbol{E}$ we can write

$$\boldsymbol{E} = \mathrm{Re}\{\underline{\boldsymbol{E}} e^{j\omega t}\}. \tag{7}$$

3.2. Model in FEM

The geometrical model was created with standard tools in ANSYS program using the automated generator of mesh and nodes; then the mathematical model is formed. The applied element is HF120.

Figure 3: Experimental construction of vircator, $P_{\max} = 250\,\mathrm{MW}$.

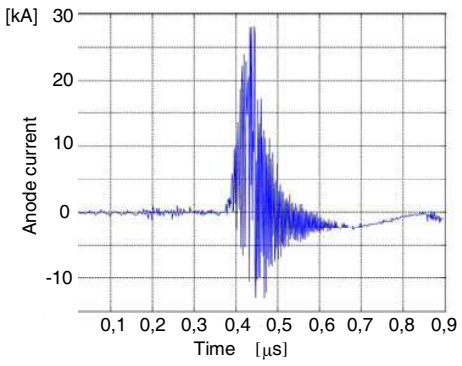


Figure 4: Waveform of vircator's anode current.

4. REALIZATION OF THE CALORIMETRIC SENSOR

Pursuant to the results obtained by numerical analysis, the calorimetric sensor was built. The prototype of the sensor was designed for the measurement of vircator with output power of $P_{\max} = 250\,\mathrm{MW}$, length of pulse $t_{\mathrm{p}} \in <10, 60> \mathrm{ns}$. Vircator is a pulse high-energy source of microwave energy based on the virtual cathode effect; its experimental construction is shown in Figure 3. Figure 4 shows the waveform of vircator's anode current by initiation. The concept was designed after consultation [7] for the supposed power and pulse length with room for absorption and damping of the possible back EMG wave.

4.1. Waveguide-fitted Calorimetric Sensor

The first prototype of the calorimetric sensor was intended for waveguide connection with a microwave vircator. Figure 5 shows the outer shell. The purpose of the calorimetric sensor was to measure the energy (power) of the emitted EMP. It was not possible to use the probe because the mode of the field, waveform and spectrum were not known.

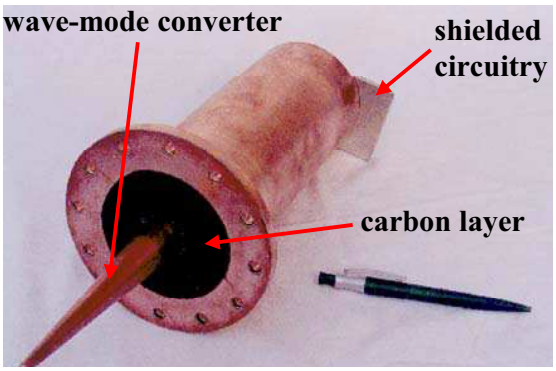


Figure 5: Waveguide-fitted calorimetric sensor.

Figure 6: Realization of optimized combined calorimetric sensor.

4.2. Free-space Combined Calorimetric Sensor

For the measurement of free-space vircator EMP, the new combined calorimetric sensor was built. The sensor operation is based on version I and version IV of the calorimetric method in Figure 2. The first part (version I) serves as the sensor of instantaneous power and the second part (version IV) serves as the sensor of pulse energy. The realization of the combined sensor is shown in Figure 6. Both parts are equipped with Horn antennas to ensure the matching of the free-space EMG wave to the sensor input.

The sensor was calibrated with an RF generator in an absorption room. The calibration was performed for microwave pulses with defined duration and power level.

Due to safety requirements, the connection between the sensor and the measuring device was ensured by means of coaxial cable of the minimum length $l_{\min} = 10\,\mathrm{m}$.

The combined calorimetric sensor was used for the measurement of vircator-emitted EMP. The supply of the vircator was provided by pulse high-voltage source powered by Marx bank. When the vircator is in the operational mode, hard RTG emission is generated in addition to the microwave emission. The energy of the electron beam is $W_{\mathrm{b}} = 1\,\mathrm{MeV}$. Therefore, safety requirements equal to those mentioned above have to be considered.

The waveform of the measured small microwave power is in Figure 7. The peak value of the vircator-emitted EMP reached $P_{\max} = 50\,\mathrm{kW}$ in this experiment.

However, vircator is able to emit EMP with peak value of hundreds of MW when supplied with pulsed power generators.

5. DESIGN OF THE MAGNETO-OPTIC METHOD

The magneto-optic (MO) method is proposed for further experiments. The magnetooptic method allows ultra-short pulses waveform measurement because of its high bandwidth. The polarization rotation of light passing the MO sensor is affected by the magnetic part of EM pulse. The rotation is due to the magnetic field and properties of the sensor material (Verdet constant). For free space measurement the MO garnet, glass or thin film may be used.

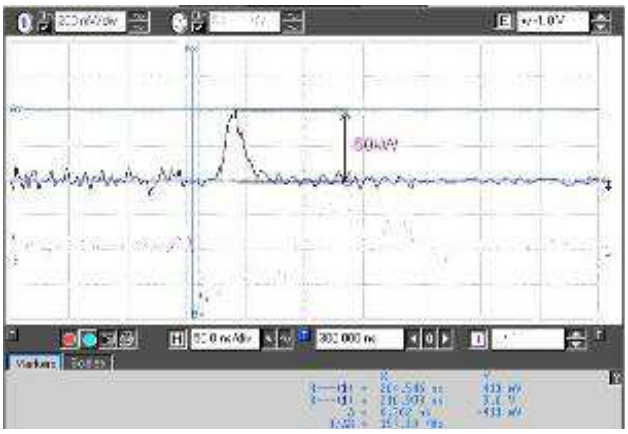


Figure 7: Measured waveform of small microwave power, $P_{\max} = 50\,\mathrm{kW}$.

The absolute measurement method utilizing the MO glass element was experimentally realized with low frequency magnetic field. Laser beam with linear polarization passes the MO glass placed in Helmholtz coil. The laser beam is subsequently fed through an analyzer and the polarization rotation is converted to intensity modulation. The intensity of light is sensed by a photodiode. The magneto-optic glass FR-5 by Hoyoa Optics was used in this experiment.

6. CONCLUSION

The overview of several methods suitable for the measurement of short solitary pulses with high power level was given. The characteristics of the designed method were discussed. Some methods were experimentally tested and evaluated. A combined calorimetric sensor for free-space measurement was built and the functionality of the calorimetric sensor was proved by real measurement of vircator-emitted EMP.

ACKNOWLEDGMENT

The paper was prepared within the framework of the research plan No. MSM 0021630513 of the Ministry of Education, Youth and Sports of the Czech Republic.

REFERENCES

1. Ott, H., *Noise Reduction Techniques in Electronic Systems*, John Wiley & Sons, 1988.
2. Booske, J. H., "Studies of non-thermal effects during intense microwave heating of crystalline solids," *Proceedings of the 3rd Symposium Microwave Processing of Materials*, 269, Pittsburgh, 1992.
3. Barker, R. J. and E. Schamiloglu, *High-power Microwave Sources and Technologies*, IEEE Press, 2001.
4. Fiala, P., *Non-conventional Sources of Electrical Energy*, BUT FEKT, Brno, 2003.
5. Kasal, M., *Multinuclear Measuring Channel of NMR Spectrometer*, FE BUT, Brno, 1984.
6. Pfeffer, M. and M. Kasal, "Automatic impulse reflectometer," *Sdělovací Technika*, Vol. 10–11, Czech Republic, 1986.
7. Šuka, P., Verbal information, UFP AV ČR, Praha 6, 2003.

A Simple Numerical Simulation of Internal Structure of Particles Test

K. Bartusek[1] and P. Fiala[2]

[1]Institute of Scientific Instruments, Academy of Sciences of the Czech Republic
Královopolská 147, 612 64, Brno, Czech Republic
[2]Department of Theoretical and Experimental Electrical Engineering
Brno University of Technology, Kolejni 4, 612 00 Brno, Czech Republic

Abstract— The article describes the test of a numerical model of the electron beam according to the present knowledge of references [1–5]. The basic configuration of the electron beam was experimentally verified at the Institute of Scientific Instruments Academy of Sciences of the Czech Republic. We prepared the numerical model which is based on the particle theory. In fact, the model respects the classical Electrodynamics Material Wave Theory (MWT). The numerical results were evaluated. The second model was prepared with respect to the theory of wave packet (Louis de Broglie) and solved again. The results of both models were the same in their quality; we evaluated electric field intensity $\boldsymbol{E}$ on the electron impact area, and they corresponded with the results from experiments.

1. INTRODUCTION

One of the problems solved at the Department of Theoretical and Experimental Electrical Engineering is the determination of water molecule properties exposed to the electric field. Nuclear magnetic resonance (NMR) is a promising experimental method for this purpose. The disadvantage of this method is the presence of a strong magnetic field. It is not easy to analyze the influence of the electric field, because of the above mentioned disadvantage.

The numerical model can serve as a means of solution to this problem. First, we can perform the simulation of the magnetic field influence and, consequently, determine the influence of the electric field. Next, we can carry out an experiment using the NMR spectrometer and compare the experimental results with the results of the simulation. This approach allows us to determine the influence of the electric field on the water molecules. If we want to build a suitable numerical model, we have to decide how to interpret the particle theory and the material theory (MT).

We based the numerical models on two of the known material theory approaches — the particle theory (PT) and the material wave theory (MWT). The numerical models were verified on a simple example and we started to define the relation between these two models. The finite element method (FEM) was used for the model solution. The results obtained by the comparison of the PT and MWT models are presented in this paper. The material wave theory phenomena [1–4] can be demonstrated and verified on the simple example described below. The principles of the material wave theory are presented in [3].

2. PHYSICAL MODEL

Numerical models based on the particle theory and the material wave theory were built. On the basis of these, the test example was built. The example consists of a pair of electrodes for electron microscope testing [6]. The test example is shown in Fig. 1.

In Fig. 1, the electrode system is supplied by a high voltage source $U(t)$. Electrons exiting the A electrode are moving toward the A' electrode. The A' electrode collects the electrons. The electrode system B–B', supplied by the u_c source, allows the correction of the electron trajectory and the impact point. This test example was solved by means of the finite element method, utilizing ANSYS system and APDL algorithms. The first of them was designed in accordance with the classical macroscopic theory. The results were obtained by the solution of partial differential equations formulated on the basis of reduced Maxwell equations [5]. The solution of the mass point dynamics has utilized the coupled electrostatic model and the Lorentz force law:

$$\boldsymbol{F} = q\boldsymbol{E} + q\boldsymbol{v} \times \boldsymbol{B} \quad \text{in } \Omega, \tag{1}$$

where $\boldsymbol{F}$ is the vector of the resultant force which acts on the particle, q is the particle charge, $\boldsymbol{E}$ is the external electric field intensity vector, $\boldsymbol{v}$ is the speed of the moving particle, $\boldsymbol{B}$ is the

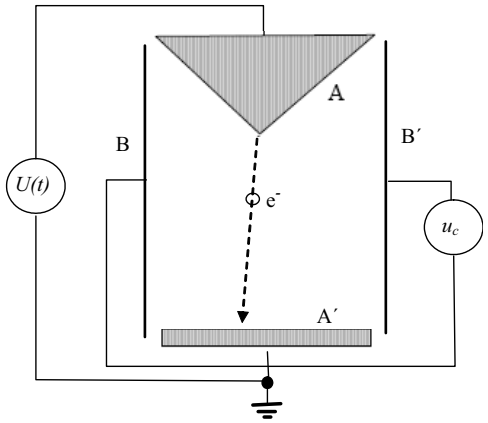


Figure 1: An example setup for the MWT test.

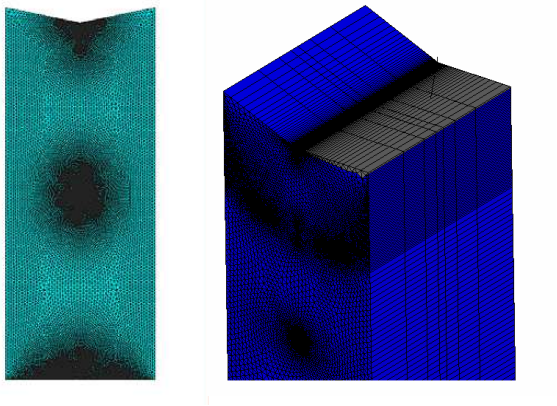


Figure 2: The FEM mesh for moving particle model.

magnetic flux density vector of the external magnetic field, Ω is the definition area of the model. The obtained results were experimentally verified at the Institute of Scientific Instruments of the Academy of Sciences of the Czech Republic.

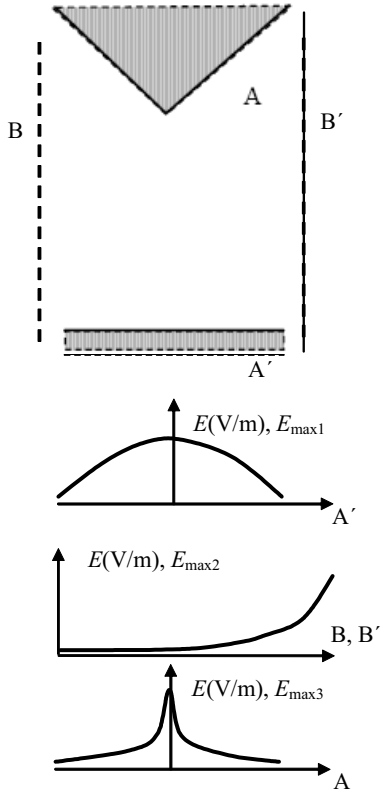


Figure 3: Boundary conditions obtained by the solution of the PT model.

Figure 3 shows the result of the numerical model of the electrode system test from Fig. 1. The simulation results correspond with the results obtained by the experiments of the Institute of Scientific Instruments. The parameters of the experiment were as follows: $U(t) = U_{max} \cdot f(x)$, where $U_{max} = 10\,\mathrm{kV}$, $u_c = 10\,\mathrm{V}$, the distance of the B–B' electrode $d_{BB'} = 20\,\mathrm{mm}$, the distance of the A–A' electrode $d_{AA'} = 50\,\mathrm{mm}$. We can evaluate the count of electrons collected by the A' electrode as the space charge q_d with spatial distribution

$$\boldsymbol{E} = \frac{1}{4\pi\varepsilon} \int_S \frac{\sigma}{|r|^2} d\boldsymbol{S} \quad \text{in} \quad \Gamma_{A'}, \tag{2}$$

where ε is the vacuum permittivity, σ is the surface charge density, $\boldsymbol{r}$ is the position vector, $\boldsymbol{S}$ is the oriented surface. After the evaluation (Fig. 4), Fig. 5 shows the module of the electric field intensity vector on the $\Gamma_{A'}$ boundary. The results of the particle theory model were compared with the results of the MWT model. The MWT model utilizes the wave equation. The wave equation is formulated with the help of Maxwell equations in the partial differential form [5]. Boundary conditions were obtained by the PT model solution. The boundary conditions are represented by intensities of the electric field on the electrode system surface in Fig. 3. Here, $E_{\max 1} = 670\,\mathrm{V/m}$, $E_{\max 2} = 2.9 \cdot 10^5\,\mathrm{V/m}$, $E_{\max 3} = 2.1 \cdot 10^7\,\mathrm{V/m}$.

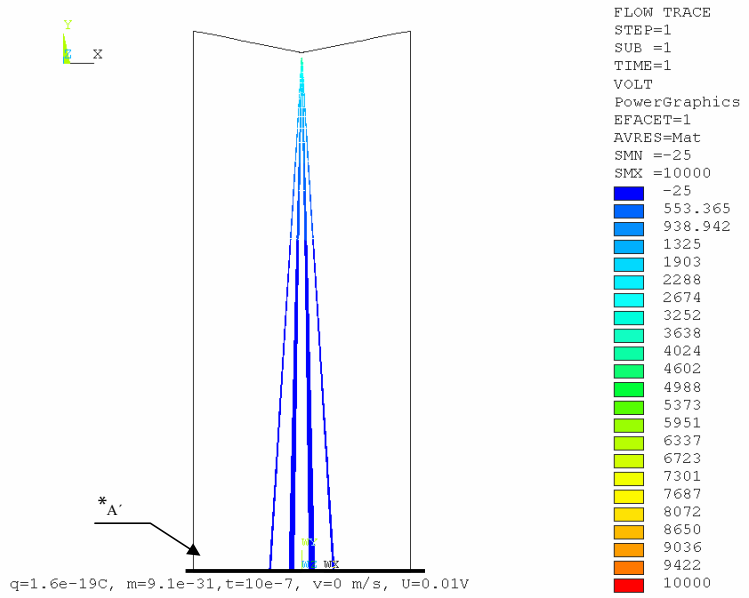


Figure 4: The resultant trajectory of moving electrons in the electrode system C particle theory model.

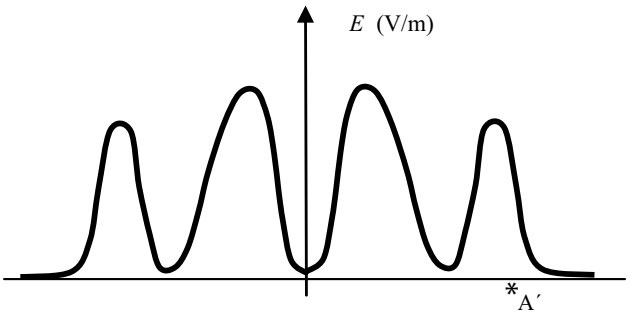


Figure 5: The electric field intensity $\boldsymbol{E}$ module — particle theory model.

The model of the source of the electromagnetic field was simplified:

$$\boldsymbol{E}(t) = \boldsymbol{E}(x, y, z) \cdot \sin(g(t)) \quad \text{in } \Gamma_A, \tag{3}$$

where $\boldsymbol{E}(x, y, z)$ is the vector of the maximal electric field intensity on the A electrode, Γ_A is the surface of the electrode, $\boldsymbol{E}(x, y, z) = \boldsymbol{E}_{\max} \cdot h(x, y, z)$, $\boldsymbol{E}_{\max} = 670\,\mathrm{V/m}$, $g(t)$ is the exciting function on the surface A, for simplicity $g(t) = \sin(2\pi f)$ and $f = 3 \cdot 10^{10}\,\mathrm{Hz}$. This simplification was considered because of the reasonable computing time. In accordance with the MWT the appropriate electron frequency should be $f = 3 \cdot 10^{27}\,\mathrm{Hz}$

$$f = \frac{c}{\lambda} \tag{4}$$

where c is the speed of light, λ represents the electron size ($\lambda = 1 \cdot 10^{-19}\,\mathrm{m}$). The solution with valid electron size is currently verified.

3. DISCUSSION

The performed experiments show that it is possible to substitute the macroscopic, microscopic material model (described with the particle theory) with the material wave theory. The powerful tool for basic material models was discovered [1–4]. It is possible to explain many experiments with the help of this new approach.

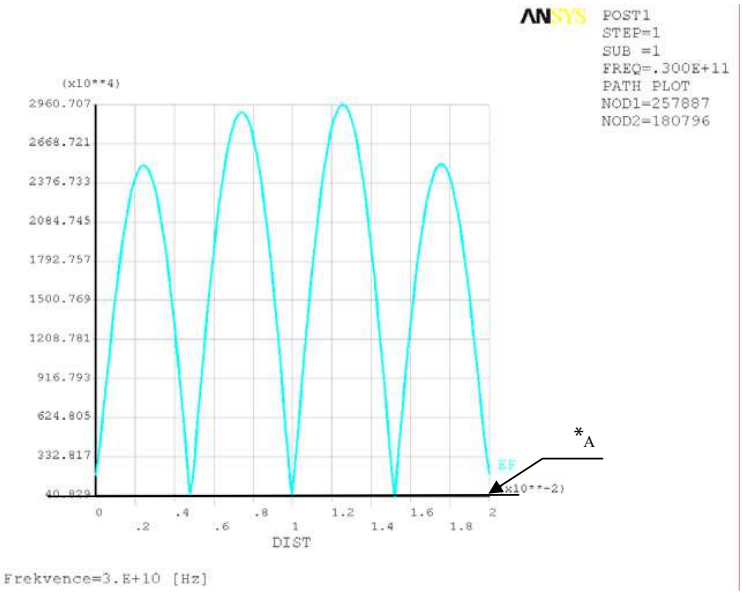


Figure 6: The electric field intensity $\boldsymbol{E}$ module — MWT model.

4. CONCLUSION

On the basis of the experiments mentioned above, the model of basic particles will be built. The next point is to set up the model of basic elements. This activity is connected with the research plan FRVS CR No. MSM 0021630513 ELCOM, No. MSM 0021630516. The aim of this research is to describe the water molecule behaviour in the electric field, with know intensity $\boldsymbol{E}$. The behaviour is experimentally tested on the NMR spectrometer.

ACKNOWLEDGMENT

The research described in the paper was financially supported by the FRVŠ research plan No. MSM 0021630513 ELCOM, No. MSM 0021630516 and grant GAAV No. B208130603.

REFERENCES

1. Van Vlaenderen, K. J. and A. Waser, "Electrodynamics with the scalar field," *Physics*, Vol. 2, 1–13, 2001.
2. Kikuchir, H., "Electrohydrodynamics in dusty and dirty plasmas, gravito-electrodynamics and EHD," Kluwer Academic Publishers, Dordrecht/Boston/London, 2001.
3. Van Vlaenderen, K. J., "A charge space as the orogin of sources, fields and potentials," *Physics*, arXiv:physics/9910022 v1, Vol. 16, 1–13, Oct. 1999.
4. Hofer, W. A., "A charge space as the origin of sources, fields and potentials," *Physics*, arXiv: quant-ph/ 9611009 v3, Vol. 17, 1–13, Apr. 1997.
5. Prosser, V a kolektiv, "Experimentální metody biofyziky," *Academia*, Praha, 1989.
6. Delong, A., "Verbal information," *Czech Academy of Science, ISI Brno*, Vol. 7, No. 2, Brno, 2006.

Numerical Model of Optimization of the Lead-acid Accumulator Grids

E. Kadlecova, **I. Behunek**, **and P. Fiala**

Department of Theoretical and Experimental Electrical Engineering
Faculty of Electrical Engineering and Communication, Brno University of Technology
Kolejni 2906/4, 612 00 Brno, Czech Republic

Abstract— This article deals with chemical processes during charging and discharging of the lead-acid accumulator. There is presented one shape of electrode grids in this work, but we researched more variants. We prepared numerical models based on the combined finite element method (FEM) and finite volume method (FVM) of those variants and computed current density distribution on the surface of electrodes. The model joins magnetic, electric and current fields, flow field and chemical nonlinear ion model. Results were obtained by means of the FEM/FVM as a main application in ANSYS software.

1. INTRODUCTION

The lead-acid battery uses lead dioxide as the active material of the positive electrode and metallic lead, in a high-surface-area porous structure, as the negative active material. Typically, a charged positive electrode contains both α-PbO_2 (orthorhombic) and β-PbO_2 (tetragonal). The equilibrium potential of the α-PbO_2 is more positive than that of β-PbO_2 by 0.01 V The α form also has a larger, more compact crystal morphology which is less active electrochemically and slightly lower in capacity per unit weight; it does, however, promote longer cycle life. Neither of the two forms is fully stoichiometric. Their composition can be represented by PbO_x, with x varying between 1.85 and 2.05. The introduction of antimony, even at low concentrations, in the preparation or cycling of these species leads to a considerable increase in their performance. The preparation of the active material precursor consists of a series of mixing and curing operations using leady lead oxide (PbO + Pb), sulfuric acid, and water. The ratios of the reactants and curing conditions (temperature, humidity, and time) affect the development of crystallinity and pore structure. The cured plate consists of lead sulfate, lead oxide, and some residual lead ($< 5\%$). The positive active material, which is formed electrochemically from the cured plate, is a major factor influencing the performance and life of the lead-acid battery. In general the negative, or lead, electrode controls cold-temperature performance (such as engine starting). The electrolyte is a sulfuric acid solution, typically about 1.28 specific gravity or 37% acid by weight in a fully charged condition. On the grids there are following reactions.

Negative Electrode

$$Pb \overset{\text{discharging}}{\underset{\text{charging}}{\Leftrightarrow}} Pb^{2+} + 2e^- \tag{1}$$

$$Pb^{2+} + SO_2^{2-} \overset{\text{discharging}}{\underset{\text{charging}}{\Leftrightarrow}} PbSO_4 \tag{2}$$

Positive Electrode

$$PbO_2 + 4H^+ + 2e^- \overset{\text{discharging}}{\underset{\text{charging}}{\Leftrightarrow}} Pb^{2+} + 2H_2O \tag{3}$$

$$Pb^{2+} + SO_4^{2-} \overset{\text{discharging}}{\underset{\text{charging}}{\Leftrightarrow}} PbSO_4 \tag{4}$$

Overall Reactions

$$Pb + PbO_2 + 2H_2SO_4 \overset{\text{discharging}}{\underset{\text{charging}}{\Leftrightarrow}} 2PbSO_4 + 2H_2O \tag{5}$$

2. MATHEMATICAL AND NUMERICAL MODEL

Electromagnetic part is derived from reduced Maxwell equations

$$\text{rot}\, \boldsymbol{H} = 0, \tag{6}$$
$$\text{div}\, \boldsymbol{B} = 0, \tag{7}$$

where $\boldsymbol{H}$ is the vector of magnetic field intensity, $\boldsymbol{B}$ is the magnetic field induction, $\boldsymbol{J}$ is vector of current density.

$$\text{rot}\, \boldsymbol{E} = 0, \tag{8}$$
$$\text{div}\, \boldsymbol{J} = 0, \tag{9}$$

where $\boldsymbol{E}$ is the vector of electric field intensity. Vector functions of electric, magnetic field are expressed by means of a scalar potentials ϕ_e, ϕ_m. Final current density from (9) $\boldsymbol{J}$ is influenced by velocity $\boldsymbol{v}$ of the flowing ion solution and the outer magnetic field

$$\boldsymbol{J} = \gamma(\boldsymbol{E} + \boldsymbol{v} \times \boldsymbol{B}). \tag{10}$$

If electrodes E_1 and E_2 have different electric potentials, then the current density $\boldsymbol{J}$ is created in the Ω area according to (10) and the current I_L flows in the ion solution

$$I_L = \iint\limits_{S_e} \boldsymbol{J} \cdot d\boldsymbol{S} = \iint\limits_{S_e} \gamma(\boldsymbol{E} + \boldsymbol{v} \times \boldsymbol{B}) \cdot d\boldsymbol{S}. \tag{11}$$

where $\boldsymbol{S}_e$ is a directed area of electrodes E_1 and E_2 into space Ω. In Eq. (11) there is the electric field intensity $\boldsymbol{E}$ for ion solution much smaller than product of $\boldsymbol{v} \times \boldsymbol{B}$, so we ignore the influence of electric field intensity. The specific force $\boldsymbol{f}$ affects the moving charge q and the force $\boldsymbol{F}$ in the whole Ω area is

$$\boldsymbol{F} = \iiint\limits_{\Omega} \boldsymbol{J} \times \boldsymbol{B} dV. \tag{12}$$

We obtain the voltage between electrodes E_1, E_2 from

$$U_L = \int\limits_{E_1}^{E_2} \frac{\boldsymbol{F}}{q} \cdot d\ell. \tag{13}$$

where the electric field intensity is derived from the force $\boldsymbol{F}$ which affects a charge q. After modification voltage on electrodes is

$$U_L = \iiint\limits_{\Omega} \left(\frac{\boldsymbol{J}(\mathbf{v})}{I_L} \times \boldsymbol{B} \right) \cdot (\boldsymbol{v}) dV. \tag{14}$$

The current density $\boldsymbol{J}(\boldsymbol{v})$ depends on immediate ion velocity between E_1 and E_2.

The magnetic field, which is expressed in (14) by the induction B, we gain from Biot-Savart law by means of different scalar magnetic potentials (DSP)

$$\boldsymbol{T} = \frac{1}{4\pi} \int\limits_{\Omega} \frac{\boldsymbol{J}_c \times \boldsymbol{R}}{|\boldsymbol{R}|^3} dV, \tag{15}$$

where $\boldsymbol{R}$ is a distance between a point where we look for the magnetic field intensity $\boldsymbol{T}$ and a point where magnetic field source is with the current density $\boldsymbol{J}_c$. We can write the magnetic field intensity $\boldsymbol{H}$ in the area as

$$\boldsymbol{H} = \boldsymbol{T} - \text{grad}\phi_m, \tag{16}$$

where $\boldsymbol{T}$ is the previous or estimated magnetic field intensity. Boundary conditions are

$$\boldsymbol{n} \cdot \mu_0 \mu_r (\boldsymbol{T} - \text{grad}\phi_m) = 0 \quad \text{on the boundary } \Gamma_{\text{Pb}-0}, \tag{17}$$

where $\boldsymbol{n}$ is the normal vector. $\Gamma_{\mathrm{Pb}-0}$ is the interface between areas Ω_{Pb} and Ω. The continuity of tangential components of the magnetic field intensity on the area interface with a ferromagnetic material is

$$\boldsymbol{n} \times (\boldsymbol{T} - \mathrm{grad}\phi_m) = 0 \quad \text{on the boundary } \Gamma_{\mathrm{Pb}-0}. \tag{18}$$

By the help of relations (6), (7) we get

$$\mathrm{div}\mu_0\mu_r \boldsymbol{T} - \mathrm{div}\mu_0\mu_r \mathrm{grad}\phi_m = 0. \tag{19}$$

We get the discretization of equation (19) by means of the approximation of the scalar magnetic potential

$$\phi_m = \sum_{k=1}^{N_\phi} \varphi_{mk} W_k(x, y, z), \quad \forall (x, y, z) \subset \Omega, \tag{20}$$

where φ_m is the nodal value of the scalar magnetic potential, W is the base function, N_φ is number of mesh nodes. We obtain the semidiscrete solution by means of the approximation (20) in the relation (19) and Galerkin method

$$\sum_{j=1}^{N_\phi} - \int_\Omega \mu \boldsymbol{t}_j \cdot \mathrm{grad}W_i + \mu \mathrm{grad}\varphi_{mj} \cdot \mathrm{grad}W_i d\Omega = 0, i = 1, \ldots, N_\varphi, \tag{21}$$

where $\boldsymbol{t}_j$ is the nodal value of the known magnetic field intensity. Equation (20) may be written briefly as

$$-[K_{Tij}] + [k_{ij}]\{\varphi\} = 0 \; i, j = 1, \ldots, N_\varphi. \tag{22}$$

Coefficients for (22) are written as

$$\begin{aligned} k_{Tij}^{em} &= -\int_{\Omega^e} \mu^e \mathrm{t}_j \cdot \mathrm{grad}W_i d\Omega \; i, j = 1, \ldots, N_e \\ k_{ij}^{em} &= -\int_{\Omega^e} \mu^e \mathrm{grad}\varphi_j \cdot \mathrm{grad}W_i d\Omega \end{aligned} \tag{23}$$

where Ω^e is the area of the selected element of mesh, μ^e is the material permeability of the selected element, N_e is the number of elements of mesh. The equation system changes into relation

$$-\left[k_{Tij}^{em}\right] + \left[k_{ij}^{em}\right]\{\varphi\} = 0 \quad e = 1, \ldots, N_e. \tag{24}$$

We can solve the equation system (24) by means of standard algorithms. The solution by means of DSP consists of two parts. Firstly, we express the distribution of magnetic field intensity $\boldsymbol{T}$ from current sources according to (16) with the respect to boundary conditions (17) and (18) in the area Ω_{Pb}. Secondly, we have to find out the solution of the magnetic intensity $\boldsymbol{H}$ distribution according to (16) from the previous step.

The **model of the electrical** or **current** field is formulated from previous equations

$$\gamma \mathrm{div}\, \mathrm{grad}\phi_e = 0. \tag{25}$$

On the interface there are conditions

$$\boldsymbol{n} \cdot \gamma(\mathrm{grad}\phi_e) = 0 \quad \text{on the boundary } \Gamma_{\mathrm{E}-\mathrm{k}}, \tag{26}$$

where $\boldsymbol{n}$ in the normal vector to the surface of the electrode. $\Gamma_{\mathrm{E}-\mathrm{k}}$ is the interface between the liquid and the electrode. The continuity of tangential components of electric field intensity on the interface is

$$\boldsymbol{n} \times (\mathrm{grad}\phi_e) = 0 \quad \text{on the boundary } \Gamma_{\mathrm{E}-\mathrm{k}}. \tag{27}$$

We can do the approximation of the scalar electric potential in the similarly way like in the relation (20) and by means of (20), (25) and the Galerkin method we get the semidiscrete solution

$$\sum_{j=1}^{N_\phi} - \int_\Omega \gamma \mathrm{grad}\varphi_{ej} \cdot \mathrm{grad}W_i d\Omega = 0, i = 1, \ldots, N_\varphi, \tag{28}$$

where φ_{ej} is the nodal value of the scalar electric potential. We can rewrite the equation system (28) by the help of

$$\left[k_{ij}^{J}\right]\{\varphi\} = 0 \; i,j = 1,\ldots,N_{\varphi}. \tag{29}$$

$$k_{ij}^{Je} = -\int_{\Omega^e} \gamma^e \mathrm{grad}\varphi_{ej} \cdot \mathrm{grad}W_i d\Omega, \tag{30}$$

where Ω^e is the area of the selected element of mesh, γ^e is the specific conductivity of liquid in the static state of the selected element, N_e is the number of elements. The relation for the voltage drop during discharging $U(t) = U_0 - \Delta u$ is

$$\Delta u = \sqrt{Z\left(\int_{\Omega}\frac{\boldsymbol{J}_e^- q_e^- \cdot \boldsymbol{v}_e^-}{\Delta V_e \gamma_e}\right)d\Omega + Z\int_{\Omega}(\boldsymbol{J}_e \times \boldsymbol{B}_e)\cdot \boldsymbol{v}_e^- d\Omega + Z\left(\int_{\Omega}\frac{\boldsymbol{J}_e^+ q_e^+ \cdot \boldsymbol{v}_e^+}{\Delta V_e \gamma_e}\right)d\Omega + Z\int_{\Omega}(\boldsymbol{J}_e \times \boldsymbol{B}_e)\cdot \boldsymbol{v}_e^+ d\Omega} \tag{31}$$

where

$$v_e^+ = \frac{E_e\gamma}{F_c \Delta V_e \sum\limits_{k=1}^{N_{\mathrm{ion}+}} c_k^+ N_{\mathrm{k}}^{+\mathrm{ion}}}, \quad v_e^- = \frac{E_e\gamma}{F_c \Delta V_e \sum\limits_{k=1}^{N_{\mathrm{ion}-}} c_k^- N_{\mathrm{k}}^{-\mathrm{ion}}},$$

$$q_e^+ = F_c \Delta V_e \sum_{k=1}^{N_{\mathrm{ion}+}} c_k^+ N_{\mathrm{k}}^{+\mathrm{ion}}, \quad q_e^- = F_c \Delta V_e \sum_{k=1}^{N_{\mathrm{ion}-}} c_k^- N_{\mathrm{k}}^{-\mathrm{ion}} \tag{32}$$

where F_c is the Faraday constant, $F_c = 96484$ C.mol^{-1}, $\boldsymbol{E}_e$ the electric field intensity in direction of ions motion in an element of mesh, c^+ the positive ions concentration, c^- the negative ions concentration, ΔV_e is the element volume, $N_{\mathrm{k}}^{+\mathrm{ion}}$ is the integer multiple of electron charge for specific positive ion, $N_{\mathrm{k}}^{-\mathrm{ion}}$ is the integer multiple of electron charge for specific negative ion, q_e^- is the whole charge of negative ions in one element, q_e^+ is the total charge of positive ions in one element, $N_{\mathrm{ion}+}$ is the number of different positive charge carriers (elements, compounds), $N_{\mathrm{ion}-}$ is the number of negative charge carriers. There are concentrations c_{k} in fluid

Positive ions	**Negative ions**
$H^+ \ldots$ 9,9193.10^2 mol.m^{-3}	$SO_4^{2-} \ldots$ 4,9597.10^3 mol.m^{-3}
$Pb^{4+} \ldots$ 2,2028.10^4 mol.m^{-3}	$OH^- \ldots$ 8,8110.10^4 mol.m^{-3}

3. NUMERICAL SOLUTION FEM/FVM

The numerical model was prepared by means of ANSYS tools [1] and main FEM/FVM solution was solved with APDL program over ANSYS system. In Fig. 1, we can see distribution of current density module J on the grid surface and in the electrolyte between positive and negative electrodes. For computation we used initial voltage was 2.3 V, minimal voltage of discharged accumulator, supposed discharging time, and time period between steps of computation.

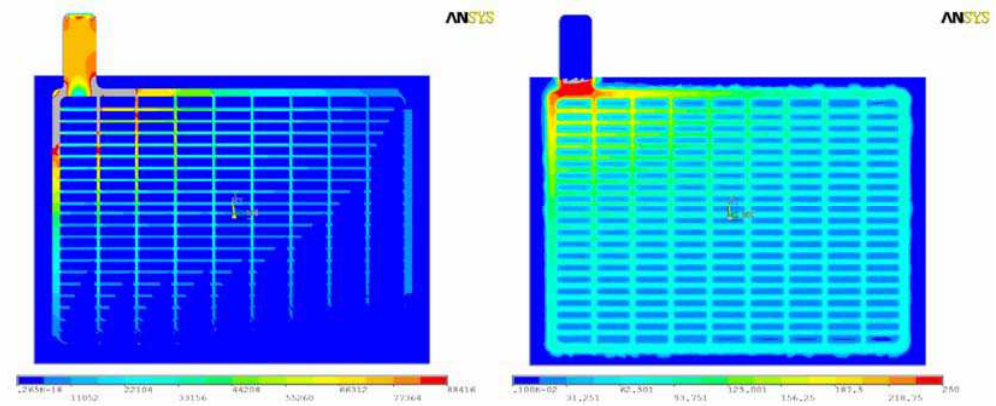


Figure 1: Current density distribution on the grid surface (left) and in the electrolyte (right), charged state.

4. CONCLUSION

This work deals with chemical processes on the electrodes and in an electrolyte of a lead-acid accumulator. There is mathematical and numerical description in the article and results of current density distribution in the electrodes area. Such analysis could be used for an optimization of grid geometry. The best shape we get if module of current density is constant on the electrode surface.

ACKNOWLEDGMENT

The paper was prepared within the framework of the research plan No. MSM 0021630516 of the Ministry of Education, Youth and Sports of the Czech Republic.

REFERENCES

1. *Ansys User's Manual*, Svanson Analysys System, Inc., Huston, USA, 2003.
2. Salkind, A. J., A. G. Cannone, and F. A. Trumbure, *Handbook of Batteries*, CRC Press, Inc., Boca Raton, Florida, USA, 2001.
3. Burian, I., L. Dědek, and J. Dědková, "Analysis and optimization of an inductive flowmeter," *8th International Symposium on Theoretical Electrical Engineering*, 506–509, Thessaloniki, September 22–23, 1995.
4. Fiala, P., "Model of induction flowmeter DN100," Research report No. 2/01, Laboratoř modelování a optimalizace polí v elektromechanických systémech. FEI VUT BRNO, Brno, Czech Republic, 21, 6, 2001.