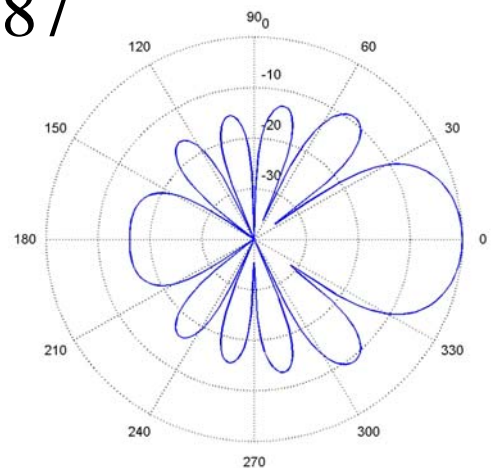
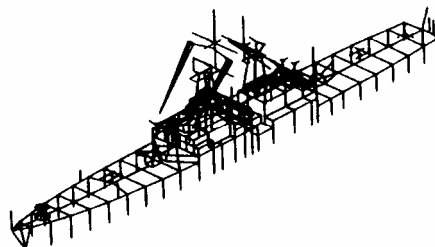
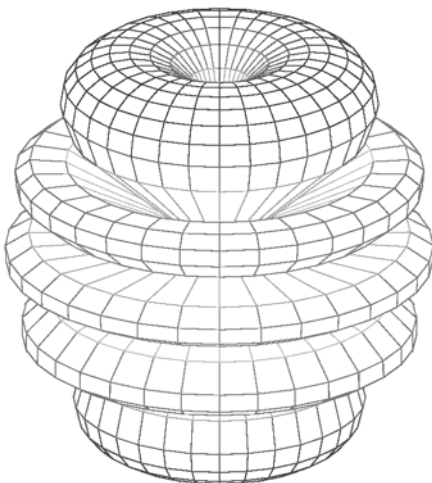
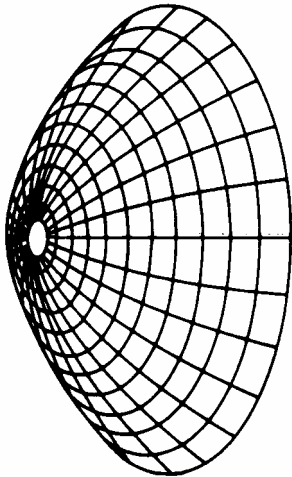
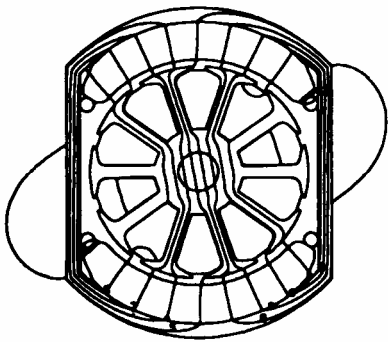
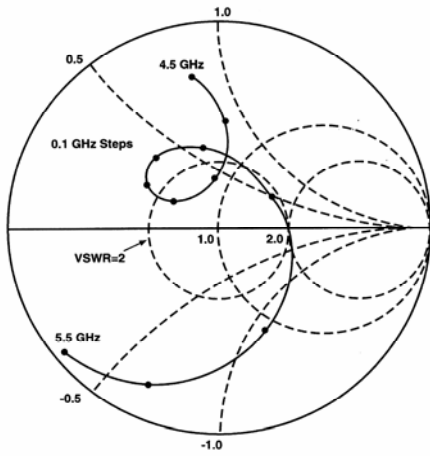


# Applied Computational Electromagnetics Society Journal

Editor-in-Chief  
**Atef Z. Elsherbeni**

September 2008  
Vol. 23 No. 3  
ISSN 1054-4887



**GENERAL PURPOSE AND SCOPE:** The Applied Computational Electromagnetics Society (*ACES*) Journal hereinafter known as the *ACES Journal* is devoted to the exchange of information in computational electromagnetics, to the advancement of the state-of-the art, and the promotion of related technical activities. A primary objective of the information exchange is the elimination of the need to “re-invent the wheel” to solve a previously-solved computational problem in electrical engineering, physics, or related fields of study. The technical activities promoted by this publication include code validation, performance analysis, and input/output standardization; code or technique optimization and error minimization; innovations in solution technique or in data input/output; identification of new applications for electromagnetics modeling codes and techniques; integration of computational electromagnetics techniques with new computer architectures; and correlation of computational parameters with physical mechanisms.

**SUBMISSIONS:** The *ACES Journal* welcomes original, previously unpublished papers, relating to applied computational electromagnetics. Typical papers will represent the computational electromagnetics aspects of research in electrical engineering, physics, or related disciplines. However, papers which represent research in applied computational electromagnetics itself are equally acceptable.

Manuscripts are to be submitted through the upload system of *ACES* web site <http://aces.ee.olemiss.edu> See “Information for Authors” on inside of back cover and at *ACES* web site. For additional information contact the Editor-in-Chief:

**Dr. Atef Elsherbeni**

Department of Electrical Engineering  
The University of Mississippi  
University, MS 386377 USA  
Phone: 662-915-5382 Fax: 662-915-7231  
Email: [atef@olemiss.edu](mailto:atef@olemiss.edu)

**SUBSCRIPTIONS:** All members of the Applied Computational Electromagnetics Society who have paid their subscription fees are entitled to receive the *ACES Journal* with a minimum of three issues per calendar year and are entitled to download any published journal article available at <http://aces.ee.olemiss.edu>.

**Back issues**, when available, are \$15 each. Subscriptions to *ACES* is through the web site. Orders for back issues of the *ACES Journal* and changes of addresses should be sent directly to *ACES*:

**Dr. Allen W. Glisson**

302 Anderson Hall  
Dept. of Electrical Engineering  
Fax: 662-915-7231  
Email: [aglisson@olemiss.edu](mailto:aglisson@olemiss.edu)

Allow four week’s advance notice for change of address. Claims for missing issues will not be honored because of insufficient notice or address change or loss in mail unless the Executive Officer is notified within 60 days for USA and Canadian subscribers or 90 days for subscribers in other countries, from the last day of the month of publication. For information regarding reprints of individual papers or other materials, see “Information for Authors”.

**LIABILITY.** Neither *ACES*, nor the *ACES Journal* editors, are responsible for any consequence of misinformation or claims, express or implied, in any published material in an *ACES Journal* issue. This also applies to advertising, for which only camera-ready copies are accepted. Authors are responsible for information contained in their papers. If any material submitted for publication includes material which has already been published elsewhere, it is the author’s responsibility to obtain written permission to reproduce such material.

# **APPLIED COMPUTATIONAL ELECTROMAGNETICS SOCIETY JOURNAL**

Editor-in-Chief  
**Atef Z. Elsherbeni**

September 2008  
Vol. 23 No. 3  
ISSN 1054-4887

**The ACES Journal is abstracted in INSPEC, in Engineering Index, DTIC, Science Citation Index Expanded, the Research Alert, and to Current Contents/Engineering, Computing & Technology.**

The first, fourth, and sixth illustrations on the front cover have been obtained from the Department of Electrical Engineering at the University of Mississippi.

The third and fifth illustrations on the front cover have been obtained from Lawrence Livermore National Laboratory.

The second illustration on the front cover has been obtained from FLUX2D software, CEDRAT S.S. France, MAGSOFT Corporation, New York.

**THE APPLIED COMPUTATIONAL ELECTROMAGNETICS SOCIETY**  
<http://aces.ee.olemiss.edu>

**ACES JOURNAL EDITOR-IN-CHIEF**

**Atef Elsherbeni**  
University of Mississippi, EE Dept.  
University, MS 38677, USA

**ACES JOURNAL ASSOCIATE EDITORS-IN-CHIEF**

**Sami Barmada**  
University of Pisa, EE Dept.  
Pisa, Italy, 56126

**Erdem Topsakal**  
Mississippi State University, EE Dept.  
Mississippi State, MS 39762, USA

**Fan Yang**  
University of Mississippi, EE Dept.  
University, MS 38677, USA

**ACES JOURNAL EDITORIAL ASSISTANTS**

**Matthew J. Inman**  
University of Mississippi, EE Dept.  
University, MS 38677, USA

**Mohamed Al Sharkawy**  
Arab Academy for Science and Technology  
ECE Dept.  
Alexandria, Egypt

**ACES JOURNAL EMERITUS EDITORS-IN-CHIEF**

**Duncan C. Baker**  
EE Dept. U. of Pretoria  
0002 Pretoria, South Africa

**Allen Glisson**  
University of Mississippi, EE Dept.  
University, MS 38677, USA

**David E. Stein**  
USAF Scientific Advisory Board  
Washington, DC 20330, USA

**Robert M. Bevensee**  
Box 812  
Alamo, CA 94507-0516, USA

**Ahmed Kishk**  
University of Mississippi, EE Dept.  
University, MS 38677, USA

**ACES JOURNAL EMERITUS ASSOCIATE EDITORS-IN-CHIEF**

**Alexander Yakovlev**  
University of Mississippi, EE Dept.  
University, MS 38677, USA

**SEPTEMBER 2008 REVIEWERS**

Mohamed Al-Sharkawy  
Francisco Jose Ares  
Ercument Arvas  
Constantine A. Balanis  
Mohamed H. Bakr  
Sami Barmada  
Malgorzata Celuch  
Deb Chatterjee  
David Chen

Magda Elshenawee  
Richard K. Gordon  
Elliott Hutchcraft  
Apisak Ittipiboon  
Yasushi Kanai  
Michiko Kuroda  
De Leeneer  
Kai Man Luk  
Richard Martell

Natalia K. Nikolova  
Antonio Orlandi  
Michele Piana  
Chris J. Railton  
C. J. Reddy  
Alan Taflove  
Christopher W. Trueman  
Fan Yang

**THE APPLIED COMPUTATIONAL ELECTROMAGNETICS SOCIETY**  
**JOURNAL**

Vol. 23 No. 3

September 2008

**TABLE OF CONTENTS**

“A Review of Perfectly Matched Absorbers for the Finite-Volume Time-Domain Method” T. Kaufmann, K. Sankaran, C. Fumeaux, and R. Vahldieck.....	184
“New Stability Criterion for Unstructured Mesh Upwinding FVTD Schemes for Maxwell’s Equations” D. Firsov and J. LoVetri.....	193
“A General Framework for Mixed Structured/Unstructured PEEC Modelling” F. Freschi and M. Repetto.....	200
“Preliminary Investigation of the NCP Parameter-Choice Method for Inverse Scattering Problems Using BIM: 2-D TM Case” Puyan Mojabi and Joe LoVetri.....	207
“Finite Element Modeling of Dual-Core Photonic Crystal Fiber” K. R. Khan and T. X. Wu.....	215
“A Parallel Numerical Method to Solve High Frequency Ghost Obstacle Acoustic Scattering Problems” L. Fatone, M. C. Recchioni, and F. Zirilli.....	220
“Analysis of a Singly-Fed Circularly Polarized Electromagnetically Coupled Patch Antenna” A. Hajiaboli and M. Popovic.....	233
“Stochastic Optimization of a Patch Antenna” S. Alfonzetti, G. Borzì, E. Diletto, and N. Salerno.....	237
“A New Broadband Microstrip Leaky-Wave Antenna” O. Losito.....	243
“Two-Element T-Array for Cross-Polarized Breast Tumor Detection” H. Kanj and M. Popovic.....	249
“Finite Array Analysis Through Combination of Macro Basis Functions and Array Scanning Methods” C. Craeye and R. Sarkis.....	255

“Robust Adaptive Beamforming Using Least Mean Mixed Norm Algorithm” R. M. Shubair, S. A. Jimaa, and A. A. Omar.....	262
“An Accurate Reduced-Order Polynomial Solution for Root-MUSIC Source Localization Using Displaced Sensor Arrays” R. M. Shubair.....	270
“The Design of a TRL Calibration Kit for Microstrip and its use for Measurement and Modeling of Active and Passive RF Components” D. Elsherbeni, L. Jordan, E. Hutchcraft, D. Kajfez, and R. K. Gordon.....	276
“A Discrete Random Medium Model for Electromagnetic Wave Interactions with Sea Spray” O. Kilic.....	286
“Abstraction of Graphics Hardware Through The Use of Modern Interfaces to Increase Performance of Linear Algebra Routines” M. Woolsey, W. E. Hutchcraft, and R. K. Gordon.....	292

# A Review of Perfectly Matched Absorbers for the Finite-Volume Time-Domain Method

T. Kaufmann, K. Sankaran, C. Fumeaux, and R. Vahldieck

Laboratory for Electromagnetic Fields and Microwave Electronics - IFH, ETH Zurich  
Zurich, CH-8092, Switzerland, Email: thomas.kaufmann@ifh.ee.ethz.ch

**Abstract** – Different implementations of planar perfectly matched absorbers are studied under the unified framework of the Finite-Volume Time-Domain (FVTD) method. This comparative analysis allows to discuss the similarities existing between the theoretical models and explores the differences in their practical implementation and numerical performance in the framework of the FVTD method. Numerical experiments for performance analysis of the different PML models are conducted in terms of discretization and angle of incidence using waveguide models. The results are compared to theoretically expected values and to the first-order Silver Müller absorbing boundary condition.

## I. INTRODUCTION

One of the biggest challenges in computational electromagnetics is to find domain truncation techniques which can accurately simulate an infinite space in a finite computational domain. The perfectly matched layer (PML) technique was introduced in [1] and it improved the accuracy of the numerical simulations by many orders of magnitude compared to previously applied traditional absorbing boundary conditions. Different implementations of the PML technique were reported in the literature giving rise to two general classes, namely non-Maxwellian (split) and Maxwellian (unsplit) absorbers. The implementation of the PML technique in conformal methods were limited to the Finite-Element Frequency-Domain or Time-Domain methods [2, 3]. A vertex-centered Finite-Volume Time-Domain (FVTD) model (variational approach) of the Bérenger PML (B-PML) was reported in [4] for scattering problems. Recently the authors introduced the cell-centered FVTD implementation of B-PML and modified Lorentz material-based PML (M-PML) techniques in [5, 6]. The generalized theory based perfectly matched layer (GT-PML) [7] and the complex frequency shifted perfectly matched layer (CFS-PML) [8, 9] were introduced for the FVTD method in [10]. The present paper extends the discussion on the unsplit perfectly matched layer (U-PML) model [11] and provides a more thorough performance evaluation. The theory of five different PML techniques, namely B-PML, M-PML, U-PML, GT-PML and CFS-PML, is studied under the unified FVTD framework and the numerical performance of the different PMLs is compared. The

abbreviations are summarized in Table 1.

The paper is arranged as follows. In Sec. II some fundamentals on the FVTD method are given in a notation that will be used throughout this paper. Sec. III focuses on both the split and unsplit PML models. Derivations of the different PML models are summarized with respect to the FVTD implementation and analytical relations between the PML models are given. The computational cost is compared in Sec. IV. Numerical experiments are presented in Sec. V for the different PML models, including an investigation of evanescent wave absorption. The conclusion in Sec. VI summarizes the findings and emphasizes the practical application range for the different PML models.

Table 1. List of abbreviations of the different PML models used in this publication.

Bérenger PML	B-PML
Modified Lorentz Material-based PML	M-PML
Unsplit anisotropic PML	U-PML
Generalized Theory based PML	GT-PML
Complex Frequency Shifted PML	CFS-PML

## II. FUNDAMENTALS OF THE FVTD METHOD

The FVTD method belongs to the general class of conformal time-domain methods. For the spatial discretization, the FVTD method employs unstructured polygons (typically tetrahedrons in 3D and triangles in 2D) which can model complex geometries using highly inhomogeneous meshes. Furthermore, curved boundaries can be modeled with high accuracy because stair-casing errors are avoided. Although this flexibility in spatial discretization is common to all conformal methods, the advantage of the FVTD method lies in the combination of an unstructured spatial discretization with an explicit time update. The method applied in this paper uses a cell-centered approach, hence field values at cell centers are updated by summing up the incoming and outgoing fluxes through each cell face. The update equation can be formulated as follows,

$$\partial_t U_i = -\frac{1}{|A_i|} \sum_{k=1}^f |S_k| \alpha_i^{-1} \mathcal{F}_{U_k^*} \cdot \mathbf{n}_k - \mathcal{L}_i \quad (1)$$

where  $\mathbf{U}_i$  denotes the cell-center field values at the  $i$ th cell,  $|A_i|$  is the cell-volume,  $|S_k|$  the area of the  $k$ th face and  $\mathcal{F}_{\mathbf{U}_k^*} \cdot \mathbf{n}_k$  the sum of the incoming and outgoing flux with the normal vector  $\mathbf{n}_k$  perpendicular to the face [12]. The “\*” in the subscript indicates that the computed flux-function across each face depends on the field quantities at the edge-center. A second-order accurate MUSCL algorithm is employed for spatial discretization. The field values at the face-center are approximated with the help of corresponding cell-center field values. For more information on the FVTD method, the reader is referred to [12]. Material properties of the cells are given in the diagonal matrix  $\alpha_i$  and a loss term  $\mathcal{L}_i$  allows to include material losses. The latter term will be used in the present work to incorporate the PML models into the FVTD method.

In order to simplify the investigation, the setup in this paper relies on a formulation using a two-dimensional (2D) transverse electric ( $TE$ ) form of the Maxwell system. The magnetic fields are assumed to be in the  $xy$ -plane and the electric-field is directed along the  $z$ -axis transverse with respect to the plane of propagation ( $xy$ -plane). Thus the field vector  $\mathbf{U}_i$  and the material parameter  $\alpha_i$  inside the  $i$ th cell become,

$$\mathbf{U}_i = \begin{pmatrix} H_{xi} \\ H_{yi} \\ E_{zi} \end{pmatrix} \quad \text{and} \quad \alpha_i = \begin{pmatrix} \mu_i & 0 & 0 \\ 0 & \mu_i & 0 \\ 0 & 0 & \varepsilon_i \end{pmatrix}. \quad (2)$$

For the time discretization, a second-order accurate explicit Lax-Wendroff scheme is employed which is based on a Predictor-Corrector algorithm [12]. There are other possibilities of time-stepping schemes such as the higher-order Runge-Kutta that can be employed within the presented framework. Although a 2D formulation is used here, the results can be generalized to 3D.

### III. PERFECTLY MATCHED LAYERS

PMLs were first introduced for the Finite-Difference Time-Domain (FDTD) method by Bérenger in [13] using a non-Maxwellian split-field formulation. Later another class of approaches using a Maxwellian unsplit formulation was developed which includes M-PML [14], U-PML [11], GT-PML [7] and CFS-PML [8]. All variations were implemented in the FDTD method.

In the framework of the FVTD method the first-order Silver-Müller absorbing boundary condition (SM-ABC) is commonly used. However, the PML technique promises improved performance for off-normal incidence on the truncating boundary.

The theoretical discussions presented in this paper assume (without loss of generality) that all PML models are used to truncate the computational domain with a planar absorber along the  $y$ -direction and to absorb uniaxially along the  $x$ -axis (see Fig. 1). In future work these planar PMLs will serve as basis for generalization to conformal geometries such as cylinders [6] or spheres. Thus, corner regions are not considered in this publication.

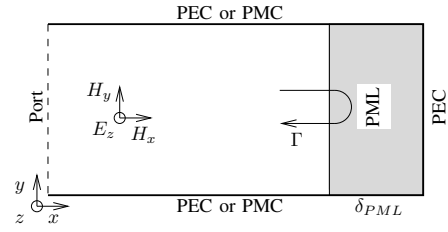


Fig. 1. Configuration of the uniaxial PML models in  $x$ -direction for the two-dimensional  $TE$  setup. The boundaries along the  $x$ -direction depend on the numerical experiment and are either PEC (rectangular waveguide) or PMC (parallel-plate waveguide).

#### A. Split PML

The Bérenger PML (also called the split-field PML) involves unphysical field splitting inside the PML domain which results in an increased number of update equations. For a complete theoretical treatment of the FVTD formulation of B-PML, refer to [5]. Due to the field splitting, an additional field component  $E_{zy}$  is introduced thus, expanding the field vector to  $\mathbf{U}_i = [H_{xi}, H_{yi}, E_{zi}, E_{zyi}]^T$ .

It is worth mentioning that in the above update equations, the fourth field equation for  $E_{zy}$  constitutes the non-hyperbolic part of the system and requires special treatment for its update. By employing the Rankine-Hugoniot jump relation discussed in [5], the field values of  $E_{zy}$  can be updated in a stable manner using the following flux term,

$$\mathcal{F}_{E_{zy}} \cdot \mathbf{n}_k = \frac{n_y}{2} (H_{xl} + H_{xr}) - \frac{n_y^2}{2} (c_r \varepsilon_r E_{zr} - c_l \varepsilon_l E_{zl}). \quad (3)$$

The loss vector  $\mathcal{L}_i$  used in the FVTD update equations for perfectly matched absorption is written as,

$$\mathcal{L}_i = \begin{pmatrix} 0 \\ (\sigma_x/\varepsilon_i) H_{yi} \\ (\sigma_x/\varepsilon_i) (E_{zi} - E_{zyi}) \\ 0 \end{pmatrix}. \quad (4)$$

#### B. Unsplit PML

In order to avoid the unphysical additional field component, unsplit PML models have been developed based on anisotropic material properties. Those models still satisfy the Maxwell equations. For achieving uniaxial absorption in  $x$ -direction, the permittivity and permeability tensors are written as follows,

$$\bar{\varepsilon} = \varepsilon [\Lambda], \quad \bar{\mu} = \mu [\Lambda] \quad \text{with} \quad [\Lambda] = \begin{pmatrix} 1/a & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & a \end{pmatrix}. \quad (5)$$

This form is derived under the condition of perfect matching as described in [7, 11, 14]. The four unsplit PML models presented in the following interpret this anisotropic matrix in different ways. In the case of



M-PML, the tensor describes a time-derivative Lorentz material model. In the case of U-PML the parameters are understood as physical material parameters which are included by explicitly using the magnetic and/or electric flux density. For GT-PML and CFS-PML, the matrix is interpreted as a complex coordinate stretching in frequency-domain which leads to integral or convolution terms in time-domain. In the following, update equations for the FVTD method are derived and compared theoretically.

1) *Modified Lorentz Material-based PML - (M-PML)*: The idea of a modified Lorentz material-based absorber was introduced for the FDTD method in [14]. The anisotropy parameter  $a$  in this case is defined as  $a = 1 + \chi_\omega^m$ , with,

$$\chi_\omega^m = \frac{\omega_0^2 [\chi_\alpha - j(\omega/\omega_0)\chi_\beta]}{\omega_0^2 - \omega^2 - j\Gamma\omega}. \quad (6)$$

The parameters  $\chi_\alpha$ ,  $\chi_\beta$ ,  $\Gamma$  and  $\omega_0$  are chosen so that the material acts as a broadband absorber. The final system of update equations for the M-PML model can be written as,

$$\partial_t H_x = -\frac{1}{\mu} \partial_y E_z + \zeta H_x - G_x \quad (7)$$

$$\partial_t H_y = \frac{1}{\mu} \partial_x E_z - \zeta H_y, \quad (8)$$

$$\partial_t E_z = \frac{1}{\varepsilon} (\partial_x H_y - \partial_y H_x) - \zeta E_z, \quad (9)$$

$$\partial_t G_x = -\zeta G_x + \zeta^2 H_x, \quad (10)$$

where  $\zeta$  is the material loss-parameter inside the absorbing layer. The fourth equation (10) is an ordinary differential equation in time and hence requires no special flux computation. Also the inherent structure of the Maxwellian system with three field components, namely  $H_x$ ,  $H_y$  and  $E_z$  is preserved and there is only an auxiliary equation for the field component  $G_x$  which causes no significant computational overhead, in contrast to the B-PML model discussed in the previous section. The above system of equations is expressed in the FVTD method in notation (1) by defining the field vector as  $\mathbf{U}_i = [H_{xi}, H_{yi}, E_{zi}, G_{xi}]^T$  and the corresponding PML loss vector  $\mathcal{L}_i$  as,

$$\mathcal{L}_i = \begin{pmatrix} G_{xi} - \zeta H_{xi} \\ \zeta H_{yi} \\ \zeta E_{zi} \\ \zeta G_{xi} - \zeta^2 H_{xi} \end{pmatrix}. \quad (11)$$

The FVTD implementation of the M-PML adapted for unstructured grid is discussed in more detail in [6].

2) *Unsplit anisotropic PML - (U-PML)*: Another approach to model unsplit PML based on the anisotropic material properties was introduced for the FDTD method in [11]. Using this model the component  $a$  in equation

(5) is defined as

$$a = 1 + \frac{\sigma_x}{j\omega\varepsilon} \quad (12)$$

where  $\sigma_x$  represents the loss term and  $\varepsilon$  denotes the permittivity. In order to include this lossy frequency-domain parameter into the conformal time-domain update scheme, the physical magnetic flux density  $\mathbf{B}$  is introduced (and/or the electric flux density  $\mathbf{D}$ , depending on the propagation mode). In the two-dimensional  $TE$  mode, the component of  $\mathbf{B}$  in the anisotropy direction takes the form,

$$B_x = \frac{\mu}{1 + \frac{\sigma_x}{j\omega\varepsilon}} H_x \quad (13)$$

and needs to be explicitly included in the equation set because of its nonlinear frequency-dependence.

Based on [11], the following update equations can be derived,

$$\partial_t H_x = -\frac{1}{\mu} \partial_y E_z + \frac{\sigma_x}{\mu\varepsilon} B_x \quad (14)$$

$$\partial_t H_y = \frac{1}{\mu} \partial_x E_z - \frac{\sigma_x}{\varepsilon} H_y, \quad (15)$$

$$\partial_t E_z = \frac{1}{\varepsilon} (\partial_x H_y - \partial_y H_x) - \frac{\sigma_x}{\varepsilon} E_z, \quad (16)$$

$$\partial_t B_x = -\partial_y E_z. \quad (17)$$

In the FVTD formulation (1), the field vector  $\mathbf{U}_i = [H_{xi}, H_{yi}, E_{zi}, B_{xi}]^T$  is then used. This leads to the following formulation for the FVTD method. The fourth flux term for the  $B_x$  field is identical to the flux of the  $H_x$  field ( $\mathcal{F}_{B_x} = \mathcal{F}_{H_x}$ ) and hence, the PML loss vector becomes,

$$\mathcal{L}_i = \begin{pmatrix} -\frac{\sigma_x}{\mu_i \varepsilon_i} B_{zi} \\ \frac{\sigma_x}{\varepsilon_i} H_{yi} \\ \frac{\sigma_x}{\varepsilon_i} E_{zi} \\ 0 \end{pmatrix}. \quad (18)$$

3) *Generalized Theory based PML - (GT-PML)*:

As opposed to the two previous approaches where the anisotropy tensor  $[\Lambda]$  described a material property, the GT-PML model considers  $[\Lambda]$  as a geometrical stretching operator ensuring perfect matching. In [7], the following complex frequency-dependant stretching factor is defined,

$$a = 1 + \frac{\omega_x''}{j\omega} \quad (19)$$

where  $\omega_x''$  describes the rate at which the field is attenuated within the PML. In the two-dimensional  $TE$  case, this leads to the usual lossy formulation for the electric field and the  $y$ -component for the magnetic field. In frequency-domain the  $x$ -component of the magnetic field becomes,

$$j\omega\mu H_x = -\partial_y E_z \cdot a = -\partial_y E_z - \frac{\omega_x''}{j\omega} \partial_y E_z. \quad (20)$$

Transforming this frequency-domain formulation into time-domain yields an integral term for the  $H_x$ -component and that results in the following GT-PML

update equations,

$$\partial_t H_x = -\frac{1}{\mu} \partial_y E_z - \frac{\omega''_x}{\mu} \int_0^t (\partial_y E_z) dt \quad (21)$$

$$\partial_t H_y = \frac{1}{\mu} \partial_x E_z - \omega''_x H_y, \quad (22)$$

$$\partial_t E_z = \frac{1}{\varepsilon} (\partial_x H_y - \partial_y H_x) - \omega''_x E_z. \quad (23)$$

In the notation of equation (1), the field vector for GT-PML is given as  $\mathbf{U}_i = [H_{xi}, H_{yi}, E_{zi}]^T$  and the corresponding lossy PML vector for the FVTD method is,

$$\mathcal{L}_i = \begin{pmatrix} -\frac{\omega''_x}{\mu_i} \int_0^t (\sum_{k=1}^f |S_k| \mathcal{F}_{H_{xk}} \cdot n_k) dt \\ \omega''_x H_{yi} \\ \omega''_x E_{zi} \end{pmatrix}. \quad (24)$$

#### 4) Complex Frequency Shifted PML - (CFS-PML):

The theory of CFS-PML was first introduced in [8] based on complex frequency shifted PML parameters. This shifting is a more general form of the factor used in the previous section, namely,

$$a = \kappa_x + \frac{\sigma_x}{\alpha_x + j\omega\varepsilon} \quad (25)$$

where  $\kappa_x$  represents the real geometrical stretching,  $\alpha_x$  is used to control the absorption for evanescent waves and  $\sigma_x$  is the physical loss.

This approach is of particular interest for damping evanescent waves since all the conventional PML models described in the previous sections exhibit a performance degradation in the evanescent regime. The complex stretching factor described in [15] can be implemented in two ways. The first approach is using an auxiliary differential equation method and the second approach involves a time-domain convolution. For the present FVTD version of CFS-PML, the second approach is utilized. As shown in [9] this results in the following update equations,

$$\partial_t H_x = -\frac{1}{\mu} \partial_y E_z \kappa_x + \Psi_{H_x} \quad (26)$$

$$\partial_t H_y = \frac{1}{\mu} \frac{\partial_x E_z}{\kappa_x} - \frac{\sigma_x}{\varepsilon \kappa_x} H_y + \Psi_{H_y}, \quad (27)$$

$$\partial_t E_z = \frac{1}{\mu \kappa_x} (\partial_x H_y) - \frac{\sigma_x}{\varepsilon \kappa_x} E_z + \Psi_{E_z}. \quad (28)$$

Apart from the standard Maxwellian fields ( $H_x$ ,  $H_y$  and  $E_z$ ) and the material parameters ( $\mu = \mu_r \mu_0$ ,  $\varepsilon = \varepsilon_r \varepsilon_0$  and  $\sigma_x$ ), two new factors, namely  $\alpha_x$  and  $\kappa_x$  are introduced. These factors give additional degrees of freedom in controlling the perfectly matched damping behavior inside the PML. The terms  $\Psi_{H_x}$ ,  $\Psi_{H_y}$  and  $\Psi_{E_z}$  in equations (26) to (28) represent the convolution operation in time for each field value and they can be written for the FVTD

method as,

$$\Psi_{H_{xi}} = -\frac{\sigma_x}{\mu_i \varepsilon_i} e^{-\frac{\alpha_x}{\varepsilon_i} \cdot t} * \sum_{k=1}^f (\mathcal{F}_{H_{xk}} \cdot n_k) |S_k| \quad (29)$$

$$\Psi_{H_{yi}} = \frac{\alpha_x \sigma_x}{\kappa_x \varepsilon_i^2} e^{-\frac{\alpha_x}{\varepsilon_i} \cdot t} * H_{yi}, \quad (30)$$

$$\Psi_{E_{zi}} = \frac{\alpha_x \sigma_x}{\kappa_x \varepsilon_i^2} e^{-\frac{\alpha_x}{\varepsilon_i} \cdot t} * E_{zi}. \quad (31)$$

This describes the continuous-time formulation of the convolutions which would be highly inefficient to implement in a discrete-time scheme since a sum over all time needs to be calculated at each time step. Thus, the iterative method proposed in [9] can be applied to simplify the convolution to one addition per time step.

This formulation eventually allows to express a loss term using the field vector  $\mathbf{U}_i = [H_{xi}, H_{yi}, E_{zi}]^T$  in equation (1) as,

$$\mathcal{L}_i = \begin{pmatrix} -\Psi_{H_{xi}} \\ \frac{\sigma_x}{\varepsilon_i \kappa_x} H_{yi} - \Psi_{H_{yi}} \\ \frac{\sigma_x}{\varepsilon_i \kappa_x} E_{zi} - \Psi_{E_{zi}} \end{pmatrix}. \quad (32)$$

These convolution operations eventually make the CFS-PML highly efficient for the absorption of evanescent waves, however, at the cost of a somewhat increased computational effort.

Guidelines to choose the parameters are given in [16], where derivations show that the PML mainly absorbs propagating modes if the term  $\frac{\sigma_x}{\alpha_x + j\omega\varepsilon}$  in equation (25) is mainly complex. In contrast, evanescent waves are best absorbed if the fraction is mainly real. Thus in the case of a waveguide the value for  $\alpha_x$  has to be chosen so that the switching frequency,

$$f_\alpha = \frac{\alpha_x}{2\pi\varepsilon} \quad (33)$$

corresponds to the cutoff frequency. The factor  $\kappa_x$  stretches the coordinate system, which in practice affects the accuracy of the simulation. In the presented work  $\kappa_x$  is fixed at  $\kappa_x = 1$  in order to avoid stretched coordinate discretization errors.

### C. Relationships between PML models

1) *Unsplit PML models:* Comparing the formulations of M-PML, U-PML and GT-PML suggests a close similarity even though the approaches differ in their physical interpretations. M-PML and U-PML both use one additional field term and GT-PML incorporates an additional integral term. When rewriting the set of equations for M-PML (7) to (10) by inserting the additional field term  $G_x = K_x + \zeta H_x$  into equation (7), the following set

emerges,

$$\partial_t H_x = -\frac{1}{\mu} \partial_y E_z - K_x \quad (34)$$

$$\partial_t H_y = \frac{1}{\mu} \partial_x E_z - \zeta H_y, \quad (35)$$

$$\partial_t E_z = \frac{1}{\varepsilon} (\partial_x H_y - \partial_y H_x) - \zeta E_z, \quad (36)$$

$$\partial_t K_x = \frac{\zeta}{\mu} \partial_y E_z, \quad (37)$$

where  $K_x$  represents the magnetic polarization current. Comparing the rewritten system in equations (34) to (37) of M-PML with that of U-PML in equations (14) to (17), their mathematical equivalence becomes apparent, although different physical quantities are considered ( $K_x$  and  $B_x$ ).

Applying a similar reformulation of GT-PML, the integral term can be written as a separate update equation identical to the one of U-PML. This rewriting is required anyhow before numerical implementation to allow for the iterative integration. In the present formulation, the Lax-Wendroff time stepping is applied. This allows for sophisticated integration methods. In this case the Simpson rule is used because it fits perfectly into the two step updating scheme. Therefore the final numerical FVTD implementations of the three PML models differ slightly, albeit they are analytically equivalent. Hence, for the numerical implementation of the scheme only minor differences in the range of the numerical precision are expected.

The relationship between the absorption parameters used in the different formulations can be expressed as follows,

$$\zeta = \omega_x'' = \frac{\sigma_x}{\varepsilon}. \quad (38)$$

2) *Split vs. Unsplit PML*: Considering further the relation between split and unsplit models, a close connection can be found between B-PML and U-PML, as pointed out in [11, 17]. Splitting the U-PML formulation, the formulation for B-PML can be retrieved, or vice-versa. Hence, numerically identical results are expected here as well.

3) *CFS-PML to GT-PML*: Finally it has to be pointed out that, as CFS-PML is a generalization of GT-PML, the convolutions of equations (29) to (31) should reduce to the integral term of GT-PML if  $\alpha_x \rightarrow 0$  and  $\kappa_x = 1$ . In fact it can be verified that,

$$\lim_{\alpha_x \rightarrow 0, \kappa_x = 1} \Psi_{H_{x_i}} = -\frac{\sigma_x}{\mu_i \varepsilon_i} * \sum_{k=1}^f (\mathcal{F}_{H_{x_k}} \cdot n_k) |S_k| \quad (39)$$

$$\lim_{\alpha_x \rightarrow 0, \kappa_x = 1} \Psi_{H_{y_i}} = 0, \quad (40)$$

$$\lim_{\alpha_x \rightarrow 0, \kappa_x = 1} \Psi_{E_{z_i}} = 0. \quad (41)$$

The convolution in equation (39) is a simple integration over time, and thus yields exactly the same term as in GT-PML. This also applies to the discrete formulation obtained by the method of [9]. Figure 2 summarizes the relationship between all the investigated PML models.

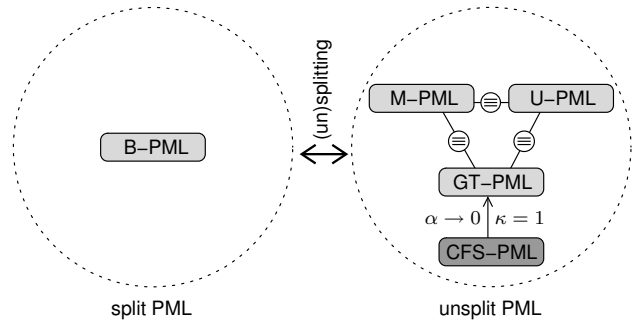


Fig. 2. Analytical relationship between the investigated PML models.

#### IV. CONSIDERATIONS ON COMPUTATIONAL EFFORT

It is shown in the previous section that the analytical formulations of B-PML, M-PML, U-PML and GT-PML are analytically equivalent. Nevertheless, the discrete models vary because of the different required operations. The numerical implementations include an additional flux term for B-PML, additional update equations for M-PML and U-PML and an integral term for GT-PML.

Table 2 shows the number of variables necessary for the investigated PML models in 2D. In free space only three field variables are updated using three flux terms. In the case of B-PML, the fourth split field requires the calculation of an additional flux term at each time step. This is not necessary for M-PML. Due to the nature of the fourth differential equation, no flux needs to be calculated, but one additional update equation is necessary. The fourth field in U-PML requires calculation of an additional flux. But since this flux is identical to the flux of the  $H_x$  field, that does not increase the computational cost. The formulation of GT-PML requires no additional update equation, and thus has only three field variables. In the numerical implementation however, the update formulation of the integral term leads to an additional variable, hence making the formulation identical to U-PML. Nevertheless the computational effort for GT-PML might increase depending on the sophistication of the

Table 2. Computational effort for two-dimensional models.  $K$  is the number of field update variables,  $M$  the number of flux variables and  $N$  the number of additional update variables.

Model	$K$	$M$	$N$	Auxiliary operation
free space	3	3	0	-
B-PML	4	4	0	flux
M-PML	4	3	0	ODE
U-PML	4	3	0	PDE
GT-PML	3	3	1	integral
CFS-PML	3	3	3	convolutions

integration method. Finally, CFS-PML adds a convolution operation to each field term, thus adding three additional operations and variables per cell in the two-dimensional  $TE$  case.

To summarize the computational effort necessary for all PML models, B-PML is slightly more costly than the basic unsplit PML models due to the calculation of the additional flux. CFS-PML significantly increases the computational costs as it requires one additional convolution operation for each field variable. In this case increased absorption for evanescent waves is achieved at the cost of increased memory and computation time.

## V. NUMERICAL PERFORMANCE COMPARISON

To validate the theoretical findings of the previous section, numerical measurements are performed using first a plane wave problem and second a waveguide problem. These particular setups were chosen to measure the influence of discretization and to investigate the reflection at off-normal angles of incidence. As the implementation of U-PML and GT-PML are identical in FVTD, U-PML results are not displayed explicitly.

### A. Plane Wave at Normal Incidence

At normal incidence the broadband performance of all the PML models is compared to that of the first-order SM-ABC. Broadband analysis provides information on the influence of the spatial discretization and the thickness of the absorbing layer in terms of the wavelength. A plane wave is simulated using a parallel-plate waveguide setup in Fig. 1 (with PMC boundaries). The scattering parameter extraction as discussed in [18] is used to retrieve the reflection coefficient  $S_{11}$  of the PML models. The results are shown in Fig. 3 with the PML parameters set to achieve a theoretical reflection [13] of  $R = -80$  dB. The model is fed with a sine-modulated Gaussian broadband pulse with effective bandwidth stretching from 1 GHz to 50 GHz. It is observed that the PML performance over the whole bandwidth remains close to the theoretically expected value, with a degradation at higher frequencies due to coarse spatial discretization. In the investigated problem, the overall performance of the SM-ABC is inferior to that of the PML models over most of the frequency range. At low frequencies (fine spatial discretization) the influence of discretization errors diminishes. It is noticed that the performance of SM-ABC at normal incidence numerically converges towards perfect absorption.

It is observed that all the investigated PML models perform identically, as expected from the discussion in the previous sections. In particular in the present configuration CFS-PML becomes numerically and theoretically identical to GT-PML. This is due to the fact that the parallel plate waveguide exhibits a cutoff frequency  $f_c = 0$  and thus, according to equation (33)  $\alpha_x = 0$ . Consequently the CFS-PML model becomes identical to GT-PML.

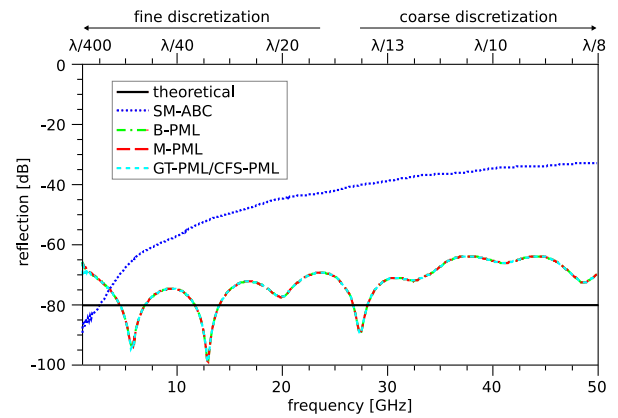


Fig. 3. Numerical results at normal incidence. The spatial discretization is shown in the upper scale.

### B. Off-Normal Incidence and Evanescent Waves

In the second example, a waveguide model (Fig. 1 with PEC boundaries) excited with a  $TE_{10}$  mode is used to compute the reflection coefficient of the investigated PML models at a range of incident angles. In addition this model permits evaluation of absorption below the cutoff frequency of the waveguide, i.e., for evanescent waves. To measure the reflection of evanescent waves, the technique introduced by Gwarek et al. in [19, 20] is applied here to extract the  $S_{11}$ -parameter. This technique uses the tangential fields ( $E_z$  and  $H_y$ ) and gives information on the physical reflection of the fields, even for evanescent waves. This stands in contrast to the traditional definition of scattering parameters based on the energy flow, where evanescent waves are understood to be totally reflected.

The investigated model has a cutoff frequency of 6.56 GHz for its fundamental  $TE_{10}$  mode and is fed with a modulated Gaussian broadband pulse with bandwidth stretching from 4 GHz to 20 GHz which spans both the evanescent and propagation regions. The spatial discretization employed for the model corresponds to a range of  $\lambda/20$  to  $\lambda/100$  (free-space wavelength) at the highest and lowest frequency components of the input signal, respectively. Simulated PML models are B-PML, M-PML, GT-PML and CFS-PML which are compared to the performance of the SM-ABC.

The results obtained are plotted in Fig. 4 with a theoretical reflection coefficient set to  $R = -80$  dB at normal incidence. Depending on the frequency, evanescent waves naturally decay inside the finite thickness of the PML. This influence is included in the depicted theoretical absorption below cutoff. Because the field pattern within a waveguide can be understood as a superposition of two plane waves travelling with a certain off-normal angle, the influence of the angle of incidence is also addressed in this experiment. The cutoff frequency corresponds to grazing incidence and at higher frequencies, the angle of incidence converges towards normal incidence. This angle is also given in the upper scale of Fig. 4 for illustration.

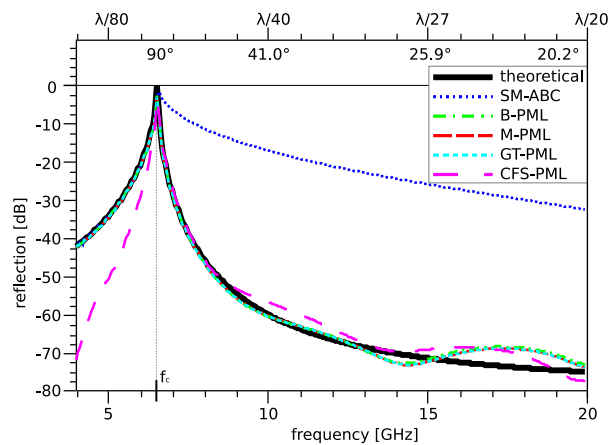


Fig. 4. Numerical results for off-normal incidence and evanescent waves in a waveguide model with thickness  $a = 22.86$  mm. The upper scale shows the spatial discretization and the angle of incidence.

It is clearly noticed that above cutoff, all PML models perform very close to the theoretically expected value. Additionally it should be noted that the B-PML, M-PML and GT-PML models perform identically, as it was expected from their theoretical identity. Below cutoff, all ABCs except CFS-PML do not absorb and therefore only the natural decay is measured. In contrast, in the case of CFS-PML, the coordinate system stretching elongates the layer and hence the decay of evanescent waves is increased. This leads to an additional absorption of up to 30 dB in this case.

## VI. CONCLUSION

In the present study different PML techniques were modeled in the unified framework of the FVTD method. The theoretical equivalence of M-PML, U-PML and GT-PML was discussed. Conditions were given that simplify CFS-PML to GT-PML. Numerical performances of all the PML models were found to be nearly identical for propagating modes. As expected, the performance of the CFS-PML was substantially better in the evanescent wave region compared to other PML models.

The improvement for CFS-PML compared to the other PML models arises for the absorption of evanescent waves and is achieved at the cost of an increased computational load. Therefore a practical application for the more costly model is only reasonable when strong evanescent waves have to be absorbed close to a source. Nevertheless the more efficient unsplit Maxwellian PML models, such as the U-PML, are sufficient for most applications in conformal time-domain methods.

Although the results presented here were obtained employing planar perfectly matched layers, this study represents a first step towards the extension of PML techniques for non-planar surfaces in the FVTD method.

## ACKNOWLEDGEMENT

This work is supported by the ETH Research Grant TH-38/04-1

## REFERENCES

- [1] J.-P. Bérenger, "A perfectly matched layer for the absorption of electromagnetic waves," *Journal of Computational Physics*, vol. 114, no. 2, pp. 185–200, 1994.
- [2] Z. Sacks, D. Kingsland, R. Lee, and J.-F. Lee, "A perfectly matched anisotropic absorber for use as an absorbing boundary condition," *IEEE Transactions on Antennas and Propagation*, vol. 43, no. 12, pp. 1460–1463, December 1995.
- [3] T. Rylander and J.-M. Jin, "Perfectly matched layer in three dimensions for the time-domain finite element method applied to radiation problems," *IEEE Transactions on Antennas and Propagation*, vol. 53, no. 4, pp. 1489–1499, April 2005.
- [4] F. Bonnet and F. Poupaud, "Bérenger absorbing boundary condition with time finite-volume scheme for triangular meshes," *Applied Numerical Mathematics*, vol. 25, no. 4, pp. 333–354, December 1997.
- [5] K. Sankaran, C. Fumeaux, and R. Vahldieck, "Cell-centered finite-volume based perfectly matched layer for time domain Maxwell system," *IEEE Transactions on Microwave Theory and Technique*, vol. 54, no. 3, pp. 1269–1276, March 2006.
- [6] —, "Uniaxial and radial anisotropy models for finite-volume Maxwellian absorber," *IEEE Transactions on Microwave Theory and Technique*, vol. 54, no. 12, pp. 4297–4304, December 2006.
- [7] L. Zhao and A. Cangellaris, "GT-PML: Generalized theory of perfectly matched layers and its application to the reflectionless truncation of finite-difference time-domain grids," *IEEE Transactions on Microwave Theory and Techniques*, vol. 44, no. 12, pp. 2555–2563, December 1996.
- [8] M. Kuzuoğlu and R. Mittra, "Frequency dependence of the constitutive parameters of causal perfectly matched absorbers," *IEEE Microwave Guided Wave Letters*, vol. 6, pp. 447–449, December 1996.
- [9] A. Roden and S. Gedney, "Convolution PML (CPML): an efficient FDTD implementation of the CFS-PML for arbitrary media," *Microwave and Optical Technology Letters*, vol. 27, no. 5, pp. 334–339, 2000.
- [10] K. Sankaran, T. Kaufmann, C. Fumeaux, and R. Vahldieck, "Different perfectly matched absorbers for conformal time-domain method: A finite-volume time-domain perspective," in *23rd Annual Review of Progress in Applied Computational Electromagnetics (ACES)*, Verona, Italy, March 2007.

- [11] S. Gedney, "An anisotropic perfectly matched layer-absorbing medium for the truncation of FDTD lattices," *IEEE Transactions on Antennas and Propagation*, vol. 44, no. 12, pp. 1630–1639, December 1996.
- [12] P. Bonnet, X. Ferrieres, B. Michielsen, P. Klotz, and J. Roumiguieres, *Time Domain Electromagnetics*. S. M. Rao, Ed., Academic Press, 1997, ch. 9, pp. 307–367.
- [13] J.-P. Bérenger, "Three-dimensional perfectly matched layer for the absorption of electromagnetic waves," *Journal of Computational Physics*, vol. 127, no. 2, pp. 363–379, September 1996.
- [14] R. W. Ziolkowski, "The design of Maxwellian absorbers for numerical boundary conditions and for practical applications using engineered artificial materials," *IEEE Transactions of Antennas and Propagation*, vol. 45, no. 4, pp. 656–671, April 1997.
- [15] W. C. Chew and W. H. Weedon, "A 3D perfectly matched medium from modified Maxwell's equations with stretched coordinates," *Microwave Optics*, vol. 114, no. 2, pp. 185–200, 1994.
- [16] J.-P. Bérenger, "Application of the CFS PML to the absorption of evanescent waves in waveguides," *IEEE Microwave and Wireless Components Letters*, vol. 12, no. 6, pp. 218–220, 2002.
- [17] S. Gedney, "Perfectly matched layer absorbing boundary conditions," in *The Finite-Difference Time-Domain Method*, A. Taflove and S. C. Hagness, Eds. Artech House Inc., 2005.
- [18] D. Baumann, C. Fumeaux, and R. Vahldieck, "Field-based scattering-matrix extraction scheme for the FVTD method exploiting a flux-splitting algorithm," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 53, no. 11, pp. 3595–3605, 2005.
- [19] W. K. Gwarek and M. Celuch-Marcysiak, "Wide-band S-parameter extraction from FD-TD simulations for propagating and evanescent modes in inhomogeneous guides," *Microwave Theory and Technique, IEEE Transactions on*, vol. 51, no. 8, pp. 1920–1928, 2003.
- [20] —, "A differential method of reflection coefficient extraction from FDTD simulations," *IEEE Microwave and Guided Waves Letters*, vol. 6, no. 5, pp. 1920–1928, 1996.



**Thomas Kaufmann** received the M.Sc. degree in electrical engineering and information technology from ETH Zurich, Switzerland in 2007 and is currently working towards the Ph. D. degree at the same university in the Laboratory for Electromagnetic Fields and Microwave Electronics (IFH). His research interest includes computational electromagnetics for microwave circuits and antennas. Mr. Kaufmann spent half a year in Toronto, Canada during an internship at InvoDane Engineering in 2004 and an exchange semester at the Norwegian University of Technology (NTNU) in Trondheim, Norway in 2005.



**Krishnaswamy Sankaran** received the B.Eng. degree (with a first-class distinction) in electrical and electronics engineering from the University of Madras, Madras, India, in 2002, the M.Sc. degree in information and communication engineering from the University of Karlsruhe TH, Germany, in 2004, and is currently working toward the Ph.D. degree from the ETH Zurich, Zurich, Switzerland. From October 2003 to May 2004, he was a Research Trainee with the European Commission, Joint Research Centre, Ispra, Italy, where he was involved in the field of radar systems engineering and remote sensing. In June 2004, he joined the ETH Zurich, where he is currently with the Laboratory for Electromagnetic Fields and Microwave Electronics (IFH). His main research interests are numerical methods for solving EM field problems, computational physics, and applied mathematics. In January 2007, he was invited by the Isaac Newton Institute for Mathematical Sciences, University of Cambridge, Cambridge, UK. Mr. Sankaran is currently the Chair of the IEEE Student Branch Zürich. He was the recipient of a full postgraduate scholarship and he was one of the recipients of the 2006 Best Student Paper Award presented at the IEEE Microwave Theory and Techniques Society International Microwave Symposium, San Francisco, CA.



**Christophe Fumeaux** received the Diploma and Ph.D. degrees in physics from the ETH Zurich, Switzerland, in 1992 and 1997, respectively. From 1998 to 2000, he was a Post-Doctoral Researcher with the School of Optics, University of Central Florida, Orlando. In 2000, he joined the Swiss Federal Office of Metrology, Bern, Switzerland, as a Scientific Staff Member. Since 2001, he has been a Research Associate with the Laboratory for Electromagnetic Fields and Microwave Electronics (IFH), ETH, Zurich, Switzerland. During the Fall of 2005, he was a Visiting Scientist with the Laboratory of Sciences and Materials for Electronics, and of Automatic (LASMEA), University Blaise Pascal, Clermont-Ferrand, France. His current main research interest concerns computational electromagnetics in the time domain for numerical analysis of microwave circuits and antennas. Dr. Fumeaux has been the chairman of the IEEE Swiss Joint Chapter on Microwave Theory and Techniques, Antennas and Propagation, and EMC since January 2006. He was the recipient of the ETH Silver Medal of Excellence for his doctoral dissertation. He was the corecipient of the outstanding paper award of the Applied Computational Electromagnetics Society (ACES) in 2004.



**Rüdiger Vahldieck** received the Dipl.-Ing. and Dr.-Ing. degrees in electrical engineering from the University of Bremen, Bremen, Germany, in 1980 and 1983, respectively. From 1984 to 1986, he was a Post-Doctoral Fellow with the University of Ottawa, Ottawa, ON, Canada. In 1986, he joined the Department of Electrical and Computer Engineering, University of Victoria, Victoria, BC, Canada, where he became a Full Professor in 1991. During the fall of 1992 and the spring of 1993, he was a Visiting Scientist with the Ferdinand-Braun-Institute für Hochfrequenztechnik, Berlin, Germany. In 1997, he became a Professor of EM-field theory with the ETH Zurich, Zurich, Switzerland, and Head of the Laboratory for Electromagnetic Fields and Microwave Electronics (IFH) in 2003. His research interests include computational electromagnetics in the general area of electromagnetic compatibility (EMC) and, in particular, for computer-aided design of microwave, millimeter-wave, and opto-electronic integrated circuits. Since 1981, he has authored or coauthored over 300 technical papers in books, journals, and conferences, mainly in the field of microwave computer-aided design. Prof. Vahldieck is the past president of the IEEE 2000 International Zurich Seminar on Broadband Communications (IZS2000). Since 2003, he has been president and general chairman of the International Zurich Symposium on Electromagnetic Compatibility. He is a member of the Editorial Board of the IEEE TRANSACTIONS ON MICROWAVE THEORY AND TECHNIQUES. From 2000 to 2003, he was an associate editor for the IEEE MICROWAVE AND WIRELESS COMPONENTS LETTERS, and from July 2003 until the end of 2005, he was the editor-in-chief. Since 1992, he has served on the Technical Program Committee (TPC) of the IEEE Microwave Theory and Techniques Society (IEEE MTT-S) International Microwave Symposium (IMS), the IEEE MTT-S Technical Committee on Microwave Field Theory, and in 1999, on the TPC of the European Microwave Conference. From 1998 to 2003, he was the chapter chairman of the IEEE Swiss Joint Chapter on Microwave Theory and Techniques, Antennas and Propagation, and EMC. Since 2005, he has been president of the Swiss Research Foundation on Mobile Communications. He was the recipient of the J. K. Mitra Award of the Institution of Electronics and Telecommunication Engineers (IETE) (in 1996) for the best research paper in 1995 and was corecipient of the Outstanding Publication Award of the Institution of Electronic and Radio Engineers in 1983. He was the corecipient of the 2004 Applied Computational Electromagnetic Society (ACES) Outstanding Paper Award.

# New Stability Criterion for Unstructured Mesh Upwinding FVTD Schemes for Maxwell's Equations

D. Firsov and J. LoVetri

Department of Electrical and Computer Engineering  
University of Manitoba  
Winnipeg, Manitoba, Canada, R3T 5V6  
firsovd@cc.umanitoba.ca lovetri@ee.umanitoba.ca

**Abstract** – A new stability criterion applicable to explicit upwind FVTD schemes for solving Maxwell's equations on unstructured meshes is derived. This criterion is based on  $L_2$ -norm estimates of specially constructed matrices  $G_i$  for each finite volume  $i$ . Each such matrix is constructed using the scalar product of the eigenvectors corresponding to the unity eigenvalues of the fluxsplitting operators associated with the facets of volume  $i$ . The new stability criterion is obtained numerically once the grid is constructed using these matrices over the mesh and is therefore mesh dependent. The new criterion gives a time-step that is larger than the time-step calculated using previously published stability criteria. On structured meshes the new criterion gives the same time-step limit as the von Neumann analysis. The method incurs a small computational expense at the beginning of each run of the algorithm. The method is generalizable but the extent to which it can be generalized to other time-evolving physical phenomenon is not considered in this paper.

**Keywords:** Finite-volume time-domain, unstructured mesh, Maxwell's equations, and stability criterion.

## I. INTRODUCTION

One of the classical drawbacks of using explicit time-stepping numerical schemes is that a stringent timestep limit must be adhered to for stability. For structured meshes this time-step limit can usually be obtained using von Neumann analysis, but it is not possible to use von Neumann analysis with unstructured meshes. Thus, several authors have used analyses based on the energy-norm in the mesh to obtain estimates for the time-step bound. Previously published maximum time-step bounds have been sufficient to ensure stability but not tight bounds: they generally restrict the time-step to a value that is smaller than necessary. Obviously, using a smaller than necessary time-step increases execution time for any particular problem, but in addition, a timestep that is too small compared to the necessary limit may also result in poorer solution accuracy. A sufficient maximum time-step criterion for FVTD upwind schemes for Maxwell's equations was presented in [1, 2]. The time-step limit

given in [2] is,

$$\Delta t = \min_i \frac{V_i}{cA_i} \quad (1)$$

for an unstructured mesh, where  $c_i$  is the speed of light in element  $i$ ,  $1 \leq i \leq N$  is a number identifying the elements in the unstructured mesh (with  $N$  elements),  $V_i$  and  $A_i$  are the volume and total facet area for the  $i$ -th element, respectively. On the other hand, in [1] the timestep limit for an unstructured mesh is given as,

$$\Delta t = \min_i \frac{2V_i}{cA_i} \quad (2)$$

twice that of equation (1) reported in [2]. Unfortunately, as was stated in [1], the bound given by equation (1) is merely a sufficient condition and not necessary: a larger time-step is possible.

For a structured cubical mesh with the edge-size of the elements  $h$  and the same speed of light  $c$  on all elements with von Newman analysis in [2] was derived,

$$\Delta t = \frac{h}{2c} \quad (3)$$

which is necessary condition and 1.5 times large then criterion equation (2) applied to the cubical elements on structured mesh. Note that a structured cubical mesh is a special case of an unstructured mesh in which all elements have the same cubical shape. But, criterion equation (2) does not reduce to equation (3) when applied to a structured cubical mesh and therefore equation (2) cannot be the necessary stability condition.

In this paper we derive the necessary stability criterion for the first-order Euler explicit scheme which can then be easily extended for higher-order time integration schemes [3]. As in [1], the derivation is based on the natural physical constraint that in a mesh that is free of sources of energy the total energy in the mesh should not increase with time. The electromagnetic energy in a particular region,  $\Omega \subset \mathbb{R}^3$ , gives rise to the mathematical concept of an energy-norm which can be calculated as,

$$\|\mathbf{u}\| = \sqrt{\frac{1}{2} \sum_{i=1}^N V_i (\epsilon_i \mathbf{E}_i \cdot \mathbf{E}_i + \mu_i \mathbf{H}_i \cdot \mathbf{H}_i)} \quad (4)$$



The key difference in the derivation that allows us to obtain the necessary criterion is that we express the summation of fluxes over facets as a single matrix operator for which the norm can be determined numerically. Therefore, once an arbitrary mesh is generated, the necessary time-step limit is obtained by computing a simple formula over each element in the mesh at the beginning of each FVTD run.

## II. THE TIME-STEP CRITERION IN TERMS OF ENERGY

Suppose we have a domain  $\Omega \subset \mathbb{R}^3$  upon which is specified an unstructured mesh  $\omega = \bigcup_{i=1}^N \Omega_i$ , where  $\Omega_i$  are the elements of the mesh each having a volume  $V_i$ . Here we suppose that the electromagnetic material parameters  $\epsilon_i$  and  $\mu_i$  are constants on each element  $\Omega_i$ . The finite-volume time-domain method is formulated in terms of a generalized solution vector containing the electric and magnetic field vectors

$$\mathbf{u}(\mathbf{x}) = [\mathbf{E}^T(\mathbf{x}) \mathbf{H}^T(\mathbf{x})]^T,$$

and solved for the averaged values

$$\mathbf{u}_i = \frac{1}{V_i} \int_{\Omega_i} \mathbf{u}(\mathbf{x}) dv,$$

on each element. An equivalent discrete energy-norm over the domain  $\Omega$  can be written as,

$$\|\mathbf{u}\| = \sqrt{\frac{1}{2} \sum_{i=1}^N V_i (\epsilon_i \mathbf{E}_i \cdot \mathbf{E}_i + \mu_i \mathbf{H}_i \cdot \mathbf{H}_i)} \quad (5)$$

where  $\mathbf{E}_i$  represents the averaged value of the electric field over the element  $i$ , and similarly for the magnetic field vector  $\mathbf{H}_i$ .

Consider now the Euler approximation for the time-dependent Maxwell's equations cast as a conservation law (see [4]). It can be written concisely as,

$$\mathbf{u}^{n+1} = \mathbf{u}^n - \Delta t L \mathbf{u}^n \quad (6)$$

where  $L$  represents the discretization of the spatial derivatives. More specifically, for the case of the FVTD method,  $L$  represents the integration of the fluxes over the facets of each element. We are interested in the maximum value of  $\delta t$  that keeps the scheme stable.

A numerical scheme is  $L^2$ -stable if the energy doesn't grow in time; that is, if the following is true,

$$\|\mathbf{u}^{n+1}\|^2 \leq \|\mathbf{u}^n\|^2. \quad (7)$$

We can define an inner-product in the mesh for Maxwell's equations as  $(\mathbf{u}, \mathbf{w}) = \sum_{i=1}^N V_i \mathbf{u}_i^T \alpha_i \mathbf{w}_i$ , where

$$\alpha_i = \begin{pmatrix} \epsilon_i & 0 \\ 0 & \mu_i \end{pmatrix},$$

and  $\epsilon_i$ ,  $\mu_i$  are permittivity and permeability matrices for volume  $i$ .

The energy-norm is obtained as  $\|\mathbf{u}\| = \sqrt{(\mathbf{u}, \mathbf{u})}$ . It can be easily verified, that  $(\mathbf{u}, \mathbf{w})$  satisfies the mathematical properties of an inner product.

Taking the inner product of equation (6) with  $\mathbf{u}^n$  we get  $(\mathbf{u}^{n+1}, \mathbf{u}^n) - (\mathbf{u}^n, \mathbf{u}^n) = -\Delta t (L \mathbf{u}^n, \mathbf{u}^n)$ , and using the property that,

$$(\mathbf{u}^{n+1}, \mathbf{u}^n) = \frac{1}{2} (\mathbf{u}^{n+1}, \mathbf{u}^{n+1}) + \frac{1}{2} (\mathbf{u}^n, \mathbf{u}^n) - \frac{1}{2} (\mathbf{u}^{n+1} - \mathbf{u}^n, \mathbf{u}^{n+1} - \mathbf{u}^n) \quad (8)$$

we can rewrite this as

$$(\mathbf{u}^{n+1}, \mathbf{u}^{n+1}) + (\mathbf{u}^n, \mathbf{u}^n) - (\mathbf{u}^{n+1} - \mathbf{u}^n, \mathbf{u}^{n+1} - \mathbf{u}^n) = -2\Delta t (L \mathbf{u}^n, \mathbf{u}^n).$$

This last equation together with,

$$(\mathbf{u}^{n+1} - \mathbf{u}^n, \mathbf{u}^{n+1} - \mathbf{u}^n) = (\Delta t L \mathbf{u}^n, \Delta t L \mathbf{u}^n)$$

and the energy constraint of equation (7) gives

$$(\mathbf{u}^{n+1}, \mathbf{u}^{n+1}) - (\mathbf{u}^n, \mathbf{u}^n) = \Delta t^2 (L \mathbf{u}^n, L \mathbf{u}^n) - 2\Delta t (L \mathbf{u}^n, \mathbf{u}^n).$$

This gives us a condition for the maximum time-step, based on the non-increasing energy stability criterion for the Euler scheme, that depends on any spatial discretization  $L$ ,

$$\Delta t (L \mathbf{u}^n, L \mathbf{u}^n) \leq 2(L \mathbf{u}^n, \mathbf{u}^n). \quad (9)$$

## III. FVTD SOLUTION OF MAXWELL'S EQUATIONS

### A. The FVTD scheme for Maxwell's equations

We can write the FVTD scheme for Maxwell's equations with Euler explicit time integration and first order spatial upwinding, as [1, 2],

$$\mathbf{u}_i^{n+1} = \mathbf{u}_i^n - \Delta t \frac{1}{V_i} \sum_{j=1}^{m_i} A_i(j) \times (T_i^+(j) B_i^+(j) \mathbf{u}_i^n + T_i^-(j) B_i^-(j) \mathbf{u}_{i_j}^n) \quad (10)$$

where  $m_i$  is the number of facets defining finite-volume  $\Omega_i$ , and  $A_i(j)$  is the area of the  $j^{\text{th}}$  facet of the  $i^{\text{th}}$  volume. The subscript  $i_j$  denotes the element neighboring facet  $j$ . The transmission operators are given as

$$T_i^\pm(j) = \alpha_i^{-1} \begin{pmatrix} \frac{2Y_i^\mp(j)}{Y_i^+(j)+Y_i^-(j)} I & 0 \\ 0 & \frac{2Z_i^\mp(j)}{Z_i^+(j)+Z_i^-(j)} I \end{pmatrix},$$

for facets between dielectrics, where

$$Y_i^+(j) = \frac{1}{Z_i^+(j)} = \sqrt{\frac{\epsilon_i}{\mu_i}},$$

$$Y_i^-(j) = \frac{1}{Z_i^-(j)} = \sqrt{\frac{\epsilon_{i_j}}{\mu_{i_j}}}.$$

For facets located on a perfect electric conductor (PEC) these become

$$T_i^+(j) = \alpha_i^{-1} \begin{pmatrix} 2I & 0 \\ 0 & 0 \end{pmatrix}, T_i^-(j) = 0,$$

or equivalently, using the image principle [4], we have the same  $T_i^\pm(j)$  and  $B_i^\pm(j)$  operators as for facets between dielectrics, but  $\mathbf{u}_i^n = [(\mathbf{E}_i^n)^T(\mathbf{H}_i^n)^T]^T$  and  $\mathbf{u}_{i_j}^n = [-(\mathbf{E}_i^n)^T(\mathbf{H}_i^n)^T]^T$ . This equivalent formulation allows us to extend the stability criterion results for meshes with PEC facets more easily. For a facet at the external boundary of the mesh, we consider the first-order absorbing boundary condition (ABC)

$$T_i^+(j) = \alpha_i^{-1} \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}, T_i^-(j) = 0.$$

The flux splitting operators are given as,

$$\begin{cases} B_i(j)^+ = \frac{1}{2} \begin{pmatrix} -S_i^2(j) & -S_i(j) \\ S_i(j) & -S_i^2(j) \end{pmatrix} \\ B_i(j)^- = \frac{1}{2} \begin{pmatrix} S_i^2(j) & -S_i(j) \\ S_i(j) & S_i^2(j) \end{pmatrix} \end{cases} \quad (11)$$

where the matrix operator  $S_i(j)$  applied to an arbitrary vector  $\mathbf{a}$  produces the cross-product of the outward normal  $\hat{\mathbf{n}}_i(j)$ , the normal to the  $j$ -th facet of element  $i$ , with  $\mathbf{a}$ , that is,  $S_i(j)\mathbf{a} = \hat{\mathbf{n}}_i(j) \times \mathbf{a}$ .

### B. Expressing $(L\mathbf{u}, \mathbf{u})$ and $(L\mathbf{u}, L\mathbf{u})$ for FVTD

For simplicity, we now consider only the case when we have the same  $\epsilon$  and  $\mu$  for all elements. For the case when we have non-uniform  $\epsilon$  and  $\mu$  the time-step will be changed in the same way as in [1] according to the speed of light in the element ( $c_i = (\epsilon_i\mu_i)^{-1/2}$ ). We write,

$$(L\mathbf{u}, \mathbf{u}) = \frac{1}{2} \sum_{i=1}^N \sum_{k=1}^{m_i} A_i(k) \times [T_i^+(k)B_i^+(k)\mathbf{u}_i + T_i^-(k)B_i^-(k)\mathbf{u}_{i_k}] \cdot \mathbf{u}_i \quad (12)$$

$$(L\mathbf{u}, L\mathbf{u}) = \frac{1}{2} \sum_{i=1}^N \frac{c}{V_i} \cdot \left\{ \sum_{k=1}^{m_i} A_i(k) (T_i^+(k)B_i^+(k)\mathbf{u}_i + T_i^-(k)B_i^-(k)\mathbf{u}_{i_k}) \cdot \sum_{j=1}^{m_i} A_i(j) (T_i^+(j)B_i^+(j)\mathbf{u}_i + T_i^-(j)B_i^-(j)\mathbf{u}_{i_j}) \right\}. \quad (13)$$

The flux-splitting operators,  $B_i^+(k)$  and  $B_i^-(k)$ , when applied to the field value at the center of an element give the flux at facet  $k$  which when summed over all facets of the element give zero [2]. Thus, we have

$$\sum_{k=1}^{m_i} A_i(k) T_i^+(k) B_i^+(k) \mathbf{u}_i + A_i(k) T_i^-(k) B_i^-(k) \mathbf{u}_i = 0,$$

which can be written as,

$$\sum_{k=1}^{m_i} A_i(k) B_i^+(k) \mathbf{u}_i = - \sum_{k=1}^{m_i} A_i(k) B_i^-(k) \mathbf{u}_i \quad (14)$$

Combining equation (13) with equation (14) allows us to write

$$(L\mathbf{u}, L\mathbf{u}) = \frac{1}{2} \sum_{i=1}^N \frac{c}{V_i} \left[ \sum_{k=1}^{m_i} A_i(k) B_i^-(k) (\mathbf{u}_{i_k} - \mathbf{u}_i) \right] \cdot \left[ \sum_{j=1}^{m_i} A_i(j) B_i^-(j) (\mathbf{u}_{i_j} - \mathbf{u}_i) \right],$$

whereas combining equation (12) with equation (14) gives

$$(L\mathbf{u}, \mathbf{u}) = \frac{1}{2} \sum_{i=1}^N \mathbf{u}_i \cdot \sum_{k=1}^{m_i} A_i(k) (B_i^-(k) (\mathbf{u}_{i_k} - \mathbf{u}_i)).$$

For our case, because  $S = -S^T$ , we have  $B = B^T$  we can write

$$B\mathbf{a} \cdot \mathbf{b} = \frac{1}{2} B\mathbf{b} \cdot \mathbf{b} + \frac{1}{2} B\mathbf{a} \cdot \mathbf{a} - \frac{1}{2} B(\mathbf{b} - \mathbf{a}) \cdot (\mathbf{b} - \mathbf{a}),$$

and therefore

$$\begin{aligned} & \mathbf{u}_i \cdot \sum_{k=1}^{m_i} A_i(k) (B_i^-(k) (\mathbf{u}_{i_k} - \mathbf{u}_i)) = \\ & \frac{1}{2} \sum_{k=1}^{m_i} A_i(k) [\mathbf{u}_{i_k} \cdot B_i^-(k) \mathbf{u}_{i_k} - \mathbf{u}_i \cdot B_i^-(k) \mathbf{u}_i - \\ & (\mathbf{u}_{i_k} - \mathbf{u}_i) \cdot B_i^-(k) (\mathbf{u}_{i_k} - \mathbf{u}_i)]. \end{aligned}$$

### C. Expressions for the maximum energy criterion

Introducing the expressions for  $(L\mathbf{u}, \mathbf{u})$  and  $(L\mathbf{u}, L\mathbf{u})$  in to formula (9) we get,

$$\begin{aligned} & \Delta t \sum_{i=1}^N \frac{c}{V_i} \left[ \sum_{k=1}^{m_i} A_i(k) B_i^-(k) (\mathbf{u}_{i_k} - \mathbf{u}_i) \right] \cdot \\ & \left[ \sum_{j=1}^{m_i} A_i(j) B_i^-(j) (\mathbf{u}_{i_j} - \mathbf{u}_i) \right] \leq \\ & \sum_{i=1}^N \sum_{k=1}^{m_i} A_i(k) [\mathbf{u}_{i_k} \cdot B_i^-(k) \mathbf{u}_{i_k} - \\ & \mathbf{u}_i \cdot B_i^-(k) \mathbf{u}_i - (\mathbf{u}_{i_k} - \mathbf{u}_i) \cdot \\ & B_i^-(k) (\mathbf{u}_{i_k} - \mathbf{u}_i)]. \end{aligned} \quad (15)$$

It is also easy to check that the following is true,

$$\begin{aligned} & \sum_{i=1}^N \sum_{k=1}^{m_i} A_i(k) [\mathbf{u}_{i_k} \cdot B_i^-(k) \mathbf{u}_{i_k} - \\ & \mathbf{u}_i \cdot B_i^-(k) \mathbf{u}_i] = \\ & - \sum_{i=1}^{N_b} \sum_{k=1}^{m_i^s} A_i(k) \mathbf{u}_i \cdot B_i^-(k) \mathbf{u}_i \geq 0 \end{aligned} \quad (16)$$

where  $N_b$  is the number of elements with facets on the domain boundary, and  $m_i^s$  is the number of facets of  $i$ -th element on that boundary. In the case if  $k$ -th facet of  $i$ -th element is PEC facet, we can write the terms in the sum on the left side of formula (16) as,

$$\begin{aligned} & \mathbf{u}_{i_k} \cdot B_i^-(k) \mathbf{u}_{i_k} - \mathbf{u}_i \cdot B_i^-(k) \mathbf{u}_i = \\ & = \begin{bmatrix} \mathbf{E}_i^n \\ \mathbf{H}_i^n \end{bmatrix} \frac{1}{2} \begin{pmatrix} S_i^2(j) & -S_i(j) \\ S_i(j) & S_i^2(j) \end{pmatrix} \begin{bmatrix} \mathbf{E}_i^n \\ \mathbf{H}_i^n \end{bmatrix} - \\ & \begin{bmatrix} -\mathbf{E}_i^n \\ \mathbf{H}_i^n \end{bmatrix} \frac{1}{2} \begin{pmatrix} S_i^2(j) & -S_i(j) \\ S_i(j) & S_i^2(j) \end{pmatrix} \begin{bmatrix} -\mathbf{E}_i^n \\ \mathbf{H}_i^n \end{bmatrix} = 0. \end{aligned} \quad (17)$$

Therefore, in equation (16) the PEC facets interior to the mesh do not contribute to the inequality and the inequality remains the same. Thus, using equation (16), we can rewrite equation (15) as a non-tight bound,

$$\begin{aligned} & \sum_{i=1}^N \frac{c\Delta t}{V_i} \left[ \sum_{k=1}^{m_i} A_i(k) B_i^-(k) (\mathbf{u}_{i_k} - \mathbf{u}_i) \right] \cdot \\ & \left[ \sum_{j=1}^{m_i} A_i(j) B_i^-(j) (\mathbf{u}_{i_j} - \mathbf{u}_i) \right] \leq \\ & - \sum_{i=1}^N \sum_{k=1}^{m_i} A_i(k) (\mathbf{u}_{i_k} - \mathbf{u}_i) \cdot \\ & B_i^-(k) (\mathbf{u}_{i_k} - \mathbf{u}_i) \leq 2(L\mathbf{u}, \mathbf{u}). \end{aligned} \quad (18)$$

This is the fundamental global inequality that imposes the stability constraint on  $\Delta t$ . If not for the ABC facets this would be a tight bound which implies that the inequality (18) defines the necessary condition on  $\Delta t$  for the FVTD algorithm on an infinite domain or inside a PEC enclosure.

It is not a simple task to derive a global constraint on  $\Delta t$  based on this formula. Therefore, we have to make due with imposing this inequality locally on a finite-volume by finite-volume basis. This removes the summation over all finite-volumes, but leaves the inner summations. Note that on a uniform mesh dealing only with the inner summation keeps the inequality (18) exact.

In order to get a manageable constraint for  $\Delta t$ , even limiting ourselves to a local constraint, requires that we somehow remove the inner summations over facets while keeping the formula exact. We proceed by first constructing a block-diagonal matrix  $Z_i = \text{diag}\{-B_i^-(k)\}_{k=1}^{m_i}$  as well as a block-row vector of  $m_i$  identity matrices  $W = \{I, \dots, I\}$ , where the dimension of  $W$  is  $6 \times 6m_i$  and  $I$  is the  $6 \times 6$  identity matrix. We also construct a column vector made up of the solution vector differences across each facet,

$$\mathbf{x} = \text{vector}\{\mathbf{u}_{i_k} - \mathbf{u}_i\}_{k=1}^{m_i}$$

which is a vector of length  $6m_i$ . Hence a new local inequality, based on inequality (18), can be written concisely using these constructions as,

$$\frac{c\Delta t}{V_i} (W Z_i \mathbf{x}, W Z_i \mathbf{x}) \leq (Z_i \mathbf{x}, \mathbf{x}), \quad \forall i. \quad (19)$$

#### D. Efficient computation of stability criterion

Using the property  $B_i^-(k)^T = B_i^-(k)$ , we have also  $Z_i = Z_i^T$ . Hence the square-root of  $Z_i$  can be expressed as

$$Z_i = Q^T \Lambda_i Q = Q^T \sqrt{\Lambda_i} Q \quad Q^T \sqrt{\Lambda_i} Q = \sqrt{Z_i} \sqrt{Z_i},$$

where  $\Lambda_i = \text{diag}\{\lambda_j\}_{j=1}^{6m_i}$  is the diagonal matrix of eigenvalues of  $Z_i$ , and

$$\sqrt{\Lambda_i} = \text{diag}\{\sqrt{\lambda_j}\}_{j=1}^{6m_i}.$$

Now defining the new variable  $\mathbf{y} = \sqrt{Z_i} \mathbf{x}$ , which means that  $\sqrt{Z_i} \mathbf{y} = \mathbf{x}$  and therefore  $\mathbf{y} \notin \text{Ker}(\sqrt{Z_i})$ , we

have equation (19) rewritten as

$$\frac{c\Delta t}{V_i} (W \sqrt{Z_i} \mathbf{y}, W \sqrt{Z_i} \mathbf{y}) \leq (\mathbf{y}, \mathbf{y}),$$

or

$$\frac{V_i}{c\Delta t} \geq \frac{(W \sqrt{Z_i} \mathbf{y}, W \sqrt{Z_i} \mathbf{y})}{(\mathbf{y}, \mathbf{y})}, \quad \forall i.$$

Finally, this last inequality can be written in terms of the original summations over the facets as,  $\forall i$

$$\begin{aligned} \frac{V_i}{c\Delta t} & \geq \left( \sum_{k=1}^{m_i} (\mathbf{y}_k, \mathbf{y}_k) \right)^{-1} \times \\ & \left[ \sum_{k=1}^{m_i} \sqrt{A_i(k)} \sqrt{-B_i^-(k)} \mathbf{y}_k \right] \cdot \\ & \left[ \sum_{j=1}^{m_i} \sqrt{A_i(j)} \sqrt{-B_i^-(j)} \mathbf{y}_j \right]. \end{aligned} \quad (20)$$

From the above equation we can evaluate the maximum  $\Delta t$  for stability by numerically evaluating the right hand side over all elements. This calculation can be simplified considerably if we use the property that  $S_i^3(k) = -S_i(k)$  and notice that

$$\begin{aligned} & [-B_i^-(k)]^2 = \\ & \frac{1}{4} \begin{pmatrix} S_i^2(k) & -S_i(k) \\ S_i(k) & S_i^2(k) \end{pmatrix} \times \begin{pmatrix} S_i^2(k) & -S_i(k) \\ S_i(k) & S_i^2(k) \end{pmatrix} = \\ & \frac{1}{2} \begin{pmatrix} -S_i^2(k) & S_i(k) \\ -S_i(k) & -S_i^2(k) \end{pmatrix} = -B_i^-(k), \end{aligned}$$

hence  $\sqrt{-B_i^-(k)} = -B_i^-(k)$ . This means that the eigenvalues of the operator  $-B_i^-(k)$  are only 1, with multiplicity 2, and 0, with multiplicity 4. Now the eigenvector  $\hat{\mathbf{e}}_k$  of  $-B_i^-(k)$  which corresponds to the eigenvalue 1 can be written (without facet index  $k$ ) as

$$\begin{pmatrix} \hat{\mathbf{a}} \\ \hat{\mathbf{b}} \end{pmatrix} = \frac{1}{2} \begin{pmatrix} -S^2 & S \\ -S & -S^2 \end{pmatrix} \begin{pmatrix} \hat{\mathbf{a}} \\ \hat{\mathbf{b}} \end{pmatrix}, \quad \begin{aligned} 2\hat{\mathbf{a}} &= -S^2 \hat{\mathbf{a}} + S \hat{\mathbf{b}} \\ S \hat{\mathbf{a}} &= -2\hat{\mathbf{b}} - S^2 \hat{\mathbf{b}} \end{aligned}$$

This is satisfied for any  $\hat{\mathbf{a}} = S \hat{\mathbf{b}} = \hat{\mathbf{n}} \times \hat{\mathbf{b}}$ . Normalizing the eigenvector for any facet  $k$  we have  $\hat{\mathbf{e}}_k = \sqrt{2}^{-1} (\hat{\mathbf{n}}_k \times \hat{\mathbf{b}}_k, \hat{\mathbf{b}}_k)^T$ , where we choose  $\hat{\mathbf{b}}_k$  as an arbitrary vector in the plane of the  $k$ -th facet (i.e.,  $\hat{\mathbf{n}}_k \cdot \hat{\mathbf{b}}_k = 0$ ). If we choose two orthogonal vectors  $\hat{\mathbf{b}}_k^1$  and  $\hat{\mathbf{b}}_k^2$  on the facet using the formula  $\hat{\mathbf{b}}_k^2 = \hat{\mathbf{n}}_k \times \hat{\mathbf{b}}_k^1$  we can write two eigenvectors for the facet as

$$\begin{aligned} \hat{\mathbf{e}}_k^1 &= \frac{1}{\sqrt{2}} (\hat{\mathbf{n}}_k \times \hat{\mathbf{b}}_k, \hat{\mathbf{b}}_k)^T, \\ \hat{\mathbf{e}}_k^2 &= \frac{1}{\sqrt{2}} (-\hat{\mathbf{b}}_k, \hat{\mathbf{n}}_k \times \hat{\mathbf{b}}_k)^T. \end{aligned}$$

The four eigenvectors corresponding to the zero eigenvalue of  $-B_i^-(k)$  can be written as

$$\begin{aligned} \hat{\mathbf{e}}_k^{3,4} &= \frac{1}{\sqrt{2}} (\hat{\mathbf{n}}_k, \pm \hat{\mathbf{n}}_k)^T, \\ \hat{\mathbf{e}}_k^{5,6} &= \frac{1}{\sqrt{2}} (-\hat{\mathbf{n}}_k \times \hat{\mathbf{b}}_k^{1,2}, \hat{\mathbf{b}}_k^{1,2})^T. \end{aligned}$$

Thus, with the single vector  $\hat{\mathbf{b}}_k$  in the plane of the  $k^{th}$  facet we can define all eigenvectors of the operator  $\sqrt{-B_i^-(k)}$ .

To efficiently evaluate formula (20) as in [2] we can decompose each  $\mathbf{y}_k$  as a sum of eigenvectors, with the only ones taking part being the ones corresponding to non-zero eigenvalues  $\sqrt{\lambda_i(k)} = 1$  of  $\sqrt{-B_i^-(k)}$

$$\mathbf{y}_k = \sum_{s=1}^2 w_k^s \hat{\mathbf{e}}_k^s.$$

The value of  $\hat{\mathbf{b}}_k^1$  can be chosen arbitrarily, for example the edge of the facet. Substituting this decomposition into the inequality (20) we get the formula

$$\frac{V_i}{c\Delta t} \geq \left( \sum_{k=1}^{m_i} \sum_{s=1}^2 (w_k^s)^2 \right)^{-1} \times \left[ \sum_{k=1}^{m_i} \sqrt{A_i(k)} \sum_{s=1}^2 w_k^s \hat{\mathbf{e}}_k^s \right] \cdot \left[ \sum_{j=1}^{m_i} \sqrt{A_i(j)} \sum_{s=1}^2 w_j^s \hat{\mathbf{e}}_j^s \right],$$

simplifying we get the formula which is used to obtain the time-step limit

$$\frac{V_i}{c\Delta t} \geq \frac{\sum_{k=1}^{m_i} \sum_{j=1}^{m_i} \sqrt{A_i(k)A_i(j)} \sum_{s_1=1}^2 \sum_{s_2=1}^2 \hat{\mathbf{e}}_j^{s_1} \cdot \hat{\mathbf{e}}_k^{s_2} w_j^{s_1} w_k^{s_2}}{\sum_{k=1}^{m_i} \sum_{s=1}^2 (w_k^s)^2}.$$

This can be written concisely as,

$$\frac{V_i}{c\Delta t} \geq \frac{(G_i \mathbf{w}, G_i \mathbf{w})}{(\mathbf{w}, \mathbf{w})} \quad (21)$$

where the elements of the matrix  $G_i \in \mathbb{R}^{2m_i \times 2m_i}$  are written as

$$G_i = \left[ g_j^i g_k^i \hat{\mathbf{e}}^j \cdot \hat{\mathbf{e}}^k \right]_{j,k=1}^{2m_i}, \quad g_{2k-1}^i = g_{2k}^i = \sqrt{A_i(k)},$$

$$\hat{\mathbf{e}}^{2k} = \hat{\mathbf{e}}_k^2, \quad \hat{\mathbf{e}}^{2k-1} = \hat{\mathbf{e}}_k^1; \quad k = 1..m_i.$$

Using this notation the stability criterion becomes

$$\Delta t \leq \min_i \frac{V_i}{c \|G_i\|}. \quad (22)$$

As in [1] this can be generalized concisely for elements with an individual  $\epsilon_i$  and  $\mu_i$  as

$$\Delta t \leq \min_i \frac{V_i}{c_i \|G_i\|}, \quad (23)$$

where  $c_i = (\epsilon_i \mu_i)^{0.5}$ .

The norm of matrix  $G_i$  can be computed relatively quickly because  $4 \leq m_i \leq 6$  for a cell-centered FVTD mesh which contains tetrahedrons, prisms, pyramids and hexahedrons.

## IV. NUMERICAL EXPERIMENTS

To test the increase in the allowed time-step due to our new limit we conducted a wide set of numerical experiments on our cell-centered FVTD code [4]. These were conducted for both unstructured as well as structured meshes. For unstructured tetrahedral mesh we had a 5-15% increase in the allowable time-step over the time-step limit given by formula (2) from [1]. For a structured cubical mesh of equation (23) gives the same result as the von Neumann method applied to the FVTD approximation of Maxwell's equations [2],

$$\Delta t \leq \frac{h}{2c}.$$

This is a 1.5 times larger time-step that allowed by (2) when it is applied to a structured cubical mesh ( $h$  taken as the cubical element edge size).

## V. SCATTERING FROM A PEC SPHERE

We present the FVTD results for scattering from a perfectly electrical conducting (PEC) sphere as an exact series solution is available in the frequency domain [5], and a time domain solution may be easily obtained using the inverse Fourier transform. The geometry of the problem is shown in Fig. 1. The radius of sphere is 3 meters. This problem was selected as a benchmark for the FVTD engine as the irregular surface of the sphere coincides with one of the primary reasons for developing finite-volume methods on irregular grids: eliminating the need for stair-stepping at arbitrarily shaped boundaries. The average edge-length of the tetrahedrons for the results shown in the figure set to 0.75 m.

An  $x$ -polarized electric-field plane-wave transient pulse  $\mathbf{E} = g(t)\hat{\mathbf{x}}$  incident in the  $z$ -direction and varying as the derivative of a Gaussian was selected where, for  $t \geq 0$ ,

$$g(t) = -2A(t-t_0)b^{-2} \exp(-b^{-2}(t-t_0)^2) \quad (24)$$

with  $A = 1$ ,  $b = 1.14 \times 10^{-8}$ s,  $t_0 = 4.0 \times 10^{-8}$ . These give a shortest free-space wavelength of about 3 meters resulting in significant energy in the resonance region of the sphere.

The results in the right side of the figure were computed using the second-order MUSCL-type scheme for two scattering locations: side-scatter and back-scatter (see Fig. 1). These are compared to the analytic solution for these same two points. The FVTD results shown were computed for the time-step given by equation (2), but almost identical results were obtained when the time-step was set to that of equations (1) and (23). Note that all the time-step criteria discussed in this paper are for the firstorder upwind schemes, but we have found that using any of these criteria for the higher-order schemes gives stable results. This was also reported by Piperno [1]. Table 1. shows the relative error of computational results for

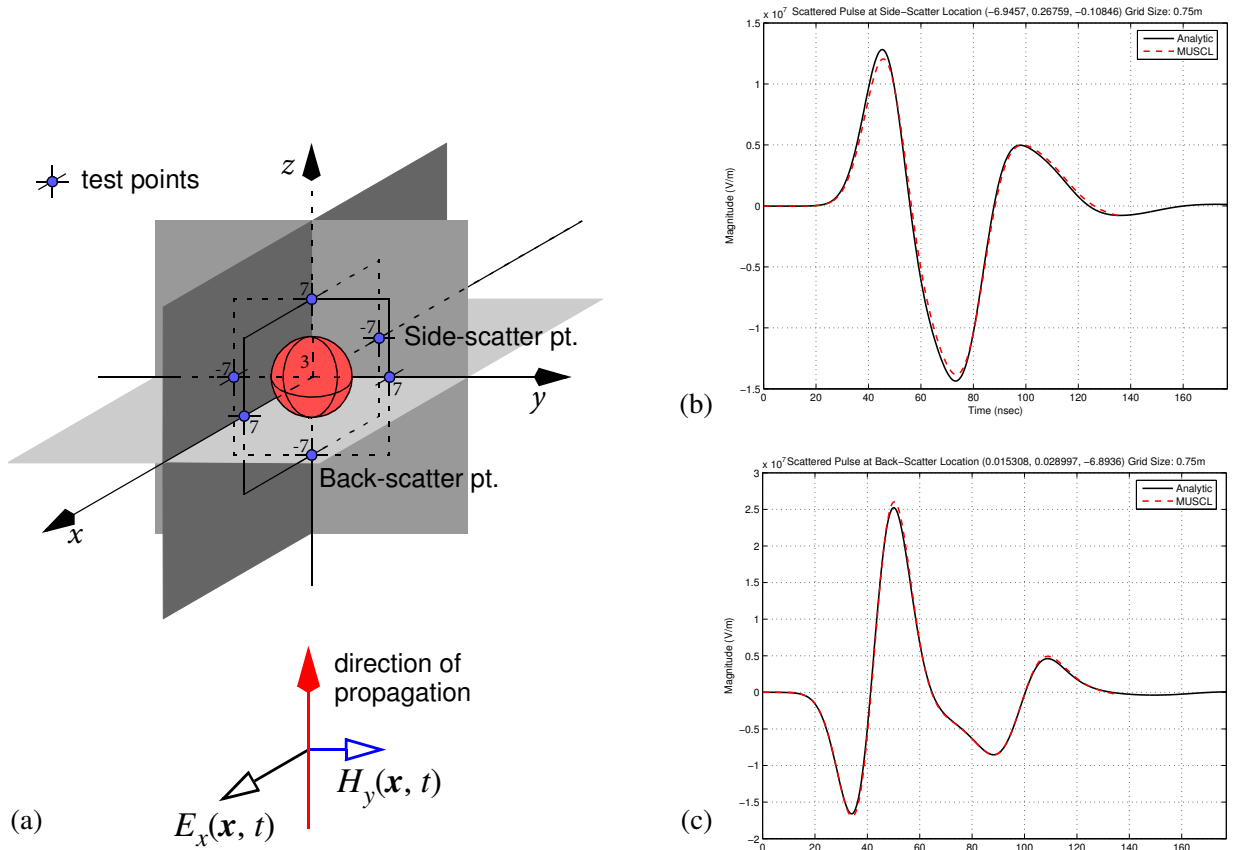


Fig. 1. Scattering from a sphere. (a) Geometry of the problem, (b) Side-scattered electric field  $E_x$ , and (c) back-scattered electric field  $E_x$ . Time-step used was that of equation (23). Analytic results are also shown.

Table 1. Comparison of PEC sphere back-scattering and side-scattering results on time interval  $(0, 1.4 \times 10^7)$ .

Time step algorithm	# of time steps (acceleration factor)	Side Scatter		Back Scatter	
		$L_2$ Error of $E_x$	$L_\infty$ Error of $E_x$	$L_2$ Error of $E_x$	$L_\infty$ Error of $E_x$
Eq. (1)	1497 (1.0)	4.15%	4.49%	4.37%	4.21%
Eq. (2)	749 (2.0)	4.52%	4.19%	4.23%	3.99%
Eq. (21)	704 (2.13)	4.63%	4.47%	4.25%	4.50%

the  $x$ -component of the scattered field at the back-scatter location  $(0, 0, -7)$ , as well as at the side-scatter location  $(-7, 0, 0)$  for the three different time-step criteria: equations (1), (2) and (23). For both test locations, the analytic solution is compared to solutions obtained using the MUSCL finite-volume methods computed on a mesh with an average cell edge length of 0.75 m. All results shown use the second-order predictor-corrector time-integration scheme.

## VI. CONCLUSION

The derivation we have provided gives a new time-step limit for the explicit upwinding finite-volume time-domain approximation scheme on arbitrary unstructured meshes for Maxwell's equations. On uniform meshes the

new criterion is necessary for stability. In fact, the time-step limit provided by the formula given herein gives the same time-step limit on structured meshes as the standard von Neumann analysis. On mixed structured and unstructured meshes, the new criterion provides the maximum time-step allowable for retaining stability. The derivation can be easily extended to other FVTD approximations of partial differential equations on unstructured meshes; which is a subject of future work.

## ACKNOWLEDGMENTS

This work was supported by the Department of National Defence and the Natural Sciences and Engineering Research Council.

## REFERENCES

- [1] S. Piperno, “ $L^2$ -Stability of the upwind first order finite volume scheme for the Maxwell equations in two and three dimensions on arbitrary unstructured meshes”, *Mathematical Modeling and Numerical Analysis*, vol. 34, no. 1, pp. 139-158, 2000.
- [2] P. Bonnet, X. Ferrieres, P. L. Michielsen, and P. Klotz, “Finite volume time domain method” in “Time domain electromagnetics”, S. M. Rao (editor), Academic Press, San Diego, 1999.
- [3] S. Gottlieb, C. W. Shu, and E. Tadmor, “Strong stability – preserving high – order time discretization methods”, *SIAM Riview*, vol. 43, no. 1, pp. 89–112, 2001.
- [4] D. Firsov, J. LoVetri, I. Jeffrey, V. Okhmatovski, C. Gilmore, and W. Chamma, “High-order FVTD on unstructured grids using an object-oriented computational engine”, *ACES Journal* vol. 22, no.1, pp. 71-82, March 2007.
- [5] R. F. Harrington, “Time-harmonic electromagnetic fields”, McGraw-Hill, 1961.



**Dmitry K. Firsov** received his diploma (with distinction) and Candidate of Science (Ph.D.) degrees in mathematics from Tomsk State University, Russia, in 1999 and 2002 respectively. In 2002 he was an Assistant Professor at Tomsk State University, Department of Theoretical Mechanics. From 2003-2005 he was a post-doctoral fellow with the Department of Mathematics at the University of Manitoba, Canada. Currently, Dmitry K. Firsov is a Research Associate at the Department of Electrical Engineering at the University of Manitoba. He is currently working in the area of numerical methods for partial differential equations.



**Joe LoVetri** received the Ph.D. degree in electrical engineering from the University of Ottawa, ON, Canada, in 1991. From 1991 to 1999, he was an Associate Professor in the Department of Electrical and Computer Engineering, University of Western Ontario, Canada. He is currently a Professor in the Department of Electrical and Computer Engineering, and Associate Dean (Research) of the Faculty of Engineering at University of Manitoba, Winnipeg, Canada. His main research interests are in time-domain CEM, modeling of EMC problems, biomedical imaging, and inverse scattering techniques.

# A General Framework for Mixed Structured/Unstructured PEEC Modelling

F. Freschi and M. Repetto

Department of Electrical Engineering, Politecnico di Torino  
c.so Duca degli Abruzzi, 24 - 10129 Torino, Italy  
fabio.freschi@polito.it, maurizio.repetto@polito.it

**Abstract** – The paper proposes a re-formulation of PEEC modeling in terms of dual discretization of surface conductors. The result is a precise formalization of the PEEC tessellation and gives the possibility of handling general meshes, both structured and unstructured. This fact allows the use of triangular meshes only where necessary, leaving orthogonal discretization whenever possible and improving computational performances. The method, referred to as dual-PEEC, is validated by solving a structure referenced in literature and finally applied to the computation of the input impedance of a spiral inductor.

**Keywords:** PEEC, unstructured mesh, and dual discretization.

## I. INTRODUCTION

Integral approaches to Maxwells equations for the modeling of interconnects and packaging structures are usually preferred to differential ones, because they solve the field equation in terms of the sources of fields, located on the metallic structures. This allows to discretize the conducting part of the domain only, accounting rigorously for the regularity conditions of fields at infinity.

Among other integral techniques, the Partial Element Equivalent Circuit (PEEC) leads to the description of the electromagnetic problems in terms of circuit parameters which describe the magnetic and electric coupling between currents and charges of the discretized structure. The main advantages of PEEC models are the possibility of integrating lumped parameters into the electromagnetic structure and the availability of both time and frequency analysis in standard SPICE-like solvers.

The classic PEEC approach is restricted to orthogonal discretization of the structures under study. Recently the scheme has been extended to non-orthogonal [1, 2] and triangular [3, 4] meshes, but the proposed techniques lack of a systematic background for their assumptions.

The aim of this work is to generalize the PEEC methodology by the use of the topological concept of *duality*. The result is the exact formalization of the PEEC tessellation gaining in addition the possibility of handling general meshes, both structured and unstructured. This fact allows the use of triangular meshes only where necessary (i.e., curved or slanting electrodes), leaving

orthogonal discretization in regular regions, with a global improvement of computational performances.

The paper is structured as follows: in Section II the original PEEC scheme is briefly described; Section III gives a general overview of dual discretization which is used as a framework of the mathematical formulation of Section IV. Two examples are investigated in Section V and finally Section VI draws some conclusions.

## II. ORIGINAL PEEC SCHEME

The first systematic description of the Partial Element Equivalent Circuit, PEEC, goes back to the middle 70s, in a work by A.E. Ruehli [5]. The original PEEC formulation is based on the circuit interpretation of the terms of the Mixed Potential Integral Equation (MPIE),

$$\frac{\vec{J}(\vec{r}, t)}{\sigma} + \frac{d\vec{A}(\vec{r}, t)}{dt} + \nabla\varphi(\vec{r}, t) = 0 \quad (1)$$

$\vec{A}$  and  $\varphi$  are the retarded magnetic vector and electric scalar potential, respectively,

$$\vec{A}(\vec{r}, t) = \frac{\mu}{4\pi} \int_{\Omega} \frac{\vec{J}(\vec{r}', t')}{|\vec{r} - \vec{r}'|} d\Omega \quad (2)$$

$$\varphi(\vec{r}, t) = \frac{1}{4\pi\epsilon} \int_{\Omega} \frac{\rho(\vec{r}', t')}{|\vec{r} - \vec{r}'|} d\Omega, \quad (3)$$

where  $t'$  is the retardation time,

$$t' = t - \frac{|\vec{r} - \vec{r}'|}{c}.$$

The solution of the electromagnetic problem is achieved by discretizing equation (1) on the system of surface cells shown in Fig. 1(a). Current and charge densities are expanded in terms of piecewise constant pulse functions. The same functions are also used as test functions in a Galerkin weighted residual scheme [6].

- Partial self and mutual inductances between nodes, representing the magnetic field couplings;
- Resistive terms between nodes, representing the Joule losses in the conductive material;
- Coefficients of potential at each node, representing the electric field couplings.

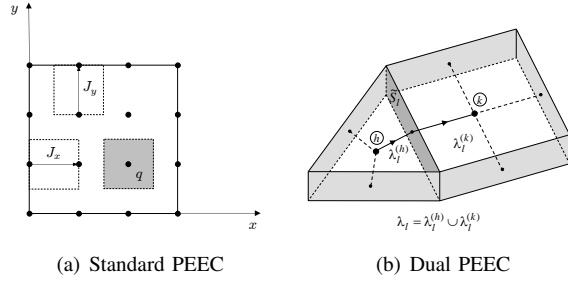


Fig. 1. (a) Standard 2D discretization of current density (dashed cells) and charge density (gray cell), (b) dual discretization: the couple *primal edge* (straight line)/*dual face* (dark gray) correspond to a two terminal circuit component.

### III. DUAL DISCRETIZATION

The work of E. Tonti has shown how the formulation of the electromagnetic field can be expressed in terms of dual relations and how this duality can be exploited in its numerical solution [7]. The conceptual framework of the finite formulation of electromagnetic fields is based on the definition of two sets of global variables: *source* and *configuration*. In the present application *source variables* are electric charge and electric current. *Configuration variables* are voltage, electric potential, magnetic vector potential. The product of a source variable by a configuration variable gives an *energy variable*. Physical variables are associated to spatial elements. In [8] it is shown that configuration variables are associated with space elements endowed with inner orientation (i.e., the orientation of the space element lies on the element itself); on the contrary source variables are associated with space elements endowed with outer orientation (i.e., the orientation of the space element depends on the space in which the element is embedded). The complex of cell with inner orientation is called *primal complex*, whereas the *dual complex* is endowed with outer orientation. Primal and dual complexes of cells are interlinked: primal nodes correspond to dual volumes, primal edges to dual faces, primal faces to dual edges and primal volumes to dual nodes. Figure 2 shows the generation of primal and dual complex and their spatial duality, assuming primal complex as simplicial. It is important to note that it is possible to assume the dual complex as simplicial, when source variables are chosen as unknowns (in this work the latter representation is adopted).

Table 1 reports the spatial assignment of the variables used in PEEC modeling: scalar potentials are assigned to primal nodes, voltage drops to primal edges, currents to dual faces and charges to dual volumes. This definition of duality relations of grids and the rigorous assignment of physical variables to spatial entities allows the possibility of giving a general framework of PEEC modeling, i.e., allowing the use of unstructured and mixed structured/unstructured complexes of cells. Besides the

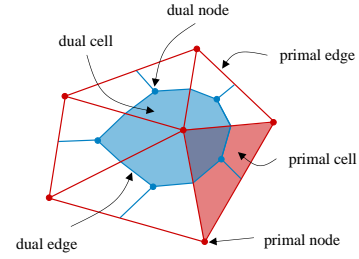


Fig. 2. 2D primal and dual complex.

generality of dual discretization, in this paper only surface discretization is investigated, leaving the generalization to volume discretization to a further work.

Making reference to Fig. 1(b) the discretization of conductors is made by prisms with triangular or rectangular basis. This tessellation constitutes the *dual complex* of cells. The *primal complex* is obtained by connecting the centroids of dual volumes by piecewise segments crossing the lateral faces of prisms.

Table 1. Variable assignment to spatial elements.

Type	Variable	Spatial Element
source	current, $i$	dual face, $\tilde{\Sigma}$
	charge, $q$	dual volume, $\tilde{\Omega}$
configuration	voltage, $u$	primal edge, $\lambda$
	potential, $\varphi$	primal node $P$

### IV. MATHEMATICAL FORMULATION

#### A. Basis Functions

When using a general approach to PEEC models, a local interpolation inside dual volumes linking the current density  $\vec{J}$  in equation (1) to the global variable  $i$  through dual faces is needed. By referring to the generic dual volume or cell  $k$  identified by a prism of thickness  $\delta_k$ , and assuming a uniform  $\vec{J}$  distribution over the electrode thickness, *facet* elements basis functions  $\vec{w}$  are chosen,

$$\vec{J}_k = \frac{1}{\delta_k} \sum_{m=1}^{N_F} i_{km} \vec{w}_{km} \quad (4)$$

where  $N_F$  is the number of dual faces in each cell (3 or 4 in our case),  $i_{km}$  is the current through the  $m$ th dual face of cell  $k$ . Many choices for  $\vec{w}$  are possible but div-conforming basis functions must be selected for edges belonging to triangles, rectangles or common between them [9].

With reference to a local frame  $(u, v, w)$  (Fig. 3(a)) the following facet functions for triangles are selected [10],

$$\vec{w}_k^3(u, v) = N_j(u, v) \nabla N_i(u, v) \times \vec{n} - N_i(u, v) \nabla N_j(u, v) \times \vec{n}$$



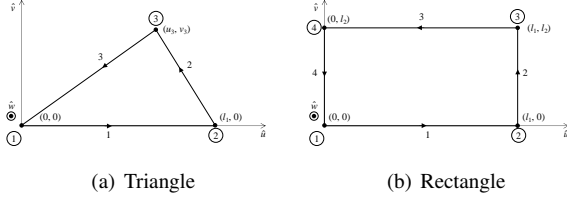


Fig. 3. Plane triangle and rectangle in local frame  $(u, v, w)$ .

where the functions  $N_i(u, v)$  and  $N_j(u, v)$  are the standard nodal functions related to the extreme nodes of edge  $k$  and  $\vec{n} = \hat{w}$  is the unit vector orthogonal to the triangle surface. The analytical expression of these shape functions is,

$$\begin{aligned}\vec{w}_1^3(u, v) &= \frac{u - u_3}{l_1 v_3} \hat{u} + \frac{v - v_3}{l_1 v_3} \hat{v} \\ \vec{w}_2^3(u, v) &= \frac{u}{l_1 v_3} \hat{u} + \frac{v}{l_1 v_3} \hat{v}, \\ \vec{w}_3^3(u, v) &= \frac{u - l_1}{l_1 v_3} \hat{u} + \frac{v}{l_1 v_3} \hat{v}.\end{aligned}$$

Figure 4(a) shows the vector plot of  $\vec{w}_1^3$  for a particular triangle in the  $x, y$  plane. For rectangular cells, special basis functions are built with the same properties of the ones for triangles.  $\vec{w}_k^4$  has only the component orthogonal to edge  $k$  and its magnitude linearly decrease from edge  $k$  to the opposite one. The analytic expression in the local  $(u, v, w)$  coordinate system is,

$$\begin{aligned}\vec{w}_1^4(u, v) &= \frac{v - l_2}{l_1 l_2} \hat{v} \\ \vec{w}_2^4(u, v) &= \frac{u}{l_1 l_2} \hat{u}, \\ \vec{w}_3^4(u, v) &= \frac{v}{l_1 l_2} \hat{v}, \\ \vec{w}_4^4(u, v) &= \frac{u - l_1}{l_1 l_2} \hat{u}.\end{aligned}$$

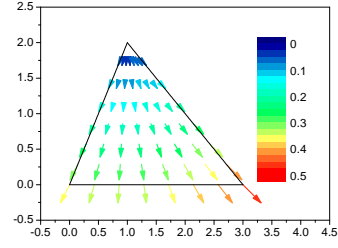
An example of a vector plot is reported in Fig. 4(b). It is easy to prove the continuity of the normal component of  $\vec{w}$  also for common edges shared by rectangles and triangles.

These basis functions have the property of being affine with respect to the  $u, v$  coordinates. This fact allows the possibility of using analytical formulas to solve the surface integrals described in the Section IV(b).

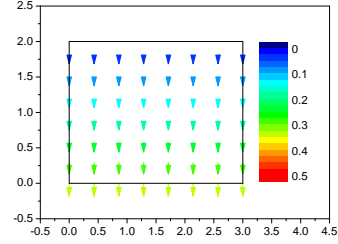
The surface charge density is expanded in terms of piecewise constant functions on each triangular/rectangular cell.

### B. Partial Element Calculation

The parameter extraction can be done by following the same rationale of the original PEEC formulation, i.e., testing MPIE equation (1) with the same basis function used to expand current and charge densities. An equivalent approach is here proposed, by making reference to energetic considerations.



(a) Triangle of vertices  $(0, 0), (3, 0), (1, 2)$



(b) Rectangle of vertices  $(0, 0), (3, 0), (3, 2), (0, 2)$

Fig. 4. Vector plot of facet element  $\vec{w}_1$  over (a) triangles and (b) rectangles.

1. *Extraction of resistances:* The Joule's power losses in the  $k$ th cell are,

$$\begin{aligned}P_k &= \int_{\Omega_k} \frac{J_k^2}{\sigma} d\Omega_k \\ &= \frac{1}{\delta_k^2 \sigma} \int_{\Omega_k} \left( \sum_{n=1}^{N_F} i_{kn} \vec{w}_{kn} \right) \cdot \left( \sum_{m=1}^{N_F} i_{km} \vec{w}_{km} \right) d\Omega_k.\end{aligned}$$

When only surface approximations are taken into account, the volume integral can be rearranged into a surface one,

$$P_k = \frac{1}{\delta_k \sigma} \int_{\Sigma_k} \left( \sum_{n=1}^{N_F} i_{kn} \vec{w}_{kn} \right) \cdot \left( \sum_{m=1}^{N_F} i_{km} \vec{w}_{km} \right) d\Sigma_k.$$

The power related to the current flowing through the generic  $\alpha$ th dual face of cell  $k$ , is,

$$P_{k\alpha} = i_{k\alpha} \sum_{m=1}^{N_F} \frac{i_{km}}{\delta_k \sigma} \int_{\Sigma_k} \vec{w}_{k\alpha} \cdot \vec{w}_{km} d\Sigma_k = i_{k\alpha} u_{k\alpha}$$

where

$$\begin{aligned}u_{k\alpha} &= \sum_{m=1}^{N_F} \frac{i_{km}}{\delta_k \sigma} \int_{\Sigma_k} \vec{w}_{k\alpha} \cdot \vec{w}_{km} d\Sigma_k \\ &= \sum_{m=1}^{N_F} R_{km} i_{km}\end{aligned}$$

$$R_{km} = \frac{1}{\delta_k \sigma} \int_{\Sigma_k} \vec{w}_{k\alpha} \cdot \vec{w}_{km} d\Sigma_k.$$

The voltage drop over half primal edge  $\alpha$  depends on currents through all dual faces of cell  $k$ .

2. *Extraction of partial inductances:* Resorting to magnetic vector potential defined in equation (2), the magnetic energy associated to  $k$ th cell is,

$$\begin{aligned} W_k &= \frac{1}{2} \int_{\Omega_k} \vec{J}_k \cdot \vec{A} d\Omega_k \\ &= \frac{1}{2} \frac{\mu_0}{4\pi} \int_{\Omega_k} \vec{J}_k \cdot \int_{\Omega_h} \sum_{h=1}^{N_{\bar{V}}} \frac{\vec{J}_h}{|\vec{r} - \vec{r}'|} d\Omega_h d\Omega_k \end{aligned}$$

where  $N_{\bar{V}}$  is the number of dual volumes. By using the expansion equation (4) and extracting only the magnetic coupling between the current  $\alpha$  of the  $k$ th cell, and current  $\beta$  of cell  $h$ , one has,

$$W_{k\alpha,h\beta} = \frac{1}{2} \frac{\mu_0}{4\pi} \int_{\Omega_k} \frac{i_{k\alpha} \vec{w}_{k\alpha}}{\delta_k} \cdot \int_{\Omega_h} \frac{i_{h\beta} \vec{w}_{h\beta}}{\delta_h |\vec{r} - \vec{r}'|} d\Omega_h d\Omega_k.$$

Finally, when considering surface complex of cells only, the partial inductance is,

$$\begin{aligned} W_{k\alpha,h\beta} &= \frac{1}{2} L_{k\alpha,h\beta} i_{k\alpha} i_{h\beta} \\ L_{k\alpha,h\beta} &= \frac{\mu_0}{4\pi} \int_{\Sigma_k} \vec{w}_{k\alpha} \cdot \int_{\Sigma_h} \frac{\vec{w}_{h\beta}}{|\vec{r} - \vec{r}'|} d\Sigma_h d\Sigma_k. \end{aligned} \quad (5)$$

3. *Extraction of coefficients of potential:* Making reference to equation (3), the electric energy related to the  $k$ th cell is,

$$\begin{aligned} W_k &= \frac{1}{2} \int_{\Omega_k} \rho_k \varphi d\Omega_k \\ &= \frac{1}{2} \frac{1}{4\pi\epsilon_0} \int_{\Omega_k} \rho_k \sum_{h=1}^{N_{\bar{V}}} \int_{\Omega_h} \frac{\rho_h}{|\vec{r} - \vec{r}'|} d\Omega_h d\Omega_k. \end{aligned}$$

With an approach similar to the one used for the extraction of partial inductances, it is possible to evaluate the electric energy due to the mutual coupling between the global charge  $q_k$  in cell  $k$  and  $q_h$  in cell  $h$ ,

$$W_{k,h} = \frac{1}{2} p_{k,h} q_k q_h$$

where

$$p_{k,h} = \frac{1}{4\pi\epsilon_0 \Sigma_k \Sigma_h} \int_{\Sigma_k} \int_{\Sigma_h} \frac{1}{|\vec{r} - \vec{r}'|} d\Sigma_h d\Sigma_k. \quad (6)$$

### C. Equivalent Circuit

Figure 5 shows the resulting two terminal component obtained by assembling all the previous contributions, and it is representative of the basic topological structure *primal edgeldual face* presented in Fig. 1(b). The unstructured PEEC two terminal components can be assembled to obtain the MNA matrix to be solved by a general purpose SPICE-like network simulator. With respect to standard PEEC model, the use of unstructured meshes introduces a local resistive mutual coupling. The use of dependent current sources allows to deal with this problem without increasing significantly the complexity of the circuit.

Figure 6 outlines the sequence of necessary steps for the application of dual-PEEC methodology to the solution of a full Maxwell problem. It is worth noting that the

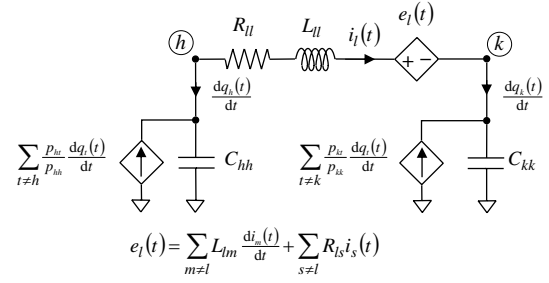


Fig. 5. Elementary branch of PEEC model.

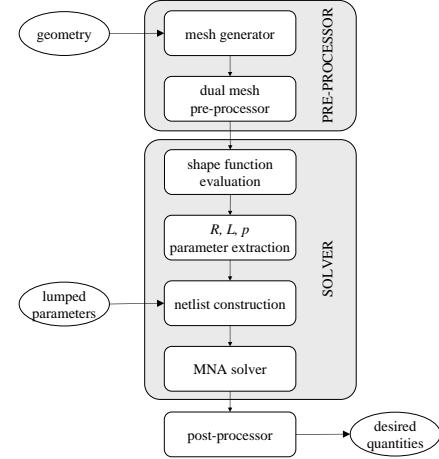


Fig. 6. Solver structure.

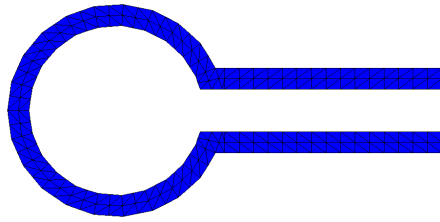
added complexity of the method due to the generation of the dual complex and the definition of basis function is negligible with respect to standard technique.

### D. Remarks

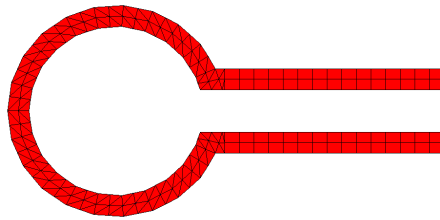
Equations (5) and (6) present a double surface integral, having the Green's function as kernel (in some cases multiplied by an affine combination of the  $(u, v)$  coordinates). To speed up the effort of the partial element computation, the inner integrals of equations (5) and (6) are solved by analytical formulas reported in [11, 12]. The outer integral is solved by using adaptive Gauss-Kronrod quadrature rule. The number of quadrature points is automatically chosen depending on the distance between cells: the larger the distance, the smaller the number of points needed to reach the convergence of the quadrature rule. Moreover, the computational time for the evaluation of inductance and coefficient of potential matrices can be reduced by considering that these matrices must be symmetric.

The standard PEEC formulation can be found as a particular case of the dual formulation when both charge and current densities are expanded in terms of piecewise constant pulse functions.

The resulting dual-PEEC circuit is similar to the standard PEEC model, with the exception of a local



(a) Triangular mesh



(b) Triangular/ quadrilateral mesh

Fig. 7. Loop antenna test case: (a) triangular discretization used as reference, (b) mixed triangular/quadrilateral discretization available with dual-PEEC modeling.

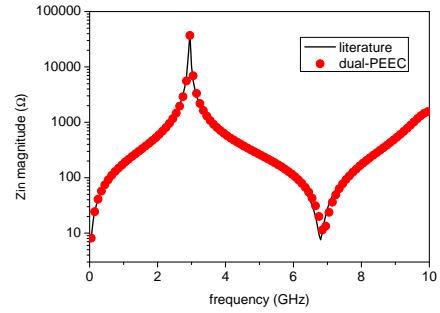
resistive coupling due to the use of unstructured meshes. The use of current controlled voltage sources allows to deal with this problem with a negligible computational effort (see Fig. 5) [13].

In frequency domain the inclusion of retardation is straightforward and makes PEEC equivalent to a full wave solution of Maxwell’s equations. It can be done by multiplying the non retarded mutual inductances and coefficients of potential by  $\exp(-j\omega\tau_{hk})$ , where  $\tau_{hk}$  is the delay time between the center coordinates of primal edges and dual volumes, respectively.

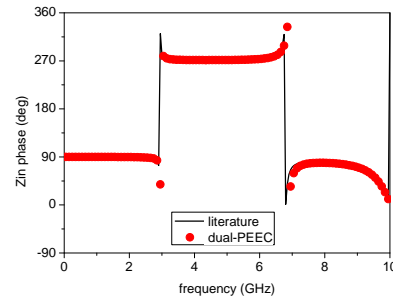
## V. EXAMPLES

### A. Model Validation

The dual-PEEC is first used to solve a benchmark loop antenna structure with known solution, already discussed in literature [14, 15]. Figure 7 shows the meshes adopted for the comparison: the structure of Fig. 7(a) is discretize by triangles and is solved by a FEM technique, whereas the mixed discretization of Fig. 7(b) is solved by the proposed procedure. Figure 8 compares the magnitude and phase of the input impedance evaluated in a full-wave regime. The results show a satisfactory agreement with respect to the literature ones.



(a) magnitude



(b) phase

Fig. 8. Comparison of input impedance evaluation: (a) magnitude, (b) phase.

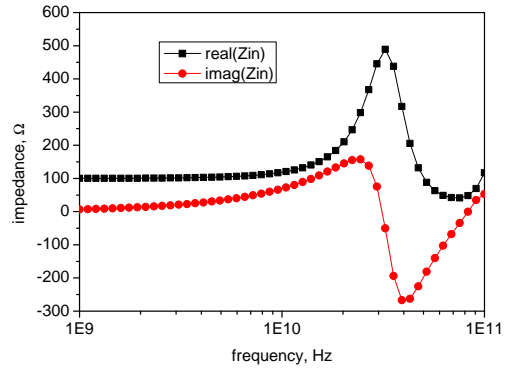


Fig. 9. Spiral inductor test case: input impedance.

### B. Spiral Inductor

The dual-PEEC technique is then used to solve the a surface metallic spiral inductor above a (finite) ground plane, ended by a lumped 100  $\Omega$  resistor. The input impedance (real and imaginary part) is reported in Fig. 9. Three different meshes are compared in Table 2 with respect to their geometrical parameters and computational time (in seconds on a Intel Centrino 1.5 GHz). The first mesh Fig. 10(a) is made by unstructured triangles (Mesh 1), the second one Fig. 10(b) with regular triangles (Mesh 2), the last one Fig. 10(c) with mixed triangles and

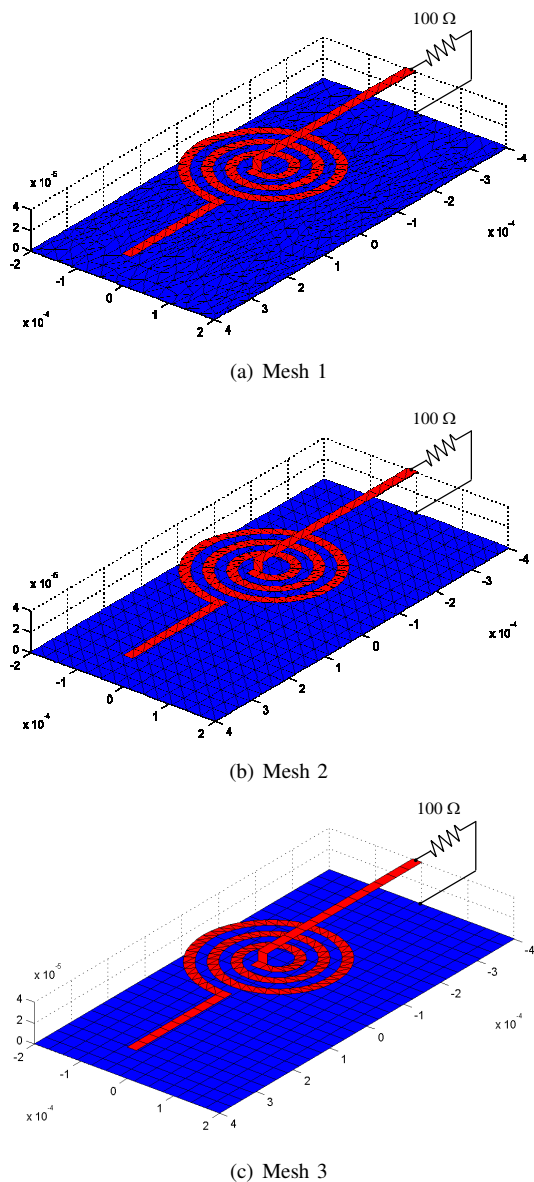


Fig. 10. Spiral inductor test case: (a) unstructured triangular mesh, (b) regular triangular mesh, (c) mixed triangular/rectangular mesh.

rectangles (Mesh 3). The three meshes share the same average edge length.

While no significant differences are observed in the values of impedance, the possibility of using together both structured and unstructured complexes of cells increases the computational efficiency of the method. In fact, given the maximum size of cell's edge (related to the shortest wavelength), the parameter extraction over rectangular dual is faster than over triangles and also the number of unknowns is lower, as shown in Table 2.

## VI. CONCLUSIONS

In this paper the generalization of PEEC modeling under the framework of dual discretization is presented.

Table 2. Comparison of complexity of three meshes in Fig. 10. In brackets are the numbers of non-floating edges. Computational times are in seconds. The solution time is intended for one frequency point.

		Mesh 1	Mesh 2	Mesh 3
Geometry	nodes	757	646	646
	faces	1222	900	539
	edges	1977 (1689)	1502 (1198)	1141 (837)
Time	inductance	2995	768	189
	coeff. pot.	884	219	68
	solution	343	127	37
	total time	4222	1114	294

Relations of duality are exploited for the exact assignment of electromagnetic quantities to spatial elements. In this way it is possible to extend standard PEEC to unstructured and mixed structured/unstructured complexes of cells. The proposed technique, named dual-PEEC has proved to be computationally efficient because triangular cells are used only where necessary (e.g. for curved or slanting conductors). The additional computational effort due to mutual resistances is negligible and can be easily handled by adding a proper term in the current controlled voltage source. Finally, dual-PEEC formulation can be improved by using the same “tricks” already tested for standard PEEC (fast multipole method, wavelet transform, model order reduction, ...).

## REFERENCES

- [1] G. Antonini, A. E. Ruehli, and J. Esch. “Non-orthogonal PEEC formulation for time and frequency domain modeling,” *IEEE International Symposium on Electromagnetic Compatibility*, pp. 1:452–456, 2002.
- [2] A. E. Ruehli, G. Antonini, J. Esch, J. Ekman, A. Mayo, and A. Orlandi. “Nonorthogonal PEEC formulation for time- and frequency-domain em and circuit modeling,” *IEEE Transactions on Electromagnetic Compatibility*, pp. 45(2):167 – 176, 2003.
- [3] A. Rong and A. C. Cangellaris. “Generalized PEEC models for three-dimensional interconnect structures and integrated passives of arbitrary shapes,” *IEEE Topical Meeting on Electrical Performance of Electronic Packaging*, pp. 225–228, 2001.
- [4] V. Jandhyala, Y. Wang, D. Gope, and R. Shi. “Coupled electromagnetic-circuit simulation of arbitrarily-shaped conducting structures using triangular meshes,” In *Proceedings International Symposium on Quality Electronic Design*, pp. 38 – 42, San Jose, CA, USA, 2002.

- [5] A. E. Ruehli, "Equivalent circuit models for three-dimensional multiconductor systems," *IEEE Transactions on Microwave Theory and Techniques*, pp. 22:216–221, 1974.
- [6] H. Heeb and A. E. Ruehli, "Three-dimensional interconnect analysis using partial element equivalent circuits," *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, pp. 39(11):974–982, Nov. 1992.
- [7] E. Tonti, "Finite formulation of electromagnetic field," *IEEE Transactions on Magnetics*, pp. 38:333–336, 2002.
- [8] E. Tonti, "Finite formulation of the electromagnetic field," *Progress in Electromagnetics Research, PIER 32* (Special Volume on Geometrical Methods for Computational Electromagnetics), 2001.
- [9] G. Miano and F. Villone, "A surface integral formulation of Maxwell equations for topologically complex conducting domains," *IEEE Transactions on Antennas and Propagation*, pp. 53(12):4001–4014, Dec. 2005.
- [10] A. Bossavit, "A rationale for 'edge-elements' in 3-D fields computations," *IEEE Transactions on Magnetics*, pp. 24(1):74–79, Jan. 1988.
- [11] D. R. Wilton, S. M. Rao, A. W. Glisson, D. H. Schaubert, O. M. Al-Bundak, and C. M. Butler, "Potential integrals for uniform and linear source distributions on Polygonal and Polyedral domains," *IEEE Transactions on Antennas and Propagation*, pp. AP-32:276–281, 1984.
- [12] R. D. Graglia, "On the numerical integration of the linear shape functions times the 3-D Green's function or its gradient on a plane triangle," *IEEE Transactions on Antennas and Propagation*, pp. 41:1448–1455, 1993.
- [13] F. Freschi, G. Gruosso, and M. Repetto, "Unstructured PEEC formulation by dual discretization," *IEEE Microwave and Wireless Components Letters*, pp. 16(10):531–533, Oct. 2006.
- [14] B. Essakhi and L. Pichon, "A finite element time domain analysis of a broad frequency band antenna," In *Proceedings of Workshop on Computational Electromagnetics in Time-Domain (CEM-TD 2005)*, pp. 76–79, Sep. 12-14, 2005.
- [15] F. Freschi, G. Gruosso, A. Maffucci, M. Repetto, F. Villone, and W. Zamboni, "Investigation of low-frequency behaviour of two surface integral full-Maxwell algorithms," *COMPEL - The International Journal for Computation and Mathematics in Electrical and Electronic Engineering*, pp. 26(3):842–858, 2007.



**Fabio Freschi** was born in Turin, Italy, in 1976. He received the Laurea degree (summa cum laude) and the PhD degree in electrical engineering from the Politecnico di Torino, Turin, Italy in 2002 and 2006, respectively. He is currently working as Assistant Professor in Fundamentals of Electrical Engineering

at the Politecnico di Torino. His main research interests are related to optimization and inverse problems in electromagnetism, numerical computation of electromagnetic fields, environmental electromagnetic fields. He published more than 40 papers on these topics.



**Maurizio Repetto** was born in Genova, Italy, in 1960. He received the degree in electrical engineering in 1985 and the Ph.D. degree in electrical engineering in 1989, both from the University of Genova, Genova, Italy. He was a Researcher at the University of Genova from 1990 to 1992. In 1992, he joined

the Politecnico di Torino, Torino, Italy, where he is now Full Professor of Fundamentals of Electrical Engineering. His main research interests are related to the numerical computation of electromagnetic fields in the area of power devices; in particular, he is involved in research projects about the analysis of ferromagnetic hysteresis and the automatic optimization of electromagnetic devices, environmental electromagnetic fields, and shielding. He is author of more than 100 publications on these topics.

# Preliminary Investigation of the NCP Parameter-Choice Method for Inverse Scattering Problems Using BIM: 2-D TM Case

Puyan Mojabi and Joe LoVetri

Department of Electrical and Computer Engineering  
University of Manitoba, Winnipeg, MB, Canada, R3T 5V6  
pmojabi@ee.umanitoba.ca and lovetri@ee.umanitoba.ca

**Abstract**—A new method of choosing the regularization parameter, originally developed for a general class of discrete ill-posed problems, is investigated for electromagnetic inverse scattering problems that are formulated using a penalty method. This so-called normalized cumulative periodogram (NCP) parameter-choice method uses more information available in the residual vector, as opposed to just its norm, and attempts to choose the largest regularization parameter that makes the residual resemble white noise. This is done by calculating the NCP of the residual for each choice of the regularization parameter, starting from large values and stopping at the first parameter which puts the NCP inside the Kolmogorov-Smirnov limits. The main advantage of this method, as compared, for example, to the  $L$ -curve and Generalized Cross-Validation (GCV) techniques, is that it is computationally inexpensive and therefore makes it an appropriate technique for large-scale problems arising in inverse imaging. In this paper, we apply this technique to the general-form Tikhonov-regularized functional arising in the 2-D/TM inverse electromagnetic problem, which is formulated via an integral equation and solved using the Born Iterative Method (BIM).

**Keywords:** Electromagnetic inverse scattering, general-form Tikhonov regularization, normalized cumulative periodogram, Born iterative method.

## I. INTRODUCTION

It is well-known that the inverse scattering problem is ill-posed; the solution to the mathematical problem is not unique and does not depend continuously on the measured data. Therefore, we usually attempt to find a solution to the ill-posed operator by adding some constraints and additional information to the system. Three general classes of handling an ill-posed system of equations are the penalty methods, various projection methods, and hybrid combinations of these (see [1, 2]). The Tikhonov method is the best-known penalty method

approach to regularizing an ill-posed system of equations. The main idea behind the standard-form Tikhonov regularization is that a regularized solution with a small norm and sufficiently small residual norm can be considered a good approximation to the desired unknown solution to the ill-posed system. The second approach, *i.e.*, projection methods, try to project the problem onto a subspace with a good basis for the solution. The most famous projection method is the so-called truncated singular value decomposition (TSVD) [3], but usually projection is achieved using iterative methods such as the conjugate gradient method, GMRES, or other Krylov subspace methods [4]. The last class of approaches are the hybrid methods [5], which are based on regularizing the projected problem. This is done because quite often the projection approach, which casts the problem in a smaller subspace, does not regularize the problem sufficiently.

The regularization in each of these methods usually requires the computationally expensive step of choosing an optimum regularization parameter. This is because the resulting solution can be very sensitive to the choice of regularization parameter. In the Tikhonov method, the regularization parameter controls the weight of the penalty term, while in the projection methods, the dimension of the subspace is considered as the regularization parameter, and therefore in the hybrid methods we need two regularization parameters: one for the dimension of the subspace and the other for regularizing the projected problem. Many regularization parameter-choice methods have been proposed in the literature, for example, the discrepancy principle, Generalized Cross-Validation (GCV), and the  $L$ -curve have been widely used. The discrepancy principle [6] uses the idea that the norm of the residual vector should not be smaller than the norm of the noise in the measured data, but is difficult to apply to electromagnetic inverse problems. Generalized cross-validation [7] is a statistical tool for choosing the regularization parameter by minimizing a specialized functional and does not require any knowledge about the noise variance in the data. The other major parameter-choice method is the  $L$ -curve

method which tries to balance the (semi) norm of the solution and the corresponding residual [8] by choosing the regularization parameter that puts one on the corner of the  $L$ -curve. All of the aforementioned parameter-choice methods are based on the norm of the residual vector. They are also computationally expensive for inverse methods, such as the Born Iterative Method (BIM), where the optimal regularization parameter must be chosen from an unknown wide range of possible values at each iteration.

In this paper, we use a new parameter-choice method for solving the inverse scattering problem which is based on the Normalized Cumulative Periodogram (NCP) of the residual vector, as opposed to just using the norm of the residual; more of the available information is used. This approach is called the NCP parameter-choice method and was recently introduced by Hansen *et. al.* [9]. The underlying idea of this method can be explained as follows: we can model the measured data, contained in a vector as the sum of a signal component and a white noise component but, due to the smoothing effect of the scattering operator, the power spectrum of the signal component will be dominated by low frequencies whereas the power spectrum of the white noise component will have the same expectation at all frequencies. Therefore, the spectral behavior of the signal component is different from the spectral behavior of the white noise and this difference can be used to find a good regularization parameter for our ill-posed problems [9]. As presented in this paper, this algorithm can be used when the residual vector of the discrete ill-posed system of equations looks statistically more like white noise as we decrease the regularization parameter. An extension of this inversion method to get around this restriction has been developed (the details are provided in [10]).

## II. FORMULATION OF THE PROBLEM

The nonlinear integral equation that encapsulates the 2-D time-harmonic, scalar inverse scattering problem for transverse magnetic (TM) fields is written as

$$E_z^s(\mathbf{r}; \mathbf{k}) = k_0^2 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} G(\mathbf{r}, \mathbf{r}'; k_0) E_z(\mathbf{r}'; \mathbf{k}) O(\mathbf{r}') dx' dy' \quad (1)$$

where  $\mathbf{r} = x\hat{a}_x + y\hat{a}_y$  represents the observation point in the Cartesian coordinate system,  $\mathbf{k} = k_x\hat{a}_x + k_y\hat{a}_y$  represents the wavevector, the wavenumber  $k_0$  is related to the wavevector by  $k_0 = |\mathbf{k}|$ .  $E_z^s(\mathbf{r}; \mathbf{k})$  is the  $z$ -component of the scattered electric field defined as the difference between the total field and the incident field. In equation (1), for a non-magnetic media,

$O(\mathbf{r}) = \varepsilon_r(\mathbf{r}) - 1$  is the contrast profile which must be recovered. The two-dimensional free-space Green's function, assuming an  $e^{-j\omega t}$  time dependence, is given as,

$$G(\mathbf{r}, \mathbf{r}'; k_0) = \frac{j}{4} H_0^{(1)}(k_0 |\mathbf{r} - \mathbf{r}'|) \quad (2)$$

where  $H_0^{(1)}(x)$  is the zeroth-order Hankel function of the first kind.

For the results given in this paper, we assume that data collection is done by a set of receivers which are located on a circle around the object and that the object is illuminated by TM plane-waves impinging on the object from different incidence angles. The geometrical configuration is the same as that described in [11]. For obtaining a solution for the contrast in equation (1), we use the Born iterative method [11]. This method proceeds by first using the Born approximation [12] to linearize the nonlinear integral equation which is then solved for the unknown contrast using an inverse solver (below we describe our inverse solver that is based on the general-form Tikhonov regularization in conjunction with the NCP parameter-choice method). The total-field inside the imaging domain, corresponding to this contrast, is then computed using a moment-method forward solver (we use Richmond's method [13]). The newly updated total-field is then used in the subsequent iteration for linearizing the integral equation and the inverse solver is again used for obtaining a new approximation to the contrast. This procedure continues until a termination condition is satisfied: when the change in the relative norm of the solution is less than a specified value.

## III. THE GENERAL-FORM TIKHONOV REGULARIZATION INVERSE SOLVER

After discretizing the linearized integral equation, we obtain a system of linear equations  $Ax = b$ , where  $A \in \mathbb{C}^{m \times n}$ ,  $b \in \mathbb{C}^m$  and  $x$  is to be found. The matrix  $A$  is a discrete representation of the linearized kernel, while  $x$  and  $b$  are column-wise stacked representations of the 2-D discrete contrast function,  $O(x, y)$ , and measured scattered field,  $E_z^s(x, y)$ , respectively. The pseudo-inverse of  $A$  is unbounded due to ill-posedness of the inverse problem. Therefore, for solving this matrix equation, we use the general-form Tikhonov regularization method, which effectively produces a regularized pseudo-inverse operator that is bounded, in conjunction with a parameter-choice method based on NCP that keeps the regularized solution as close as possible to the exact solution. The general-form

Tikhonov regularization method can be represented concisely as producing a solution  $x_\lambda$  given as,

$$x_\lambda = A_\lambda^\dagger b = \min_x \left\{ \|Ax - b\|_2^2 + \lambda^2 \|L(x - x_0)\|_2^2 \right\} \quad (3)$$

where  $\lambda$  is the regularization parameter, and  $L \in \mathbb{C}^{k \times n}$  is called the regularization matrix which can be any matrix whose nullspace does not intersect with the nullspace of  $A$  [14]. The vector  $x_0$  is generally taken as a guess of the solution, and in our case we take it to be the most recent value of the contrast (at the previous iteration). We take  $L$  to be either the identity matrix or the Laplacian operator with zero boundary conditions for the unknown contrast profile. In these two cases, the nullspace of  $L$  is trivial and will not intersect with the numerical nullspace of the ill-posed operator, making the solution to equation (3) unique.

#### IV. THE NCP PARAMETER-CHOICE METHOD

Consider the measured data, *i.e.*, the scattered electric field, as a matrix  $B \in \mathbb{C}^{p \times q}$  where  $p$  denotes the number of different angles at which the TM plane wave illuminates the object and  $q$  denotes the angle at which the  $q^{\text{th}}$  receiver is located (on a circle around the object). As mentioned previously, the measured values in the matrix  $B$  consist of signal and noise components and therefore this matrix can be represented as a signal component matrix  $\bar{B}$  and a white noise component matrix  $E$ :  $B = \bar{B} + E$ .

For simplicity of discussion assume that in equation (3),  $L = I$  where  $I$  is the identity matrix and  $x_0 = 0$ , then the residual vector of Tikhonov solution can be written as,

$$r_\lambda = b - Ax_\lambda = U\Lambda U^H \bar{b} + U\Lambda U^H e \quad (4)$$

where  $\Lambda = \text{diag}\{\lambda^2/(\lambda^2 + \sigma_i^2)\}$  and  $U$  is the matrix of left singular vectors,  $u_i$ , of the matrix  $A$ , with each  $u_i$  corresponding to  $\sigma_i$ , a singular value. The vectors  $\bar{b}$  and  $e$  are obtained by stacking the columns of  $\bar{B}$  and  $E$  into a vector of length  $p \times q$ . For the case where  $L \neq I$ , the singular values will be substituted by generalized singular values of the pair  $(A, L)$  and  $U$  will be the orthonormal matrix in the decomposition of  $A$  using the generalized singular value decomposition of  $(A, L)$  [15].

The diagonal components  $\lambda^2/(\lambda^2 + \sigma_i^2)$  look like a high-pass filter when plotted against the index  $i$ , because the singular values decrease rapidly for ill-posed

problems. The regularization parameter  $\lambda$  determines the “cut-off” index  $k$  of this highpass characteristic; the smaller the value of  $\lambda$ , the larger the cut-off index. This means that as we decrease the regularization parameter  $\lambda$ , the first term in the residual,  $U\Lambda U^H \bar{b}$  will have little contribution from these initial vectors since it can be written as,

$$\sum_{i=1}^n u_i \left[ \frac{\lambda^2}{\lambda^2 + \sigma_i^2} u_i^H \bar{b} \right] = \sum_{i=1}^k u_i \left[ \frac{\lambda^2}{\lambda^2 + \sigma_i^2} u_i^H \bar{b} \right] + \sum_{i=k+1}^n u_i \left[ \frac{\lambda^2}{\lambda^2 + \sigma_i^2} u_i^H \bar{b} \right] \approx \sum_{i=k+1}^n u_i \left[ \frac{\lambda^2}{\lambda^2 + \sigma_i^2} u_i^H \bar{b} \right] \quad (5)$$

It has been argued by Hansen *et al.* [9] that  $\bar{b}$  has few significant (*i.e.*, non-zero) components in the Singular Value Decomposition (SVD) basis corresponding to the first few left singular vectors of  $A$  and the remaining components are almost zero. That is why the smaller the  $\lambda$  (*i.e.*, the larger the cut-off index), the less contribution from the first term in the residual vector (in equation (4)). Thus, using a cut-off index that suppresses all of the significant components of  $\bar{b}$  in the residual means that we have used as much information as possible in the solution, and choosing the smallest such index ensures a stable solution (giving an acceptable trade-off between the regularization and perturbation errors). The regularization parameter corresponding to this cut-off index can be considered as the optimum regularization parameter because it singles out the most stable solution whose residual does not have any important component of  $\bar{b}$ . The residual vector for this optimum regularization parameter will be dominated by  $U\Lambda U^H e$ , but  $U\Lambda U^H e$  behaves statistically like white noise for ill-posed problems because,

$$\text{cov}\{U\Lambda U^H e\} = U\Lambda U^H [\text{cov}\{e\}] U\Lambda U^H = U\Lambda U^H [\eta^2 I](U\Lambda U^H) = \eta^2 (U\Lambda^2 U^H) \approx \eta^2 I_k \quad (6)$$

where  $\eta$  is the standard deviation of the additive white noise and  $I_k$  is the identity matrix with the first  $k$  diagonal elements set to zero. For ill-posed problems  $k$  is very small and this covariance will be very similar to the covariance of white noise. Therefore, the optimum regularization parameter can be considered as the largest  $\lambda$  which makes the residual vector behaves statistically like white noise.

The metric that is used to see if the residual “looks” like white noise is the NCP of the residual [16]. So the regularization procedure that we follow is to start with a large  $\lambda$ , which is usually less than one, in which case the NCP of the residual vector will look like that of the data—meaning that we have a lot of the data information



left in the residual. We then decrease  $\lambda$  until the NCP of the residual first becomes like that of white noise (*i.e.*, a curve between the Kolmogorov-Smirnov (KS) limits for white noise, which are bounds around a straight line). Once this happens, we can be sure that all the important information available in  $\bar{b}$  has been used in calculating  $x_\lambda$ . Notice that if we decrease  $\lambda$  further, the residual is still white noise (or slightly highpass filtered white noise) but the solution is more likely to be unstable due to perturbation errors.

One note regarding our use of the NCP parameter-choice method is that NCP is usually defined for real vectors—because it is generally used as a statistical time-series analysis tool [16]—but here we use the same definition for the NCP of a complex vector. For finding the NCP of the vector  $r_\lambda$ , we first find the power spectrum of this vector as,

$$P = [\text{DFT}\{r_\lambda\}] \odot [\text{DFT}\{r_\lambda\}]^* \quad (7)$$

where  $\odot$  denotes the Hadamard product (*i.e.*, element-wise multiplication) and  $*$  denotes complex conjugate. The components of the NCP vector,  $C \in \mathbb{R}^{n-1}$ , can be calculated as

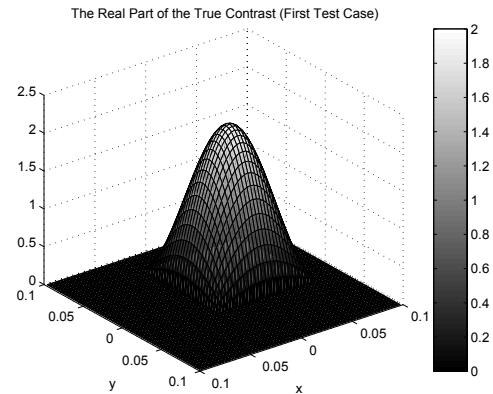
$$C_i = [\|P\|_1 - P_1]^{-1} \sum_{j=2}^{i+1} P_j, \quad i = 1, \dots, n-1. \quad (8)$$

In our case, the Kolmogorov-Smirnov (KS) limit lines as a function of index  $i$  are given as  $i/(n-1) \pm \delta$  where, for a significance level of 5 percent, we set  $\delta = 1.36/\sqrt{n}$ .

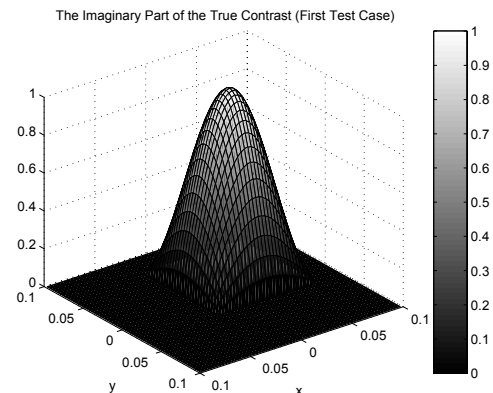
As was mentioned above,  $\bar{b}$ , and consequently  $b$ , have only a few components that are significant in the SVD basis of the ill-posed operator. Due to the similarity of the SVD basis and the Fourier basis [9],  $b$  will also have only a few significant components in the Fourier basis. At the index location where these significant components occur there will also occur step changes in the NCP of  $b$ . This means that the NCP of  $b$  will look like a staircase plot where the step-locations correspond to the location of the significant components of  $b$  in its Fourier basis. In [9], a permutation matrix,  $\Pi$ , has been introduced to reorder the elements of the power spectrum of  $b$  such that all the significant information inside the NCP of  $b$ , *i.e.*, the steps in the original NCP plot, are moved to the first elements of the NCP vector. We've observed that using this permutation matrix has no effect on choosing the optimum regularization parameter.

## V. NUMERICAL RESULTS

We present the results for two cases: a sinusoidal contrast with an amplitude of  $2 + j1$  and also two spatially separated sinusoidal contrasts of amplitudes  $2 + j1$  and  $1.0$ . Figures 1 and 2 show the true contrast function for the first and second test cases respectively. The synthetic data was produced by an MoM solver with triangular meshes — 3448 meshes over the imaging domain — and then white noise was added such that the signal to noise ratio  $\text{SNR} = \|\bar{b}\|_2 / \|e\|_2$  was 8 in the first case and 2 in the second case. In the different iterations of the BIM, the forward solution was obtained by Richmond's method [13] using a pulse basis over the imaging domain (the number of pulses over the imaging

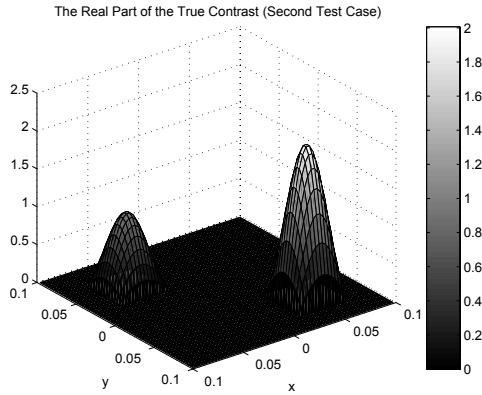


(a)

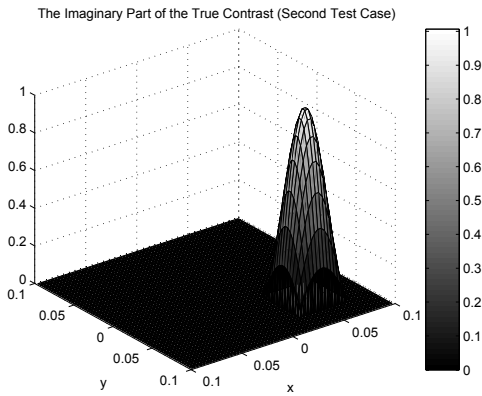


(b)

Fig. 1. True contrast function for the first test case (a) real part (b) imaginary part.



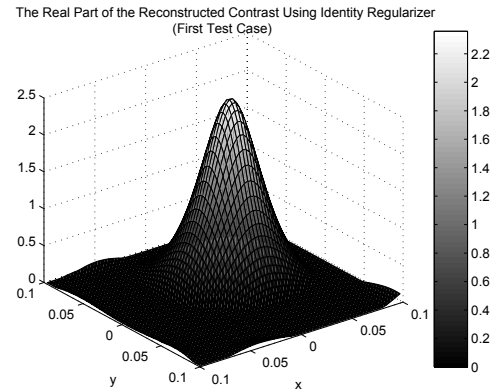
(a)



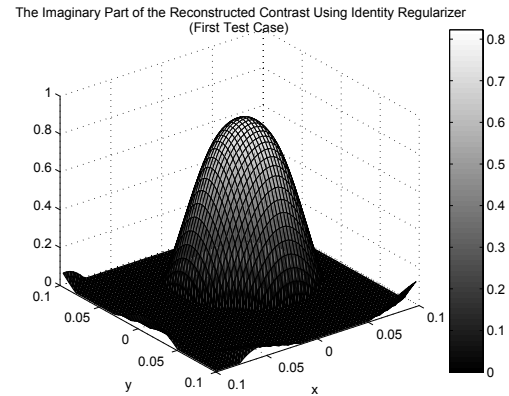
(b)

Fig. 2. True contrast function for the second test case (a) real part (b) imaginary part.

domain is  $60 \times 60$ ). The resulting reconstructions for both test cases using the identity and the Laplacian operator as the regularization matrices are shown in Figures 3 to 6. It seems that choosing  $L$  as the Laplacian operator with zero boundary conditions allows the reconstruction of the peak contrast of both test cases better than setting  $L$  to the identity matrix. On the other hand, using the identity operator as the regularization matrix makes the computations faster because we can use the multishift conjugate gradient least squares (CGLS) method for solving equation (3) [17]. In Fig. 7, the NCP of the synthetic data,  $b$ , is shown. The NCP of some residual vectors corresponding to six different regularization parameters are shown for the Born approximation of the first test case in Fig. 8. As seen in Fig. 8, for large values of  $\lambda$  the NCP of the residual looks like the NCP of  $b$ , showing that we have not used



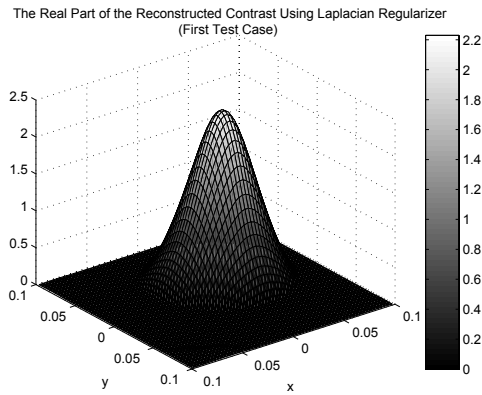
(a)



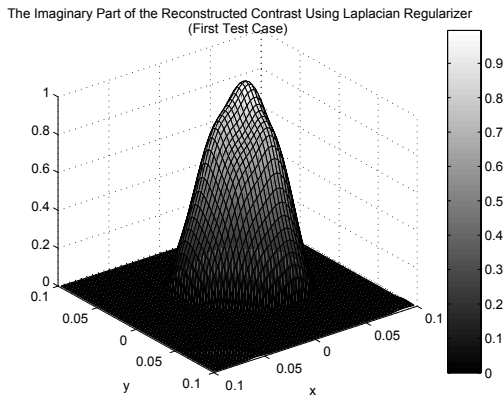
(b)

Fig. 3. Reconstructed contrast of the first test case using the identity regularizer (SNR = 8) (a) real part (b) imaginary part.

all of the available information in reconstructing the profile. As we decrease the  $\lambda$ , we include less information in the residual and more for the solution. The first NCP which fits the Kolmogorov-Smirnov limits is the NCP corresponding to  $\lambda = 0.0039$ . As a comparison, we've solved these two problems using the  $L$ -curve parameter-choice method and it seems that  $L$ -curve chooses a smaller regularization parameter in comparison with the NCP method. For example, in the Born approximation of the first test case, the  $L$ -curve chooses  $\lambda = 0.0028$  as the optimum regularization parameter (as compared to the NCP's  $\lambda = 0.0039$ ). This simply reflects the fact that there is no unique solution to the inverse problem. The  $L$ -curve for the Born approximation of the first case is shown in Fig. 9 using 100 different  $\lambda$ 's .



(a)

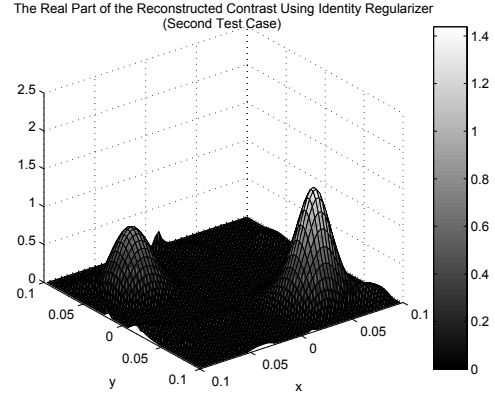


(b)

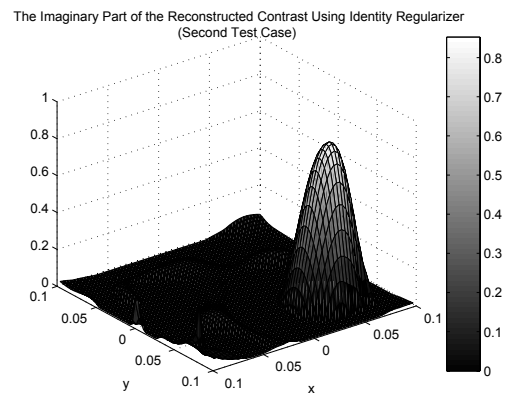
Fig. 4. Reconstructed contrast of the first test case using the Laplacian regularizer (SNR = 8) (a) real part (b) imaginary part.

## VI. CONCLUSION

We've shown that the NCP parameter-choice method can be very useful in large-scale inverse problems because it is based on the Fast Fourier Transform (FFT), and the SVD of the matrix does not need to be computed. The main advantage of the NCP method is that more than just the norm of the residual is used to determine the optimum regularization parameter. Also, the implementation of the NCP algorithm is much easier than the  $L$ -curve and GCV parameter-choice methods because finding the corner of the  $L$ -curve and the minimum of the GCV functional is difficult. We've also found that using the Laplacian operator as the regularization matrix seems to result in a better reconstruction compared to the identity operator;



(a)



(b)

Fig. 5. Reconstructed contrast of the second test case using the identity regularizer (SNR = 2) (a) real part (b) imaginary part.

although, using the identity operator as the regularization matrix makes the computations faster because this allows the use of the multishift CGLS algorithm for solving equation (3).

The NCP method, as presented here, is only applicable when the residual tends to behave like white noise as we decrease the regularization parameter. This requirement can be checked by finding the Tikhonov solution for a very small regularization parameter, say  $\lambda = 10^{-6}$ , and checking the NCP of the corresponding residual. If it satisfies the NCP criteria, then this method should work well. On the other hand, if it does not satisfy the NCP criteria for such a small  $\lambda$ , then the method must be modified. The details of the required modifications are provided in [10].

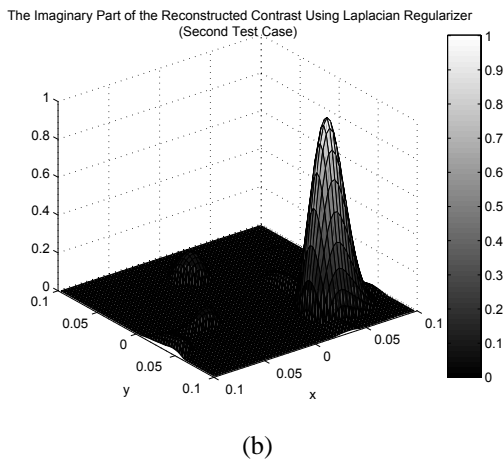
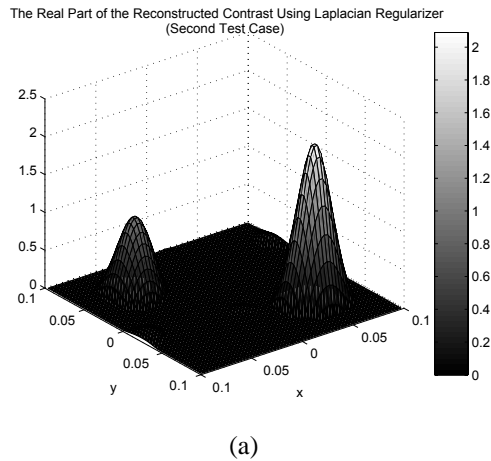


Fig. 6. Reconstructed contrast of the second test case using the Laplacian regularizer (SNR = 2) (a) real part (b) imaginary part.

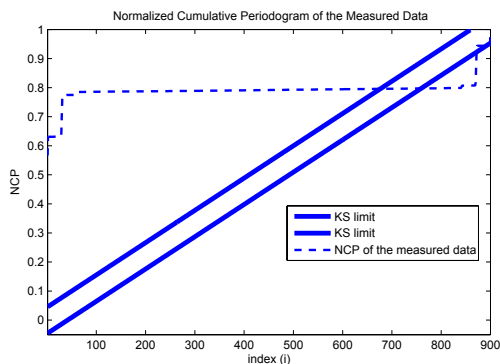


Fig. 7. The NCP of the measured data,  $b$ , for the first test case and Kolmogorov-Smirnov (KS) limits.

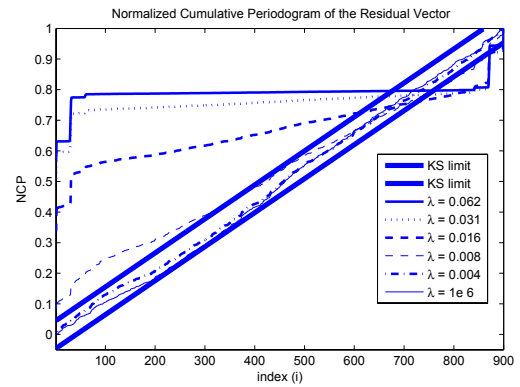


Fig. 8. The NCP of the residual vector for six different regularization parameters at the first iteration of the BIM for the first test case using the identity regularizer and Kolmogorov-Smirnov (KS) limits.

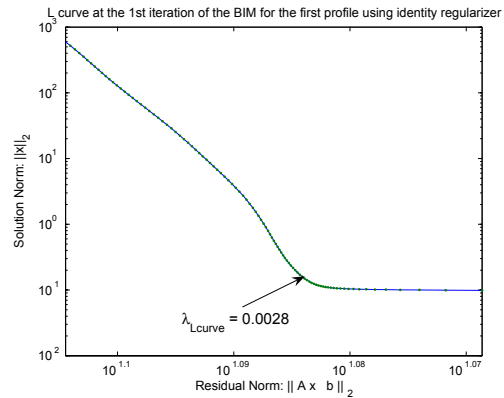


Fig. 9. The  $L$ -curve at the first iteration of the BIM for the first test case using the identity regularizer.

### ACKNOWLEDGEMENT

The authors would like to acknowledge the financial support of the Natural Sciences and Engineering Research Council of Canada. We would also like to thank Prof. P. C. Hansen for making some of his recent work on the NCP method available to us.

## REFERENCES

- [1] A. N. Tikhonov and V. Y. Arsenin, *Solution of Ill-Posed Problems*, Scripta Series in Mathematics, John Wiley & Sons, New York, 1977.
- [2] P. C. Hansen, *Rank-Deficient and Discrete Ill-Posed Problems*, SIAM, Philadelphia, 1998.
- [3] P. C. Hansen, "Truncated Singular Value Decomposition Solutions to Discrete Ill-Posed Problems With Ill-determined Numerical Rank," *SIAM, J. Sci. Stat. Comput.*, vol. 11, no. 3, pp. 503-518, 1990.
- [4] M. E. Kilmer and Dianne P. O'Leary "Choosing Regularization Parameters In Iterative Methods For Ill-Posed Problems," *SIAM J. Matrix. Anal. Appl.*, vol. 22, no. 4, pp. 1204-1221, 2001.
- [5] D. P. O'Leary and J. A. Simmons, "A Bidiagonalization-Regularization Procedure For Large Scale Discretization of Ill-Posed Problems," *SIAM J. Sci. Statist. Comput.*, vol. 2, pp. 474-489, 1981.
- [6] V. A. Morozov, *Methods For Solving Incorrectly Posed Problems*, Springer-Verlag, 1984.
- [7] G. Golub, M. Heath, and G. Wahba, "Generalized Cross-Validation As A Method For Choosing A Good Ridge Parameter," *Technometrics*, vol. 21, pp. 215-223, 1979.
- [8] P. C. Hansen, "Analysis of Discrete Ill-Posed Problems By Means of The *L*-Curve," *SIAM, Review*, vol. 34, pp. 561-580, 1992.
- [9] P. C. Hansen, M. E. Kilmer, and R. H. Kjeldsen, "Exploiting Residual Information In The Parameter Choice For Discrete Ill-Posed Problems," *BIT, Numerical Mathematics*, 2006.
- [10] P. Mojabi and J. LoVetri, "Adapting the Normalized Cumulative Periodogram Parameter-Choice Method to the Tikhonov Regularization of 2-D/TM Inverse Scattering Problems Using Born Iterative Method," *Progress in Electromagnetic Research M*, vol. 1, pp. 111-138, 2008.
- [11] Y. M. Wang and W. C. Chew, "An Iterative Solution of Two-Dimensional Electromagnetic Inverse Scattering Problem," *Int. J. Imaging Syst. Technol.*, vol. 1, no. 1, pp. 100-108, 1989.
- [12] M. Born and E. Wolf, *Principles of Optics*, 7th ed., Cambridge University Press, 1999.
- [13] J. H. Richmond, "Scattering by a Dielectric Cylinder of Arbitrary Cross Section Shape," *IEEE Trans. Antennas. Propag.*, vol. 13, pp. 334-341, 1965.
- [14] H. W. Engl, M. Hanke, and A. Neubauer, *Regularization of Inverse Problems*, Kluwer Academic Publishers, Dordrecht, 2000.
- [15] P. C. Hansen, "Regularization, GSVD and Truncated GSVD," *BIT*, vol. 29, pp. 491-594, 1989.
- [16] W. A. Fuller, *Introduction to Statistical Time Series*, 2nd ed., Wiley, New York, 1976.
- [17] J. V. D. Eshof and Gerard L. G. Sleijpen, "Accurate Conjugate Gradient Methods for Families of Shifted Systems," *SIAM, Applied Numerical Mathematics*, vol. 49, pp. 17-37, 2004.



**Puyan Mojabi** received the B.Sc. degree in Electrical and Computer Engineering from the University of Tehran, Iran, in 2002 and the M.Sc. degree in Electrical Engineering from Iran University of Science and Technology in 2004. Currently, he is working toward the Ph.D. degree in Electrical Engineering at the University of Manitoba, Winnipeg, MB, Canada. His current research interests are computational electromagnetics and inverse problems.



**Joe LoVetri** received the Ph.D. degree in electrical engineering from the University of Ottawa, ON, Canada, in 1991. From 1991 to 1999, he was an Associate Professor in the Department of Electrical and Computer Engineering, University of Western Ontario, Canada. He is currently a Professor in the Department of Electrical and Computer Engineering, and Associate Dean (Research) of the Faculty of Engineering at University of Manitoba, Winnipeg, Canada. His main research interests are in time-domain CEM, modeling of EMC problems, GPR, and microwave imaging techniques.

# Finite Element Modeling of Dual-Core Photonic Crystal Fiber

K. R. Khan and T. X. Wu

School of Electrical Engineering and Computer Science  
University of Central Florida, Orlando, Florida, FL32816, USA  
tomwu@mail.ucf.edu

**Abstract** – In this paper, coupling characteristics of dual-core photonic crystal fiber (PCF) are studied extensively using vector finite element method, which has the potential to realize wavelength selective MUX-DEMUX for wavelength division multiplexing (WDM) application. Dispersion characteristic is also reported and demonstrates the wavelength region where it can support short duration soliton like pulses.

**Keyword:** Photonic Crystal, Waveguide, and FEM.

## I. INTRODUCTION

Photonic crystal fiber (PCF) has recently attracted a considerable amount of attention, because of their unique properties that are not present in conventional optical fibers. A PCF has a central region of pure silica (core) surrounded by air holes. It is a regular morphological microstructure incorporated into the material to radically alter its optical properties [1]. Here a wavelength dependant effective volume average index difference between the defect regions will form the core, and the surrounding region, which contains air holes will be acting as the cladding. This effective-index guidance does not depend on having a periodic array of holes. Even other arrangements could serve a similar function [1, 2]. Index-guiding PCF guides light by total internal reflection between a solid core and a cladding region with multiple air-holes [1]. On the other hand, a perfectly periodic structure exhibiting a photonic band gap (PBG) effect at the operating wavelength to guide light in a low index core-region [1]. In this paper, we will focus on index-guiding PCFs, also called holey fibers (HFs).

HF possess numerous unusual properties such as wide single-mode wavelength, bend-loss edge at short wavelength, controlled effective-core-area at single-mode region, and anomalous group-velocity dispersion at visible and near-infrared wavelengths [2, 3]. It has been shown that the PCF with two adjacent defect area (served as two core), can be used as an optical fiber coupler [4-6]. These PCF couplers have the possibility of realizing a multiplexer-demultiplexer (MUX-DEMUX). In this paper, wavelength dependent coupling characteristics of dual-core PCF couplers are evaluated by using a vector finite element method (FEM) [7, 8]. This gives

understanding of the PCF based MUX-DEMUX for wavelength selective application such as WDM.

## II. FINITE ELEMENT FORMULATION FOR GUIDED MODE

The vector finite element method is used to compute the mode spectrum of an electromagnetic waveguide with arbitrary cross section [9, 10]. It eliminates the disadvantages of the scalar finite element approach of having undesired spurious modes or non-physical solutions and is characterized by easy implementation of boundary conditions at material interfaces [9]. Recent study shows that some double curl finite element formulations are not immune to spurious modes even though they are not observed frequently. It is due to the fact that the initial conditions (forcing term) corresponding to the physical situation eliminates frequent observation of spurious modes [9,10]. We discretized the continuous spectrum by enclosing the structure with an electrical wall as shown in Fig. 1. We have studied dual core PCF and evaluate coupling characteristics using vector FEM. Figure 1 shows the dual core PCF geometry. Some dimension of PCF geometry such as air hole diameter  $d$ , pitch between two adjacent holes  $\Lambda$  and core separation  $C$  were adjusted to obtain the desired values for coupling length and dispersion.

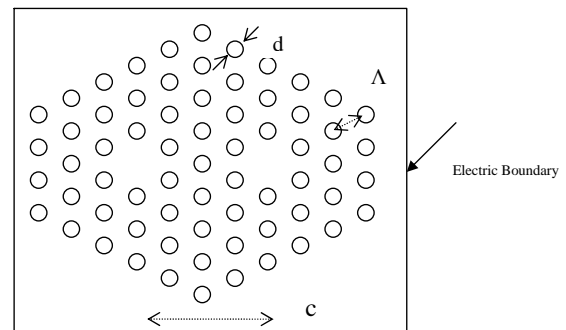


Fig. 1. Arbitrary shaped waveguide with electrical wall.

The vector finite element formulation can be illustrated by using either the  $\mathbf{E}$  or  $\mathbf{H}$  field; here we explain the case for the  $\mathbf{E}$  field, which is the same for the

**H** field. The vector wave equation for the **E** field is given by,

$$\nabla \times \left( \frac{1}{\mu_r} \nabla \times \mathbf{E} \right) - k_0^2 \varepsilon_r \mathbf{E} = 0 \quad (1)$$

where  $\mu_r$  and  $\varepsilon_r$  are, respectively, the permeability and permittivity of the material in the waveguide.  $k_0$  is the free space wave number. The transverse and longitudinal components are separated and are written as,

$$\nabla_t \times \left( \frac{1}{\mu_r} \nabla_t \times \mathbf{E}_t \right) + \frac{1}{\mu_r} (k_z^2 \nabla_t E_z + k_z^2 \mathbf{E}_t) = k_0^2 \varepsilon_r \mathbf{E}_t \quad (2)$$

$$-\frac{1}{\mu_r} [\nabla_t \cdot (\nabla_t E_z + \mathbf{E}_t)] = k_0^2 \varepsilon_r E_z. \quad (3)$$

Since the vector Helmholtz equation is divided into two parts, equations (2) and (3), vector-based tangential edge elements, shown in Fig. 2 (a), can be used to approximate the transverse fields, and nodal-based elements, shown in Fig. 2 (b), can be used to approximate the longitudinal component.

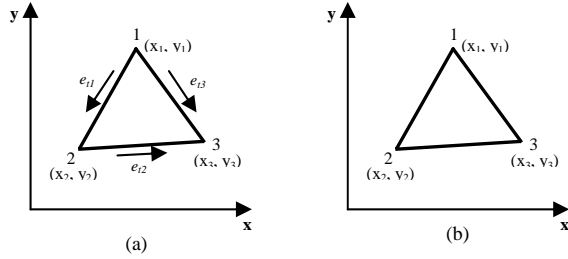


Fig. 2. Configurations of (a) tangential edge elements and (b) node elements.

For a single triangular element shown in Fig. 2, the transverse electric field can be expressed as a superposition of edge elements. The edge elements permit a constant tangential component of the basis function along one triangular edge while simultaneously allowing a zero tangential component along the other two edges [9]. Three such functions overlapping each triangular element provide the complete expansion, that is,

$$\mathbf{E}_t = \sum_{m=1}^3 e_{tm} \mathbf{W}_{tm} \quad (4)$$

where  $m$  indicates the  $m$ -th edge of the triangle and  $\mathbf{W}_{tm}$  is the edge element for edge  $m$  given by,

$$\mathbf{W}_{tm} = L_{tm} (\alpha_i \nabla_t \alpha_j - \alpha_j \nabla_t \alpha_i) \quad (5)$$

$L_{tm}$  is the length of edge  $m$  connecting nodes  $i$  and  $j$  and  $\alpha_i$  is the first-order shape function associated with nodes 1, 2, and 3. The longitudinal component is written as,

$$E_z = \sum_{i=1}^3 e_{zi} \alpha_i. \quad (6)$$

After simple manipulations the integral equation for each element can be written in matrix form as,

$$\begin{bmatrix} S_{e(t)} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} e_t \\ e_z \end{bmatrix} = -k_z^2 \begin{bmatrix} T_{e(tt)} & T_{e(tz)} \\ T_{e(zt)} & T_{e(zz)} \end{bmatrix} \begin{bmatrix} e_t \\ e_z \end{bmatrix} \quad (7)$$

where

$$S_{e(t)} = \frac{1}{\mu_r} \iint_{\Delta} (\nabla_t \times \mathbf{W}_{tm}) \cdot (\nabla_t \times \mathbf{W}_{tn}) ds - k_0^2 \varepsilon_r \iint_{\Delta} (\mathbf{W}_{tm} \cdot \mathbf{W}_{tn}) ds, \quad (8)$$

$$T_{e(tt)} = \frac{1}{\mu_r} \iint_{\Delta} (\mathbf{W}_{tm} \cdot \mathbf{W}_{tn}) ds, \quad (9)$$

$$T_{e(tz)} = \frac{1}{\mu_r} \iint_{\Delta} (\mathbf{W}_{tm} \cdot \nabla \alpha_j) ds, \quad (10)$$

$$T_{e(zt)} = \frac{1}{\mu_r} \iint_{\Delta} (\nabla \alpha_i \cdot \mathbf{W}_{tm}) ds, \quad (11)$$

$$T_{e(zz)} = \frac{1}{\mu_r} \iint_{\Delta} (\nabla \alpha_i \cdot \nabla \alpha_j) ds - k_0^2 \varepsilon_r \iint_{\Delta} \alpha_i \alpha_j ds. \quad (12)$$

These element matrices are assembled over all the triangular elements in the cross section of the structure to obtain a global eigenvalue equation [9, 10]. Solving the above equation yields the eigenvalues or the longitudinal propagation constants  $k_z$ , from which the effective refractive index  $n_e$  is obtained by using the relation,

$$n_e = \left( \frac{k_z}{k_0} \right). \quad (13)$$

### III. CHARACTERISTICS OF DUAL-CORE PCF

In brief signal power is exchanged between the coupled cores due to weak overlap of adjacent electric field. Here light confined into one of the PCF core moves to the other waveguide after propagating a distance known as coupling length  $L_c$  due to the different propagation constants of the even and odd modes of the coupler [11]. Coupling length  $L_c$  is determined by the following equation,

$$L_c = \frac{\pi}{\beta_e - \beta_o} \tag{14}$$

Here  $\beta_e$  is the dispersion coefficient for even mode;  $\beta_o$  is the dispersion coefficient for odd mode. Figure 3 shows the coupling length  $L_c$  with the hole pitch for  $d/\Lambda = 0.7$ . It is shown from numerical results that it is possible to realize significantly shorter MUX-DEMUX PCFs, compared to conventional optical fiber couplers. In the conventional fiber coupler with core spacing and core radius ratio of 3 we found that the coupling length of is 1cm. If the spacing between the cores increases then the coupling length will also increase [12]. On the other hand PCF coupler with  $d/\Lambda = 0.7$  has coupling length of few mm (at 1.50  $\mu\text{m}$  it is between 2-4 mm). The advantage of having short coupling length is that the device becomes more miniaturized.

Figure 4 shows the wavelength dependency of  $L_c$ . It is observed that the coupling length decreases sharply at shorter wavelength up to 0.5  $\mu\text{m}$  then the slope decreases up to 2  $\mu\text{m}$  then it get almost constant at higher wavelength which comparable to the pitch (in this case 2.5  $\mu\text{m}$ ). This is because of the sharp change of material dispersion of silica glass at the short wavelength contributes in a higher extent in evaluation of even and odd wave number. The difference between these two wave numbers varies rapidly at short wavelength but at higher wavelength this remains fairly constant.

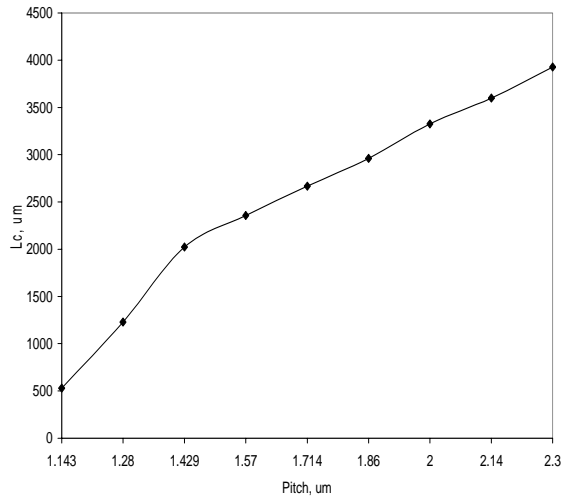


Fig. 3. Coupling lengths of PCF couplers with  $d/\Lambda = 0.7$  at 1500 nm.

It is also possible to significantly change the coupling length by altering the cladding geometry as well as the core separation. For the similar dual core PCF with  $d=0.8 \mu\text{m}$  and  $d/\Lambda=0.4$ , as shown in Fig. 5,  $L_c$  shows similar characteristics with the wavelength but this time it is longer than the one of the PCF with  $d/\Lambda=0.7$  shown in Fig. 4. For a constant  $d/\Lambda$  ratio we can vary  $L_c$  by

varying  $d$ . Figure 6 shows that the  $L_c$  can be increased by increasing  $d$ .

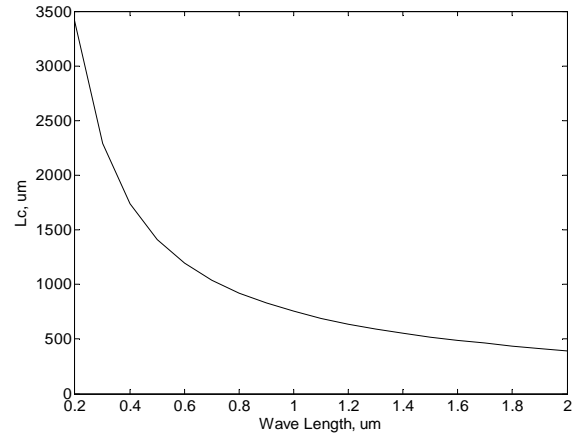


Fig. 4. Coupling lengths of PCF couplers ( $d=0.8 \mu\text{m}$  and  $d/\Lambda = 0.7$ ).

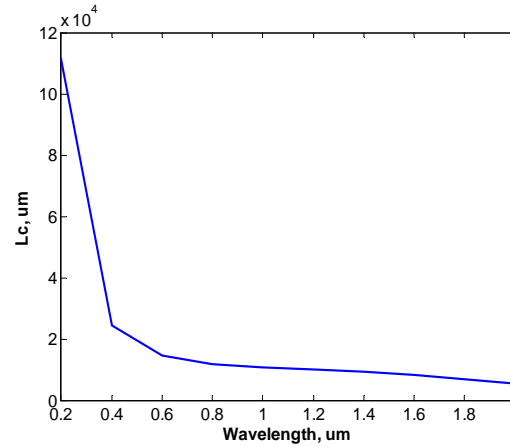


Fig. 5. Coupling length of dual core PCF coupler ( $d=0.8\mu\text{m}$  and  $d/\Lambda=0.4$ ).

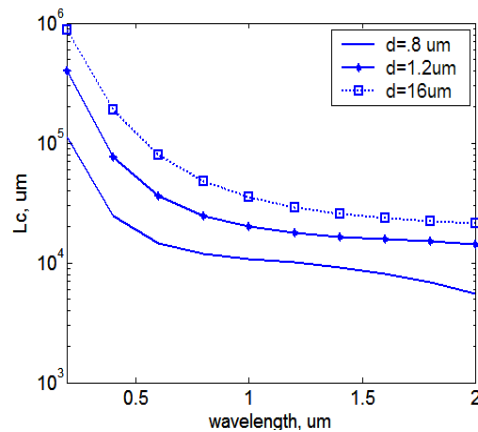


Fig. 6. Coupling length comparison for dual core PCF coupler ( $C=4 \Lambda$ ,  $d/\Lambda=0.4$ ).



#### IV. WAVEGUIDE DISPERSION

When electromagnetic wave interacts with the bound electron of the dielectric materials the medium response is frequency dependant and this manifests itself through the change of refractive index  $n(\omega)$ . It is due to the characteristic resonance frequency at which the bound electron oscillation of the dielectric medium absorbs the electromagnetic radiation [2, 13]. It is the dispersion of optical waveguide which is the most critical for short pulse propagation because the different spectral component associated with the pulse travels at different speed  $c/n(\omega)$ . The parameter  $D$  which is commonly used in optical fiber literature is called the group velocity dispersion and can be expressed as,

$$D = \left( \frac{d\beta}{d\lambda} \right) = - \frac{2\pi c}{\lambda^2} \beta_2 \quad (15)$$

were  $\beta_1$  and  $\beta_2$  are the first and second derivative of wave number  $\beta$  with respect to  $\omega$ . Fig 6 shows the dispersion profile for the same coupler. There is a sign change around  $1.0 \mu\text{m}$ . These wavelength, when the group velocity dispersion shifts from normal ( $D$  is +ve) to anomalous ( $D$  is -ve), is called the zero dispersion wavelength and treated as a very important design parameter for device supporting short pulse propagation. Soliton pulse propagation through this optical waveguide is dependant on the delicate balance between nonlinearity and the dispersion. Nonlinear phase modulation tries to compress the pulses and the dispersion causes pulse broadening. If the power and dispersion is properly balanced then input pulse can propagate without any distortion. In order to support soliton pulses the dispersion has to be in the anomalous region [13, 14].

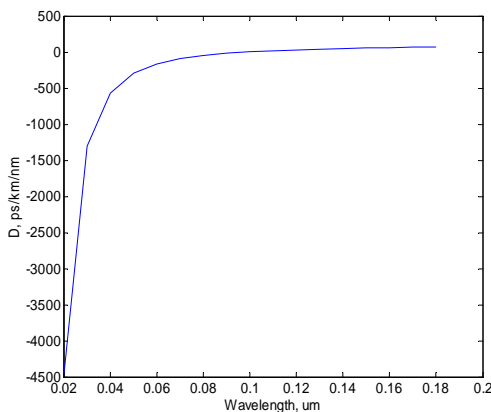


Fig. 7. Dispersion characteristics of dual core PCF ( $d=0.8$  and  $d/\Lambda = 0.7$ ).

From the dispersion results shown in Fig. 7 we found that the wavelengths beyond  $1.0 \mu\text{m}$  have anomalous dispersion. At those wavelengths the waveguide is able to

support soliton pulses. Therefore  $L_c$  along with the dispersion parameters determines the waveform while it propagates.

#### V. CONCLUSION

We have numerically demonstrated the coupling and dispersion characteristics of dual core PCF which have a potential application in the wavelength selective system. With short coupling length compared to regular fiber coupler the device can be used as a MUX-DEMUX or power coupler in the WDM system. Dispersion profile demonstrates the wavelength region where it can support short duration soliton like pulses.

#### REFERENCES

- [1] J. C. Knight, "Photonic crystal fibres," *Nature*, vol. 424, pp. 847-51, 2003.
- [2] K. Saitoh, Y. Sato, and M. Koshiba, "Coupling characteristics of dual-core photonic crystal fiber couplers," *Opt. Express*, vol. 11, pp. 3188-3195, 2003.
- [3] M. J. Gander, R. McBride, J. D. C. Jones, D. Mogilevtsev, T. A. Birks, J. C. Knight, and P. St. J. Russell, "Experimental measurement of group velocity dispersion in photonic crystal fibre," *Electron. Lett.* vol. 35, pp. 63-64, 1999.
- [4] B. J. Mangan, J. C. Knight, T. A. Birks, P. St. J. Russell, and A. H. Greenaway, "Experimental study of dualcore photonic crystal fibre," *Electron. Lett.*, vol. 36, pp. 1358-1359, 2000.
- [5] F. Fogli, L. Saccomandi, P. Bassi, G. Bellanca, and S. Trillo, "Full vectorial BPM modeling of index guiding photonic crystal fibers and couplers," *Opt. Express*, vol. 10, pp. 54-59, 2002.
- [6] B. H. Lee, J. B. Eom, J. Kim, D. S. Moon, U.-C. Paek, and G.-H. Yang, "Photonic crystal fiber coupler," *Opt. Lett.*, vol. 27, pp. 812-814, 2002.
- [7] K. Saitoh and M. Koshiba, "Full-vectorial imaginary-distance beam propagation method based on a finite element scheme: application to photonic crystal fibers," *IEEE J. Quantum Electron.*, vol. 38, pp. 927-933, 2002.
- [8] K. Saitoh, M. Koshiba, T. Hasegawa, and E. Sasaoka, "Chromatic dispersion control in photonic crystal fibers: application to ultra-flattened dispersion," *Opt. Express*, vol. 11, pp. 843-852, 2003.
- [9] J. Jin, *The Finite Element Method in Electromagnetics*, New York: J. Wiley & Sons, 2002.
- [10] C. J. Reddy, M. D. Desphande, C. R. Cockrell, and F. B. Beck, "Finite element method for eigenvalue problems in electromagnetics," *NASA Technical Paper* 3485, December 1994.

- [11] K. Saitoh, Y. Sato, and M. Koshiba, "Coupling characteristics of dual-core photonic crystal fiber couplers," *Opt. Express*, vol. 11, pp. 3188-3195, 2003.
- [12] R. Tewari and K. Thyagarajan, "Analysis of tunable single-mode fiber directional couplers using simple and accurate relations," *J. of Lightwave Technology*, vol. 4, pp.386, 1986.
- [13] G. P Agarwal, *Nonlinear Fiber Optics*, 3<sup>rd</sup> edition Academic press, 2001.
- [14] S. Trillo, S. Wibnitz, E. Wright, and G. I. Stegeman, "Soliton switching in fiber nonlinear directional couplers," *Opt. Lett.*, vol. 13, no. 8, pp 672-674, 1988.



**Kaisar R. Khan** completed B.Sc. in electrical engineering from Bangladesh Institute of Technology, Rajshahi. After that he worked as an engineer for the nationalized telecommunication service provider of Bangladesh (BTTB) until he join the University of Texas at El Paso to complete his MSEE. Currently, He is working for the Ph.D. in electrical engineering at the University of Central Florida. His present research interest includes computational electromagnetics, optical waveguides and nonlinear photonic crystal fiber.



**Thomas Wu** received the M.S. and Ph.D. degrees in Electrical Engineering from the University of Pennsylvania in 1997 and 1999. In the fall of 1999, he joined the School of Electrical Engineering and Computer Science, University of Central Florida (UCF) as an assistant professor. Now he is serving as an associate professor. Professor Wu's current research interests include RF and photonics integrated circuits and packaging, liquid crystal device, computational physics, and nano electronics. He was chairman of IEEE Orlando Section in 2004, and chairman of IEEE MTT and AP Joint Chapter from 2003 to 2004.

# A Parallel Numerical Method to Solve High Frequency Ghost Obstacle Acoustic Scattering Problems

<sup>1</sup>L. Fatone, <sup>2</sup>M. C. Recchioni, and <sup>3</sup>F. Zirilli

<sup>1</sup> Dipartimento di Matematica Pura e Applicata, University of Modena e Reggio Emilia, Modena, Italy

<sup>2</sup> Dipartimento di Scienze Sociali “D. Serrani”, University Politecnica delle Marche, Ancona, Italy

<sup>3</sup> Dipartimento di Matematica “G. Castelnuovo”, University of Roma “La Sapienza”, Roma, Italy

**Abstract** – A highly parallelizable numerical method for time dependent high frequency acoustic scattering problems involving realistic smart obstacles is proposed. A scatterer becomes smart when hit by an incoming wave reacts circulating on its boundary a pressure current to pursue a given goal. A pressure current is a quantity whose physical dimension is pressure divided by time. In particular in this paper we consider obstacles that when hit by an incoming acoustic wave try to generate a virtual image of themselves in a location in space different from their actual location. The virtual image of the obstacle (i.e.: the ghost obstacle) is seen outside a given set containing the obstacle and its virtual image in the apparent location. We call this problem ghost obstacle scattering problem. We model this acoustic scattering problem and several other acoustic scattering problems concerning other types of smart obstacles as optimal control problems for the wave equation. Using the Pontryagin maximum principle the first order optimality conditions associated to these control problems are formulated. The numerical method proposed to solve these optimality conditions is a variation of the operator expansion method and reduces the solution of the optimal control problem to the solution of a sequence of systems of integral equations. These systems of integral equations are solved using suitable wavelet bases to represent the unknowns, the data and the integral kernels. These wavelet bases are made of piecewise polynomial functions and have the property that the matrices that represent the integral operators on these wavelet bases can be approximated satisfactorily with very sparse matrices. This property of the wavelet bases makes possible to approximate the optimal control problems considered with linear systems of equations with hundreds of thousands or millions of unknowns and equations that can be stored and solved with affordable computing resources, that is it makes possible to solve satisfactorily problems with realistic obstacles hit by waves of small wavelength. We validate the method proposed solving some test problems, these problems are optimal control problems involving a “smart” simplified version of the NASA space shuttle hit by incoming waves with small wavelengths compared to its characteristic dimension. We consider test problems with ratio between the characteristic dimension of the obstacle and the wavelength of the time harmonic component of the incoming wave up to approximately sixty. The numer-

ical results obtained are very satisfactory. The website: <http://www.econ.univpm.it/recchioni/scattering/w16> contains stereographic and virtual reality applications showing some numerical experiments relative to the problems studied in this paper. A more general reference to the work in acoustic and electromagnetic scattering of the authors and of their coauthors is the website: <http://www.econ.univpm.it/recchioni/scattering>.

**Keywords:** Acoustic obstacle scattering, smart obstacles, open loop control, operator expansion method, and wavelet expansion.

## I. INTRODUCTION

In this paper we propose a highly performing parallel numerical method to solve an acoustic time dependent scattering problem involving a realistic smart obstacle. The smart obstacle considered is an obstacle that when hit by an incoming acoustic wave and reacts circulating on its boundary a pressure current (i.e., a field that is dimensionally pressure divided by time) in order to generate a virtual image of a possibly different obstacle (i.e., a ghost obstacle) in a location in space different from its actual one. That is, this kind of smart obstacle when hit by an incoming wave generates a scattered field that, outside a suitable set containing the smart obstacle and the ghost obstacle in the apparent location, resembles to the field scattered in the same circumstances by the ghost obstacle located in the apparent location that is in a position in space different from the position of the obstacle.

Let  $\mathbf{R}$  be the set of real numbers and  $\mathbf{R}^3$  be the three dimensional real Euclidean space, the acoustic time dependent direct scattering problem involving the smart obstacle that we want to solve can be stated as follows:

Given an incoming acoustic field propagating in  $\mathbf{R}^3$ , an obstacle  $\Omega \subset \mathbf{R}^3$  non empty characterized by an acoustic boundary impedance  $\chi$ , a ghost obstacle  $\Omega_G \subset \mathbf{R}^3$  non empty characterized by an acoustic boundary impedance  $\chi_G$  such that  $\Omega \cap \Omega_G = \emptyset$  and a set  $\Omega_\epsilon \subset \mathbf{R}^3$ , such that  $\Omega, \Omega_G \subset \Omega_\epsilon$ , find a pressure current circulating on the boundary of  $\Omega$  such that the field scattered by  $\Omega$  when hit by the incoming acoustic field appears, outside  $\Omega_\epsilon$ , “as similar as possible” to the field scattered in the same circumstances by  $\Omega_G$ .

For simplicity in the numerical examples presented in Section IV we limit our attention to the case when  $\Omega_G$  is the set  $\Omega$  translated in space and  $\chi = \chi_G$  is a constant. We denote with  $\partial\Omega$ ,  $\partial\Omega_G$  the boundary of  $\Omega$  and  $\Omega_G$  respectively. This is the formulation of the ghost obstacle scattering problem considered in the numerical examples and it is a simplified version of a more general formulation of the problem that is given in Section II. Indeed, in the numerical examples we solve the ghost obstacle scattering problem in this simplified formulation since we consider problems that are already difficult for other reasons. In fact we consider problems that involve obstacles with complex geometry, that is realistic obstacles, and incoming waves of small wavelengths.

We formulate the ghost obstacle scattering problem as an optimal control problem for the wave equation and, under suitable assumptions, applying the Pontryagin maximum principle, we derive the first order necessary optimality condition relative to the optimal control problem considered. This condition is formulated as an exterior value problem for two coupled wave equations. Assuming that the incoming and the scattered fields and some auxiliary variables can be represented as superposition of time harmonic waves, we reduce the solution of this exterior problem to the solution of a set of exterior problems for two coupled Helmholtz equations. Finally using a perturbation expansion known as operator expansion method (see for example [1, 2]) we reduce the solution of this set of exterior problems for two coupled Helmholtz equations to the solution of a set of systems of first kind integral equations. This approach has been used to solve several direct acoustic and electromagnetic scattering problems involving several kinds of smart obstacles (i.e., undetectable obstacles [2], masked obstacles [3, 4, 5], ghost obstacles [6, 7]). Moreover some attempts to solve inverse acoustic scattering problems involving smart obstacles with ad hoc methods have been made with promising results (i.e.: for furtive and masked obstacles [8, 9] and for ghost obstacles [10, 11]).

A common feature of the work contained in the papers mentioned previously is that the numerical methods proposed assume implicitly that the smart obstacles considered must have simple geometries. In particular they assume that their shape must be not too far from being the shape of a sphere and that the incoming waves used to illuminate the obstacles must be a superposition of time harmonic waves with wavelengths not too small, let us say approximately not smaller than the characteristic length of the obstacles. These assumptions are due to the following facts: the “special” surfaces used in the development of the operator expansion method are chosen to be surfaces of spheres and the set of coupled integral equations coming from the first order optimality condition is solved projecting the systems of integral equations on a vector space generated by a finite subset of the spherical harmonic functions (see [12], p. 77). The use of spherical surfaces as “special” surfaces of the operator expansion

method and of the spherical harmonic functions as function basis to approximate the integral equations makes the corresponding computational method very efficient. In fact this method reduces the solution of the integral equations to the solution of diagonal systems of linear equations. However, it has some disadvantages, one of them is that in practical computations only the first few hundreds spherical harmonic functions can be used so that only linear systems involving few hundreds equations and unknowns can be considered. That is in the scattering problems that can be solved satisfactorily using spherical surfaces as “special” surfaces and the spherical harmonic functions as function basis, the obstacles must have a shape not too far from being the shape of a sphere and the ratios between the characteristic length of the obstacles and the wavelengths present in the Fourier decomposition of the incoming field considered must be at most of a few units.

Moreover, the first order necessary optimality condition given in [6, 7] for the direct ghost scattering problem is derived assuming that  $\Omega$  and  $\Omega_\epsilon$  are star-like obstacles with respect to the same point and that the set  $\Omega_\epsilon$  is a “magnification” of the obstacle  $\Omega$  itself. This last assumption in [6, 7] is used to make an expansion involving the surface  $\partial\Omega_\epsilon$ , boundary of  $\Omega_\epsilon$ , using as “base point” the surface  $\partial\Omega$  to get a set of integral equations on  $\partial\Omega$ . These assumptions restrict the choice of the shape of the smart obstacles and the choice of the distance between the smart and the ghost obstacle that can be considered. In fact, only when the distance between the smart and the ghost obstacle is sufficiently small we have that the expansion involving the surface  $\partial\Omega_\epsilon$  mentioned above is convergent.

In this paper we overcome these restrictions giving a new formulation of the first order necessary optimality condition. In particular we remove the assumptions that the set  $\Omega$  and the set  $\Omega_\epsilon$  are starlike with respect to the same point and that  $\Omega_\epsilon$  is a magnification of  $\Omega$ . Moreover we remove the assumptions that the “special” surfaces of the operator expansion method are spherical surfaces and that  $\Omega$  is starlike. As a consequence are removed the restrictions that the smart obstacles considered must have a shape not too far from being the shape of a sphere and that the wavelengths of the time harmonic waves contained in the Fourier decomposition of the incoming field are at least of the same order of magnitude of the characteristic length of the obstacles. A first attempt of removing the restrictions on the shape of the scatterer and on the wavelengths of the time harmonic waves contained in the incoming field in the case of smart obstacles that pursue the goal of being undetectable has been made in [13]. In particular in [13] the set of the coupled integral equations coming from the first order necessary optimality condition of the control problem associated to the furtivity problem is derived and is solved discretizing the integral equations using a wavelet basis introduced in [10]. The use of this wavelet basis makes possible

to solve with affordable computing resources scattering problems involving realistic obstacles when the ratio between the characteristic dimension of the obstacle and the wavelength of the incoming wave goes up to (approximately) sixty at the price of solving non diagonal sparse linear systems having a condition number that increases when the number of unknowns increases. These sparse linear systems contain hundreds of thousands or millions of unknowns and equations. The ghost scattering problem considered in this paper is more difficult than the furtivity problem studied in [13]. In fact it requires a reformulation of the optimal control problem and the use of wavelet bases with improved sparsification properties able to approximate the integral operators involved in the integral equations coming from the first order necessary optimality condition with very sparse matrices. Hence, in this paper, a new wavelet basis, introduced in [14], is used to approximate the set of coupled integral equations coming from the new formulation of the first order optimality condition for the acoustic ghost obstacle scattering problem. This new wavelet basis is made of piecewise constant functions, and generalizes the Haar's basis. Furthermore, the computation of the matrix elements of the matrices that approximate the integral operators consists in the computation of four dimensional integrals independent one from the other. This type of computation is very well suited for parallel computing or even for distributed computing. The computation of these matrix elements is by far the dominant part of the computation in the solution of the ghost obstacle scattering problem when the integral equations that translate the optimal control problem must be approximated in high dimensional vector spaces. Concluding we can say that we have developed an efficient computational method that reduces the solution of the direct ghost obstacle scattering problem involving realistic obstacles and small wavelengths (see Section IV where a simplified version of the NASA space shuttle is considered as obstacle) to the solution of a set of sparse linear systems of equations with hundreds of thousands or millions of unknowns and equations that can be stored and solved using affordable computing resources. The computational method takes care of the ill-conditioning of the sparse linear systems obtained solving them with an iterative procedure that uses suitably chosen starting points. Thank to the use of a parallel implementation of the computational method based on FORTRAN 77 as programming language, on MPI as message passing library, to the use of the SP5 IBM machine with 168 processors (144 processors dedicated to batch running) of CASPUR (Roma, Italy) computing center and to the use of stereographic and virtual reality animations to represent the numerical results obtained, we can solve efficiently very difficult scattering problems and we can represent satisfactorily the results obtained. Some stereographic and virtual reality applications relative to numerical experiments on test problems can be found in the website

<http://www.econ.univpm.it/recchioni/scattering/w16>.

A more general reference to the work in acoustic and electromagnetic scattering of the authors and of their coauthors is the website: <http://www.econ.univpm.it/recchioni/scattering>.

In Section II we formulate the time dependent ghost obstacle scattering problem as an optimal control problem for the wave equation and we derive the corresponding first order optimality condition using the Pontryagin maximum principle (see [7]). In Section III, under suitable assumptions, we reduce the first order optimality condition derived in Section II first to a set of time harmonic problems for a system of two coupled Helmholtz equations and later to a set of systems of integral equations. Moreover we describe the algorithm developed to solve the systems of integral equations using a wavelet basis introduced in [14]. Finally, in Section IV starting from synthetic data the algorithm discussed in Section III is used to solve numerically in some test cases the ghost obstacle scattering problem.

## II. THE TIME DEPENDENT GHOST OBSTACLE DIRECT SCATTERING PROBLEM

Let  $\mathbf{C}$  be the set of complex numbers, for  $\eta \in \mathbf{C}$  let  $|\eta|$ ,  $\bar{\eta}$  denote the modulus and the complex conjugate of  $\eta$  respectively. Let  $\underline{x} = (x, y, z)^T \in \mathbf{R}^3$  be a generic vector where the superscript T means transposed,  $(\cdot, \cdot)$  be the Euclidean scalar product in  $\mathbf{R}^3$  and  $\|\cdot\|$  be the corresponding Euclidean vector norm. In the following with abuse of notation occasionally we will use the real Euclidean scalar product of complex vectors.

Let us formulate the scattering problem that we want to study. Let  $\Omega \subset \mathbf{R}^3$ ,  $\Omega_G \subset \mathbf{R}^3$  be two bounded simply connected open sets with locally Lipschitz boundaries  $\partial\Omega$ ,  $\partial\Omega_G$ , and let  $\bar{\Omega}$  and  $\bar{\Omega}_G$  be their closures. Furthermore, let  $\Omega$ ,  $\Omega_G$  be such that:  $\Omega \neq \emptyset$ ,  $\Omega_G \neq \emptyset$  and  $\Omega \cap \Omega_G = \emptyset$ . We denote with  $\underline{n}(\underline{x}) = (n_1(\underline{x}), n_2(\underline{x}), n_3(\underline{x}))^T \in \mathbf{R}^3$ ,  $\underline{x} \in \partial\Omega$ , the outward unit normal vector to  $\partial\Omega$  in  $\underline{x} \in \partial\Omega$ . Since  $\Omega$  has a locally Lipschitz boundary  $\underline{n}(\underline{x})$ ,  $\underline{x} \in \partial\Omega$ , exists almost everywhere (see [15] Lemma 2.42 p. 88); similar statements hold for the outward unit normal vector to  $\partial\Omega_G$ .

We assume that  $\Omega$  and  $\Omega_G$  are characterized by constant acoustic boundary impedances  $\chi \geq 0$  and  $\chi_G \geq 0$  respectively. The case  $\chi = +\infty$  and/or  $\chi_G = +\infty$  (i.e.: the case of acoustically hard obstacles) can be treated with simple modifications of the formulae that follow. We refer to  $(\Omega; \chi)$  as to the obstacle, and to  $(\Omega_G; \chi_G)$  as to the ghost obstacle. Without loss of generality, we can assume that the origin of the coordinate system lies in  $\Omega$ .

We consider an acoustic incident wave  $u^i(\underline{x}, t)$ ,  $(\underline{x}, t) \in \mathbf{R}^3 \times \mathbf{R}$ , propagating in a homogeneous isotropic medium in equilibrium, with no source terms present, satisfying the wave equation in  $\mathbf{R}^3 \times \mathbf{R}$  with wave propagation velocity  $c > 0$ . We denote with  $u^s(\underline{x}, t)$ ,  $(\underline{x}, t) \in (\mathbf{R}^3 \setminus \bar{\Omega}) \times \mathbf{R}$ , and with  $u_G^s(\underline{x}, t)$ ,  $(\underline{x}, t) \in$

$(\mathbf{R}^3 \setminus \overline{\Omega_G}) \times \mathbf{R}$ , the waves scattered, respectively, by the obstacle  $(\Omega; \chi)$  and by the ghost obstacle  $(\Omega_G; \chi_G)$  when hit by  $u^i(\mathbf{x}, t)$ ,  $(\mathbf{x}, t) \in \mathbf{R}^3 \times \mathbf{R}$ .

The scattered acoustic field  $u^s(\mathbf{x}, t)$ ,  $(\mathbf{x}, t) \in (\mathbf{R}^3 \setminus \overline{\Omega}) \times \mathbf{R}$ , when the obstacle  $\Omega$  is not smart (i.e. is an obstacle that does not react to the incident wave circulating a pressure current on its boundary to pursue a goal) is defined as the solution of the following exterior problem for the wave equation (see [1]),

$$\Delta u^s(\underline{x}, t) - \frac{1}{c^2} \frac{\partial^2 u^s}{\partial t^2}(\underline{x}, t) = 0, \quad (\underline{x}, t) \in (\mathbf{R}^3 \setminus \overline{\Omega}) \times \mathbf{R} \quad (1)$$

with the boundary condition (see [12] p. 66),

$$-\frac{\partial u^s}{\partial t}(\underline{x}, t) + c\chi \frac{\partial u^s}{\partial \underline{n}(\underline{x})} = g(\underline{x}, t), \quad (\underline{x}, t) \in \partial\Omega \times \mathbf{R} \quad (2)$$

where  $g(\underline{x}, t)$  is given by,

$$g(\underline{x}, t) = \frac{\partial u^i}{\partial t}(\underline{x}, t) - c\chi \frac{\partial u^i}{\partial \underline{n}(\underline{x})}(\underline{x}, t), \quad (\underline{x}, t) \in \partial\Omega \times \mathbf{R} \quad (3)$$

the condition at infinity,

$$u^s(\underline{x}, t) = O\left(\frac{1}{r}\right), \quad r \rightarrow +\infty, \quad t \in \mathbf{R} \quad (4)$$

and the radiation condition,

$$\frac{\partial u^s}{\partial r}(\underline{x}, t) + \frac{1}{c} \frac{\partial u^s}{\partial t}(\underline{x}, t) = o\left(\frac{1}{r}\right), \quad r \rightarrow +\infty, \quad t \in \mathbf{R} \quad (5)$$

where  $r = \|\underline{x}\|$ ,  $\underline{x} \in \mathbf{R}^3$ ,  $\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$  is the Laplace operator and  $O(\cdot)$  and  $o(\cdot)$  are the Landau symbols. We note that  $g(\underline{x}, t)$ ,  $(\underline{x}, t) \in \partial\Omega \times \mathbf{R}$ , is defined almost everywhere, and that the boundary condition of equation (2) can be adapted to deal with the limit case of acoustically hard obstacles, i.e. obstacles such that  $\chi = +\infty$  (see [1, 2]). The obstacle  $(\Omega; \chi)$  that scatters the field  $u^s$  solution of equations (1), (2), (4), and (5) is called passive obstacle. The field  $u_G^s(\mathbf{x}, t)$ ,  $(\mathbf{x}, t) \in (\mathbf{R}^3 \setminus \overline{\Omega_G}) \times \mathbf{R}$ , scattered by the (passive) ghost obstacle is defined as the solution of equations (1), (2), (4), and (5) when we replace  $\Omega$  with  $\Omega_G$  and  $\chi$  with  $\chi_G$ . Note that we always consider the ghost obstacle  $(\Omega_G; \chi_G)$  as a passive obstacle.

Remember that in the ghost obstacle problem the smart obstacle  $(\Omega; \chi)$  when hit by an incoming wave tries to generate a scattered wave that resembles, outside a given set containing  $\Omega$  and  $\Omega_G$ , to the wave scattered by  $(\Omega_G; \chi_G)$  in the same circumstances. Our goal is to model the ghost obstacle problem as an optimal control problem for the wave equation introducing a control variable  $\psi(\mathbf{x}, t)$ ,  $(\mathbf{x}, t) \in \partial\Omega \times \mathbf{R}$ , acting on the boundary of the obstacle. This is done replacing the boundary condition of equation (2) with the following boundary condition,

$$-\frac{\partial u^s}{\partial t}(\underline{x}, t) + c\chi \frac{\partial u^s}{\partial \underline{n}(\underline{x})} = g(\underline{x}, t) + (1 + \chi)\psi(\underline{x}, t), \quad (\underline{x}, t) \in \partial\Omega \times \mathbf{R}. \quad (6)$$

We note that the physical dimension of the control function  $\psi$  is pressure divided by time, so that we call  $\psi$  ‘‘pressure current’’.

Let  $\Omega_\epsilon$  be a bounded simply connected open set containing  $\Omega$  and  $\Omega_G$  with locally Lipschitz boundary  $\partial\Omega_\epsilon$  and let  $ds_{\partial\Omega_\epsilon}$ ,  $ds_{\partial\Omega}$  be the surface measures on  $\partial\Omega_\epsilon$  and  $\partial\Omega$  (see [15] Lemma 1.1 p. 119-120), respectively.

As done in [7, 6] we define the cost functional,

$$F_{\lambda, \mu, \epsilon}(\psi) = \int_{\mathbf{R}} dt \left\{ \int_{\partial\Omega_\epsilon} (1 + \chi)\lambda c [u^s(\underline{x}, t) - u_G^s(\underline{x}, t)]^2 ds_{\partial\Omega_\epsilon}(\underline{x}) + \int_{\partial\Omega} (1 + \chi)\mu \varsigma \psi^2(\mathbf{x}, t) ds_{\partial\Omega}(\underline{x}) \right\} \quad (7)$$

where  $\lambda \geq 0$ ,  $\mu \geq 0$  are dimensionless constants such that  $\lambda + \mu = 1$ , and  $\varsigma$  is a nonzero positive dimensional constant. Note that in the first addendum of equation (7) we have introduced as a factor the propagation velocity  $c$  that does not appear in the functional used in [6, 7], the presence of this factor simplifies some of the formulae that follow. We model the direct ghost obstacle scattering problem via the following optimal control problem,

$$\min_{\psi \in \mathcal{C}} F_{\lambda, \mu, \epsilon}(\psi). \quad (8)$$

Subject to the constraints of equations (1), (4), (5), and (6). The set  $\mathcal{C}$  is the space of the admissible controls and is a vector space that we leave unspecified in this paper (see [7] for further details). The obstacle  $(\Omega; \chi)$  that generates the scattered field  $u^s$  solution of equations (8), (1), (4), (5), and (6) is a smart (or active) obstacle that we call of ghost obstacle type.

As shown in [2, 7] the cases  $\mu = 0$  and  $\mu = 1$  are trivial. The choice of the cost functional of equation (7) is motivated by the fact that when  $0 < \mu < 1$  we have  $\lambda > 0$ , that is, when  $0 < \mu < 1$  the minimization of  $F_{\lambda, \mu, \epsilon}$  makes small on  $\partial\Omega_\epsilon$  for all times the difference between the field scattered by the smart obstacle and the field scattered by the ghost obstacle and makes small on  $\partial\Omega$  for all times the control function used to achieve this goal. Note that forcing the two scattered fields to be similar on  $\partial\Omega_\epsilon$  for all times implies that they remain similar in  $\mathbf{R}^3 \setminus \overline{\Omega_\epsilon}$  for all times. So that an observer located in  $\mathbf{R}^3 \setminus \overline{\Omega_\epsilon}$  that observes the scattered field is induced to believe that the obstacle generating the scattered field is the ghost obstacle  $(\Omega_G; \chi_G)$  instead than the obstacle  $(\Omega; \chi)$ . From now on we restrict our attention to the case  $0 < \mu < 1$ .

We note that the functional of equation (7) is defined via two integrals, one integral on the boundary of  $\Omega$  and

the other integral on the boundary of  $\Omega_\epsilon$ . This fact makes difficult to write conveniently the first order optimality condition associated to the optimal control problem of equations (8), (1), (4), (5), and (6). In reference [7] this difficulty has been solved assuming that the boundary of  $\Omega$  is a star like surface and that the boundary of  $\Omega_\epsilon$  is obtained by a simple transformation of the boundary of  $\Omega$  that previously we have called magnification. In this paper we modify this assumption since we deal with obstacles (see Fig. 3(b)) whose boundaries are not starlike surfaces with respect to a point in their interior.

The numerical experiments presented in this paper are done using cylindrical coordinates since the simplified NASA space shuttle (Fig. 3(b)), that is the obstacle considered in Section IV, can be described conveniently using these coordinates. So that, for simplicity, we use cylindrical coordinates also in this section and in Section III in the exposition of the solution method. When necessary more general curvilinear coordinate systems can be considered as shown, for example, in [16].

Let us introduce the canonical cylindrical coordinate system in  $\mathbf{R}^3$ , that is  $(r_1, \phi, z)$ , with origin in a point in the interior of the smart obstacle that will be specified later and let  $\hat{\phi} = (\cos \phi, \sin \phi, 0)^T \in \mathbf{R}^3$ ,  $\phi \in [0, 2\pi)$  and  $\hat{z} = (0, 0, z)^T = z\mathbf{e}_3 \in \mathbf{R}^3$ ,  $z \in \mathbf{R}$ ,  $\mathbf{e}_3 = (0, 0, 1)^T \in \mathbf{R}^3$ . Note that with abuse of notation, in order to keep the exposition simple in formulae equations (9)-(34) we have used the coordinates  $(r_1, \phi, z)$  to denote several cylindrical coordinate systems obtained one from the others with suitable translations, in the specific contexts considered the coordinate systems used will be understandable. We assume that:

(a) The obstacle  $(\Omega; \chi)$  is given by,

$$\begin{aligned} \Omega &= \{ \underline{x} = (r_1 \cos \phi, r_1 \sin \phi, z)^T \quad (9) \\ &= r_1 \hat{\phi} + z\mathbf{e}_3 \in \mathbf{R}^3 \mid 0 \leq r_1 < \xi(\phi, z), \\ &\phi \in [0, 2\pi), z \in [z_i, z_f] \}, \quad (10) \end{aligned}$$

where  $z_i, z_f$  are two real numbers such that  $z_i < z_f$  and  $\xi(\phi, z) > 0$ ,  $\phi \in [0, 2\pi)$ ,  $z \in (z_i, z_f)$  is a single valued function defined in  $[0, 2\pi) \times [z_i, z_f]$  sufficiently regular to make sense out for the manipulations that follow. As a consequence of equation (9) we have,

$$\begin{aligned} \partial\Omega &= \{ \underline{x} = (r_1 \cos \phi, r_1 \sin \phi, z)^T = \\ &\xi(\phi, z) \hat{\phi} + z\mathbf{e}_3 \in \mathbf{R}^3, \phi \in [0, 2\pi), z \in [z_i, z_f] \}. \quad (11) \end{aligned}$$

(b) The sets  $\Omega_G$  and  $\partial\Omega_G$  are given by,

$$\begin{aligned} \Omega_G &= \{ \underline{x} = \underline{p}_G + (r_1 \cos \phi, r_1 \sin \phi, s_G(z))^T = \\ &\underline{p}_G + r_1 \hat{\phi} + s_G(z)\mathbf{e}_3 \in \mathbf{R}^3 \mid 0 \leq r_1 < \xi_G(\phi, s_G(z)), \\ &\phi \in [0, 2\pi), z \in [z_i, z_f] \}, \quad (12) \end{aligned}$$

and

$$\begin{aligned} \partial\Omega_G &= \{ \underline{x} = \underline{p}_G + (r_1 \cos \phi, r_1 \sin \phi, s_G(z))^T = \\ &\underline{p}_G + \xi_G(\phi, s_G(z)) \hat{\phi} + s_G(z)\mathbf{e}_3 \in \mathbf{R}^3, \\ &\phi \in [0, 2\pi), z \in [z_i, z_f] \}, \quad (13) \end{aligned}$$

where  $s_G(z) = z_i^G + (z - z_i)(z_f^G - z_i^G)/(z_f - z_i)$  is a linear function of  $z$ ,  $z_i^G, z_f^G$  are constants such that  $z_i^G < z_f^G$ ,  $\xi_G(\phi, s_G(z))$ ,  $\phi \in [0, 2\pi)$ ,  $z \in [z_i, z_f]$  is a single valued function and  $\underline{p}_G \in \mathbf{R}^3$  is a suitable vector (see Fig. 1).

(c) The sets  $\Omega_\epsilon$  and  $\partial\Omega_\epsilon$  are given by,

$$\begin{aligned} \Omega_\epsilon &= \{ \underline{x} = \underline{p} + (r_1 \cos \phi, r_1 \sin \phi, s(z))^T = \\ &\underline{p} + r_1 \hat{\phi} + s(z)\mathbf{e}_3 \in \mathbf{R}^3, 0 \leq r_1 < \xi_\epsilon(\phi, s(z)), \\ &\phi \in [0, 2\pi), z \in [z_i, z_f] \}, \quad (14) \end{aligned}$$

and

$$\begin{aligned} \partial\Omega_\epsilon &= \{ \underline{x} = \underline{p} + (r_1 \cos \phi, r_1 \sin \phi, s(z))^T = \\ &\underline{p} + \xi_\epsilon(\phi, s(z)) \hat{\phi} + s(z)\mathbf{e}_3 \in \mathbf{R}^3, \\ &\phi \in [0, 2\pi), z \in [z_i, z_f] \}, \quad (15) \end{aligned}$$

where  $s(z) = z_i^* + (z - z_i)(z_f^* - z_i^*)/(z_f - z_i)$  is a linear function of  $z$  and  $z_i^*, z_f^*$  are constants such that  $z_i^* < z_f^*$ . Finally,  $\underline{p} \in \mathbf{R}^3$  is a suitable vector (see Fig. 1) and  $\xi_\epsilon$  is a single valued function such that  $\Omega \subset \Omega_\epsilon$  and  $\Omega_G \subset \Omega_\epsilon$ .

Figure 1 shows an example of the ghost obstacle experiment in a simple situation where  $\Omega$  is a sphere of center the origin and  $\Omega_G$  is a translation of  $\Omega$ . Note that only for simplicity we have chosen the origin of the cylindrical coordinate system in the center of mass of the smart obstacle.

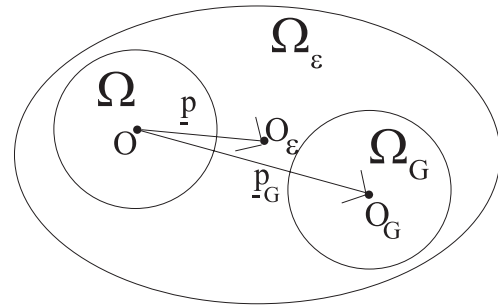


Fig. 1. The ghost obstacle experiment:  $\Omega$  obstacle,  $\Omega_G$  ghost obstacle,  $\Omega_\epsilon$  auxiliary set,  $O$  point contained in  $\Omega$  and origin of the coordinate system,  $O_G$  point contained in  $\Omega_G$ ,  $O_\epsilon$  point contained in  $\Omega_\epsilon$ .

Note that in a suitable cylindrical coordinate system with origin in the point  $O_G$  of the set  $\Omega_G$  or in the point  $O_\epsilon$  of the set  $\Omega_\epsilon$  we have respectively the following representation formulae for  $\Omega_G$  and  $\Omega_\epsilon$ ,

$$\begin{aligned} \Omega_G &= \{ \underline{x} = (r_1 \cos \phi, r_1 \sin \phi, z)^T = \\ &r_1 \hat{\phi} + z\mathbf{e}_3 \in \mathbf{R}^3 \mid 0 \leq r_1 < \xi_G(\phi, z), \\ &\phi \in [0, 2\pi), z \in [z_i^G, z_f^G] \}, \quad (16) \end{aligned}$$

and

$$\begin{aligned} \Omega_\epsilon &= \{ \underline{x} = (r_1 \cos \phi, r_1 \sin \phi, z)^T = \\ & r_1 \hat{\phi} + z \underline{e}_3 \in \mathbf{R}^3 \mid 0 \leq r_1 < \xi_\epsilon(\phi, z), \\ & \phi \in [0, 2\pi), z \in [z_i^*, z_f^*] \}. \end{aligned} \quad (17)$$

Remember that as already said with abuse of notation in the previous formulae,  $(r_1, \phi, z)$  denotes several cylindrical coordinate systems and that given  $\Omega$ ,  $\Omega_G$  such that  $\Omega \neq \emptyset$ ,  $\Omega_G \neq \emptyset$  and  $\Omega \cap \Omega_G = \emptyset$ , we must choose  $\Omega_\epsilon$  such that  $\Omega \subset \Omega_\epsilon$  and  $\Omega_G \subset \Omega_\epsilon$ .

Under assumptions (a) and (b) and other technical assumptions (see [2, 7]) applying the Pontryagin maximum principle, we find that the optimal scattered field  $\tilde{u}^s(\underline{x}, t)$ ,  $(\underline{x}, t) \in (\mathbf{R}^3 \setminus \bar{\Omega}) \times \mathbf{R}$ , and the corresponding adjoint variable  $\tilde{\varphi}(\underline{x}, t)$ ,  $(\underline{x}, t) \in (\mathbf{R}^3 \setminus \bar{\Omega}) \times \mathbf{R}$ , satisfy the first-order necessary optimality condition associated to the optimal control problem of equations (8), (1), (4), (5), and (6). That is, they are the solution of the following exterior problem for a system of two coupled wave equations,

$$\Delta \tilde{u}^s(\underline{x}, t) - \frac{1}{c^2} \frac{\partial^2 \tilde{u}^s}{\partial t^2}(\underline{x}, t) = 0, \quad (\underline{x}, t) \in (\mathbf{R}^3 \setminus \bar{\Omega}) \times \mathbf{R} \quad (18)$$

$$\tilde{u}^s(\underline{x}, t) = O\left(\frac{1}{r}\right), \quad r \rightarrow +\infty, \quad t \in \mathbf{R}, \quad (19)$$

$$\frac{\partial \tilde{u}^s}{\partial r}(\underline{x}, t) + \frac{1}{c} \frac{\partial \tilde{u}^s}{\partial t}(\underline{x}, t) = o\left(\frac{1}{r}\right), \quad r \rightarrow +\infty, \quad t \in \mathbf{R}, \quad (20)$$

$$\begin{aligned} -\frac{\partial \tilde{u}^s}{\partial t}(\underline{x}, t) + c\chi \frac{\partial \tilde{u}^s}{\partial \underline{n}(\underline{x})}(\underline{x}) &= g(\underline{x}, t) - \frac{(1+\chi)}{\varsigma} \tilde{\varphi}(\underline{x}, t), \\ (\underline{x}, t) &\in \partial\Omega \times \mathbf{R}, \end{aligned} \quad (21)$$

$$\Delta \tilde{\varphi}(\underline{x}, t) - \frac{1}{c^2} \frac{\partial^2 \tilde{\varphi}}{\partial t^2}(\underline{x}, t) = 0, \quad (\underline{x}, t) \in (\mathbf{R}^3 \setminus \bar{\Omega}) \times \mathbf{R}, \quad (22)$$

$$\tilde{\varphi}(\underline{x}, t) = O\left(\frac{1}{r}\right), \quad r \rightarrow +\infty, \quad t \in \mathbf{R}, \quad (23)$$

$$\frac{\partial \tilde{\varphi}}{\partial r}(\underline{x}, t) - \frac{1}{c} \frac{\partial \tilde{\varphi}}{\partial t}(\underline{x}, t) = o\left(\frac{1}{r}\right), \quad r \rightarrow +\infty, \quad t \in \mathbf{R}, \quad (24)$$

$$\begin{aligned} -\mu \frac{\partial \tilde{\varphi}}{\partial t}(\underline{x}, t) - c\chi\mu \frac{\partial \tilde{\varphi}}{\partial \underline{n}(\underline{x})}(\underline{x}) &= \\ -\lambda c(1+\chi) f_\epsilon(\underline{x})(\tilde{u}^s(\underline{p}_\epsilon(\underline{x}, t)) - u_G^s(\underline{p}_\epsilon(\underline{x}, t))), \\ (\underline{x}, t) &\in \partial\Omega \times \mathbf{R}, \end{aligned} \quad (25)$$

$$\lim_{t \rightarrow -\infty} \tilde{u}^s(\underline{x}, t) = 0, \quad \underline{x} \in \mathbf{R}^3 \setminus \bar{\Omega}, \quad (26)$$

$$\lim_{t \rightarrow +\infty} \tilde{\varphi}(\underline{x}, t) = 0, \quad \underline{x} \in \mathbf{R}^3 \setminus \Omega, \quad (27)$$

where  $\underline{p}_\epsilon(\underline{x})$  is a point belonging to  $\partial\Omega_\epsilon$  given by,

$$\begin{aligned} \underline{p}_\epsilon(\underline{x}) &= \frac{\xi_\epsilon(\phi, s(z))}{\xi(\phi, z)} [\underline{x} - (\underline{x}, \underline{e}_3)\underline{e}_3] + s((\underline{x}, \underline{e}_3))\underline{e}_3, \\ \underline{x} &= \xi \hat{\phi} + z \underline{e}_3 \in \partial\Omega \end{aligned} \quad (28)$$

where  $s(z) = z_i^* + (z - z_i)(z_f^* - z_i^*) / (z_f - z_i)$ ,  $z \in [z_i, z_f]$ , and  $f_\epsilon(\underline{x})$ ,  $\underline{x} \in \partial\Omega$ , is the function defined by,

$$\begin{aligned} f_\epsilon(\underline{x}(\phi, z)) &= \frac{v_\epsilon(\phi, z)}{v(\phi, z)}, \quad \underline{x} = \xi(\phi, z)\hat{\phi} + z\underline{e}_3 \in \partial\Omega, \\ \phi &\in [0, 2\pi), \quad z \in [z_i, z_f] \end{aligned} \quad (29)$$

with

$$\begin{aligned} v(\theta, \phi) &= \\ \left[ \xi^2(\phi, z) + \left( \frac{\partial \xi}{\partial \phi} \right)^2(\phi, z) + \xi^2(\phi, z) \left( \frac{\partial \xi}{\partial z} \right)^2(\phi, z) \right]^{\frac{1}{2}}, \\ \phi &\in [0, 2\pi), \quad z \in [z_i, z_f], \end{aligned} \quad (30)$$

$$\begin{aligned} v_\epsilon(\theta, \phi) &= \frac{(z_f^* - z_i^*)}{(z_f - z_i)} \cdot \left[ \xi_\epsilon^2(\phi, s(z)) + \right. \\ \left. \left( \frac{\partial \xi_\epsilon}{\partial \phi} \right)^2(\phi, s(z)) + \xi_\epsilon^2(\phi, s(z)) \left( \frac{\partial \xi_\epsilon}{\partial s(z)} \right)^2(\phi, s(z)) \right]^{\frac{1}{2}}, \\ \phi &\in [0, 2\pi), \quad z \in [z_i, z_f]. \end{aligned} \quad (31)$$

The relation between  $\tilde{\varphi}$  solution of equations (18)-(27) and the optimal control  $\tilde{\psi}$  solution of problem of equations (8), (1), (4), (5), and (6) is the following one,

$$\tilde{\psi}(\underline{x}, t) = -\frac{1}{\varsigma} \tilde{\varphi}(\underline{x}, t), \quad (\underline{x}, t) \in \partial\Omega \times \mathbf{R}. \quad (32)$$

For future convenience, we point out that,

$$ds_{\partial\Omega} = v(\phi, z) d\phi dz, \quad \phi \in [0, 2\pi), \quad z \in [z_i, z_f] \quad (33)$$

and

$$ds_{\partial\Omega_\epsilon} = v_\epsilon(\phi, z) d\phi dz, \quad \phi \in [0, 2\pi), \quad z \in [z_i, z_f]. \quad (34)$$

Note that in order to guarantee conditions of equations (26) and (27) we must choose the incoming field in a suitable class of functions (see [2]). This will be done in Section III. We note that the boundary conditions of equations (21) and (25) can be slightly modified to deal with the limit case  $\chi = +\infty$ .

### III. THE NUMERICAL SOLUTION OF THE FIRST ORDER OPTIMALITY CONDITION

Let  $B = \{ \underline{x} \in \mathbf{R}^3 \mid \|\underline{x}\| < 1 \}$  and  $\partial B$  be the boundary of  $B$ . We assume that  $u^i$ ,  $\tilde{u}^s$ ,  $u_G^s$  and  $\tilde{\varphi}$  can be approximated in a compact subset of the time axis by finite sums of time harmonic waves, that is,

$$\begin{aligned} u^i(\underline{x}, t) &\approx \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} \left[ a_{i,j} e^{-i\omega_i t} e^{i\omega_i(\underline{x}, \underline{\alpha}_j)/c} \right] \\ (\underline{x}, t) &\in \mathbf{R}^3 \times \mathbf{R}, \end{aligned} \quad (35)$$



$$\begin{aligned} \tilde{u}^s(\mathbf{x}, t) &\approx \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} \left[ a_{i,j} e^{-\imath \omega_i t} u_{\omega_i, \underline{\alpha}_j}^s(\mathbf{x}) \right], \\ (\mathbf{x}, t) &\in (\mathbf{R}^3 \setminus \bar{\Omega}) \times \mathbf{R}, \end{aligned} \quad (36)$$

$$\begin{aligned} u_G^s(\mathbf{x}, t) &\approx \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} \left[ a_{i,j} e^{-\imath \omega_i t} u_{G, \omega_i, \underline{\alpha}_j}^s(\mathbf{x}) \right], \\ (\mathbf{x}, t) &\in (\mathbf{R}^3 \setminus \bar{\Omega}) \times \mathbf{R}, \end{aligned} \quad (37)$$

$$\begin{aligned} \tilde{\varphi}(\mathbf{x}, t) &\approx \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} \left[ a_{i,j} e^{-\imath \omega_i t} \varphi_{\omega_i, \underline{\alpha}_j}(\mathbf{x}) \right], \\ (\mathbf{x}, t) &\in (\mathbf{R}^3 \setminus \bar{\Omega}) \times \mathbf{R}, \end{aligned} \quad (38)$$

where  $\imath \in \mathbf{C}$  is the imaginary unit,  $N_1, N_2$  are positive integers,  $a_{i,j} \in \mathbf{R}$ ,  $\omega_i \in \mathbf{R}$ ,  $\underline{\alpha}_j \in \partial B$ ,  $i = 1, 2, \dots, N_1$ ,  $j = 1, 2, \dots, N_2$  are suitable quantities,  $u_{G, \omega_i, \underline{\alpha}_j}^s$ ,  $i = 1, 2, \dots, N_1$ ,  $j = 1, 2, \dots, N_2$  are suitable functions and  $u_{\omega_i, \underline{\alpha}_j}^s(\mathbf{x})$ ,  $\varphi_{\omega_i, \underline{\alpha}_j}(\mathbf{x})$ ,  $\mathbf{x} \in \mathbf{R}^3 \setminus \Omega$ ,  $i = 1, 2, \dots, N_1$ ,  $j = 1, 2, \dots, N_2$ , are functions to be determined. Substituting the right hand side of equations(35)-(38) into equations (18)-(27) and defining  $\tilde{\zeta} = c\zeta$  we obtain that the space dependent part  $u_{\omega, \underline{\alpha}}^s(\mathbf{x})$ ,  $\mathbf{x} \in \mathbf{R}^3 \setminus \Omega$ , of the time harmonic components of  $\tilde{u}^s$ , and the space dependent part  $\varphi_{\omega, \underline{\alpha}}(\mathbf{x})$ ,  $\mathbf{x} \in \mathbf{R}^3 \setminus \Omega$ , of the time harmonic components of  $\tilde{\varphi}$ ,  $(\omega, \underline{\alpha}) = (\omega_i, \underline{\alpha}_j)$ ,  $i = 1, 2, \dots, N_1$ ,  $j = 1, 2, \dots, N_2$ , are solutions of the following set of systems of Helmholtz: for  $i = 1, 2, \dots, N_1$ ,  $j = 1, 2, \dots, N_2$ , and  $(\omega, \underline{\alpha}) = (\omega_i, \underline{\alpha}_j)$ , we have,

$$\left( \Delta u_{\omega, \underline{\alpha}}^s + \frac{\omega^2}{c^2} u_{\omega, \underline{\alpha}}^s \right)(\underline{x}) = 0 \quad (39)$$

$$\left( \Delta \varphi_{\omega, \underline{\alpha}} + \frac{\omega^2}{c^2} \varphi_{\omega, \underline{\alpha}} \right)(\underline{x}) = 0, \underline{x} \in \mathbf{R}^3 \setminus \bar{\Omega}, \quad (40)$$

$$\begin{aligned} \imath \omega u_{\omega, \underline{\alpha}}^s(\underline{x}) + c\chi \frac{\partial u_{\omega, \underline{\alpha}}^s}{\partial \underline{n}(\underline{x})}(\underline{x}) + c \frac{(1+\chi)}{\tilde{\zeta}} \varphi_{\omega, \underline{\alpha}}(\mathbf{x}) &= \\ b_{\omega, \underline{\alpha}}(\underline{x}), \underline{x} \in \partial\Omega, \end{aligned} \quad (41)$$

$$\begin{aligned} \imath \omega \mu \varphi_{\omega, \underline{\alpha}}(\underline{x}) - c\mu\chi \frac{\partial \varphi_{\omega, \underline{\alpha}}}{\partial \underline{n}(\underline{x})}(\underline{x}) + c\lambda(1+\chi) \cdot \\ \left( u_{\omega, \underline{\alpha}}^s(\underline{p}_\epsilon(\underline{x})) - u_{G, \omega, \underline{\alpha}}^s(\underline{p}_\epsilon(\underline{x})) \right) = 0, \underline{x} \in \partial\Omega, \end{aligned} \quad (42)$$

with the conditions at infinity,

$$\frac{\partial u_{\omega, \underline{\alpha}}^s(\underline{x})}{\partial r} - \imath \frac{\omega}{c} u_{\omega, \underline{\alpha}}^s(\underline{x}) = o\left(\frac{1}{r}\right), \quad r \rightarrow +\infty, \quad (43)$$

$$\frac{\partial \varphi_{\omega, \underline{\alpha}}(\underline{x})}{\partial r} + \imath \frac{\omega}{c} \varphi_{\omega, \underline{\alpha}}(\underline{x}) = o\left(\frac{1}{r}\right), \quad r \rightarrow +\infty, \quad (44)$$

where  $b_{\omega, \underline{\alpha}}(\mathbf{x}) = -\imath \omega e^{\imath \omega(\mathbf{x}, \underline{\alpha})/c} (1 + \chi(\underline{n}(\mathbf{x}), \underline{\alpha}))$ ,  $\mathbf{x} \in \partial\Omega$ . We remind that  $\underline{p}$  has been defined in (28).

Using equations (32) and (38) the relation of the adjoint variables  $\varphi_{\omega_i, \underline{\alpha}_j}$ ,  $i = 1, 2, \dots, N_1$ ,  $j = 1, 2, \dots, N_2$ , with the optimal control variable  $\tilde{\psi}$  can be expressed as follows,

$$\begin{aligned} \tilde{\psi}(\mathbf{x}, t) &\approx -\frac{c}{\tilde{\zeta}} \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} a_{i,j} e^{-\imath \omega_i t} \varphi_{\omega_i, \underline{\alpha}_j}(\mathbf{x}), \\ (\mathbf{x}, t) &\in \partial\Omega \times \mathbf{R}, 0 < \mu < 1. \end{aligned} \quad (45)$$

We propose a variation of the operator expansion method

presented in [17] to solve, using affordable computing resources, equations (39)-(44) when  $(\Omega; \chi)$  has a complex geometry and at least some of the wavelengths contained in the incoming wave packet of equation (35) (i.e., some of the quantities  $|2\pi/(\omega_i/c)|$ ,  $i = 1, 2, \dots, N_1$ ) are small compared with the characteristic dimension of the obstacle. In the numerical experiments presented in Section IV we consider problems where the ratio between the characteristic dimension of the obstacle and the wavelength of the time harmonic component of the incident waves goes up to approximately sixty.

As done in equation (35) let us consider acoustic incoming time harmonic plane waves whose space dependent part is given by,

$$u_{\omega, \underline{\alpha}}^i(\mathbf{x}) = e^{\imath \omega(\mathbf{x}, \underline{\alpha})/c}, \quad \mathbf{x} \in \mathbf{R}^3 \quad (46)$$

where  $c > 0$  is the wave propagation velocity,  $\omega \neq 0$  is the frequency of the wave,  $\underline{\alpha} \in \partial B$  is the wave propagation direction. Let us define the wave number  $k$  as  $k = \omega/c$ , we remind that  $2\pi/|k|$  is the wavelength of the plane wave of equation (46). Later we will choose  $\omega = \omega_i$ ,  $i = 1, 2, \dots, N_1$ , and  $\underline{\alpha} = \underline{\alpha}_j$ ,  $j = 1, 2, \dots, N_2$ .

Let us describe briefly the basic steps of the operator expansion method (see [2, 16, 17] for more details). Let  $(r_1, \phi, z)$  be the canonical cylindrical coordinate system introduced in Section II and let us assume that  $\partial\Omega$  is given by formula (11). The use of the operator expansion method to solve equations (39)-(44) is based on the following assumptions:

(a<sub>1</sub>) there exists a bounded simply connected open set  $\Omega_c$  such that  $\bar{\Omega}_c \subset \Omega$  with locally Lipschitz boundary  $\partial\Omega_c$ , given by,

$$\begin{aligned} \partial\Omega_c &= \{ \underline{x} = (r_1 \cos \phi, r_1 \sin \phi, z)^T \in \mathbf{R}^3 \mid \\ r_1 &= \xi_c(\phi, z), \phi \in [0, 2\pi), z \in [\tilde{z}_i, \tilde{z}_f] \} \end{aligned} \quad (47)$$

where  $\tilde{z}_i, \tilde{z}_f$  are two given real numbers such that  $\tilde{z}_i < \tilde{z}_f$ ,  $[\tilde{z}_i, \tilde{z}_f] \subset (z_i, z_f)$  and  $\xi_c$  is a single valued function sufficiently regular to make sense out of the formulae that follow. Note that we have:  $0 < \xi_c(\phi, z) < \xi(\phi, z)$ ,  $\phi \in [0, 2\pi)$ ,  $z \in (\tilde{z}_i, \tilde{z}_f)$ ,

(b<sub>1</sub>) for  $(\omega, \underline{\alpha}) = (\omega_i, \underline{\alpha}_j)$ ,  $i = 1, 2, \dots, N_1$ ,  $j = 1, 2, \dots, N_2$ , the functions  $u_{\omega, \underline{\alpha}}^s$ ,  $\varphi_{\omega, \underline{\alpha}}$  that solve the exterior problem (39)-(44) can be written as

single layer potentials with density functions (to be determined) supported on  $\partial\Omega_c$ , that is,

$$u_{\omega,\underline{\alpha}}^s(\mathbf{x}) = \int_{\partial\Omega_c} ds_{\partial\Omega_c}(\mathbf{y}) \Phi_{\underline{\omega}}^s(\mathbf{x}, \mathbf{y}) c_{\omega,\underline{\alpha}}(\mathbf{y}), \quad \mathbf{x} \in \mathbf{R}^3 \setminus \Omega \quad (48)$$

$$\varphi_{\omega,\underline{\alpha}}(\mathbf{x}) = \int_{\partial\Omega_c} ds_{\partial\Omega_c}(\mathbf{y}) \overline{\Phi_{\underline{\omega}}^s(\mathbf{x}, \mathbf{y})} f_{\omega,\underline{\alpha}}(\mathbf{y}), \quad \mathbf{x} \in \mathbf{R}^3 \setminus \Omega, \quad (49)$$

where  $ds_{\partial\Omega_c}$  is the surface measure defined on  $\partial\Omega_c$  that we assume to be given by:  $ds_{\partial\Omega_c}(\mathbf{y}(\phi, z)) = g_c(\phi, z) d\phi dz$ ,  $(\phi, z) \in U' = (0, 2\pi) \times (\tilde{z}_i, \tilde{z}_f)$ , where  $g_c$  is a sufficiently regular positive function,  $\Phi_{\underline{\omega}}^s(\mathbf{x}, \mathbf{y}) = \frac{e^{i\frac{\omega}{c}\|\underline{x}-\underline{y}\|}}{4\pi\|\underline{x}-\underline{y}\|}$ ,  $\underline{x}, \underline{y} \in \mathbf{R}^3$ ,  $\underline{x} \neq \underline{y}$ , is the fundamental solution of the Helmholtz operator on  $\mathbf{R}^3$  satisfying the radiation condition (43) and finally  $c_{\omega,\underline{\alpha}}(\mathbf{y})$ ,  $f_{\omega,\underline{\alpha}}(\mathbf{y})$ ,  $\mathbf{y} \in \partial\Omega_c$  are the density functions to be determined mentioned previously. We note that  $\overline{\Phi_{\underline{\omega}}^s}$ , the complex conjugate of  $\Phi_{\underline{\omega}}^s$ , satisfies the radiation condition (44);

- (c<sub>1</sub>) there exists a surface  $\partial\Omega_r$  boundary of a bounded simply connected open set  $\Omega_r$  representable with a formula analogous to formula (11) when we replace the function  $\xi$  with a suitable single valued function  $\xi_r$  such that  $\Omega_c \subset \Omega_r$  and such that the statements contained in (d<sub>1</sub>) hold. We refer to the surface  $\partial\Omega_r$  as “reference surface”;
- (d<sub>1</sub>) let  $\mathbf{y}_{\xi_c}(\underline{v}') = (\xi_c(\underline{v}') \cos(\phi'), \xi_c(\underline{v}') \sin(\phi'), z'^T$ ,  $\underline{v}' = (\phi', z')^T \in U'$  denote a (generic) point of  $\partial\Omega_c$ , we assume that the following perturbative expansions of  $u_{\omega,\underline{\alpha}}^s$  and  $\varphi_{\omega,\underline{\alpha}}$  hold,

$$u_{\omega,\underline{\alpha}}^s(\mathbf{x}) = \int_{U'} d\underline{v}' g_c(\underline{v}') \left( \Phi_{\underline{\omega}}^s(\mathbf{x}, \mathbf{y}_{\xi_c}(\underline{v}')) \cdot \sum_{s=0}^{+\infty} \tilde{c}_{k,\underline{\alpha},s}(\underline{v}') (\xi(\underline{v}') - \xi_r(\underline{v}'))^s \right), \quad \mathbf{x} \in \mathbf{R}^3 \setminus \Omega \quad (50)$$

$$\varphi_{\omega,\underline{\alpha}}(\mathbf{x}) = \int_{U'} d\underline{v}' g_c(\underline{v}') \left( \overline{\Phi_{\underline{\omega}}^s(\mathbf{x}, \mathbf{y}_{\xi_c}(\underline{v}'))} \cdot \sum_{s=0}^{+\infty} \tilde{f}_{k,\underline{\alpha},s}(\underline{v}') (\xi(\underline{v}') - \xi_r(\underline{v}'))^s \right), \quad \mathbf{x} \in \mathbf{R}^3 \setminus \Omega, \quad (51)$$

where  $d\underline{v}' = d\phi' dz'$  is the usual Lebesgue measure on  $U'$ . With abuse of notation we require that  $\tilde{c}_{k,\underline{\alpha},s}(\xi - \xi_r)^s = O((\xi - \xi_r)^s)$ ,  $\tilde{f}_{k,\underline{\alpha},s}(\xi - \xi_r)^s = O((\xi - \xi_r)^s)$  as  $\xi \rightarrow \xi_r$ ,  $s = 0, 1, 2, \dots$

Note that the surfaces  $\partial\Omega_c$  and  $\partial\Omega_r$  introduced here have been called “special” surfaces of the operator expansion method in Section I. We note that  $u_{\omega,\underline{\alpha}}^s$ ,  $\varphi_{\omega,\underline{\alpha}}$  given by equations (48) and (49) satisfy the

Helmholtz equations (39) and (40) and the “radiation” conditions at infinity of equations (43) and (44) for any choice of the density functions  $c_{\omega,\underline{\alpha}}$  and  $f_{\omega,\underline{\alpha}}$  that make possible differentiation under the integral sign. Using assumptions (a<sub>1</sub>)-(d<sub>1</sub>), substituting equations (50) and (51) into the boundary conditions (41), (42) and imposing the boundary conditions (41), (42) order by order in perturbation theory we obtain a sequence of systems of integral equations, that is a system made of two integral equations at each order in the expansion in powers of  $(\xi - \xi_r)$ . In fact remind that  $k = \omega/c$  is the wave number of the incoming plane wave and let  $U = (0, 2\pi) \times (z_i, z_f)$ ,  $\mathbf{x}_{\xi_r}(\underline{v})$ ,  $\underline{v} \in U$ , be a (generic) point belonging to  $\partial\Omega_r$ ,  $\nabla_{\mathbf{x}}$  be the gradient operator with respect to  $\mathbf{x} \in \mathbf{R}^3$ , for  $\nu = 0, 1, \dots$  let  $\hat{\phi}_k(\underline{v}) = (1/\nu k) \chi n(\underline{x}_{\xi}(\underline{v}))$ ,  $\underline{v} \in U$ ,  $\underline{Q}^{\nu}(\underline{v}, \mathbf{y}) = \frac{\partial^{\nu}}{\partial r_1^{\nu}} \nabla_{\mathbf{x}} \Phi_k((r_1 \cos \phi, r_1 \sin \phi, z)^T, \mathbf{y}) \big|_{r_1=\xi_r(\underline{v})}$ , and  $\mathcal{L}^{\nu}(\underline{v}, \mathbf{y}) = \frac{\partial^{\nu}}{\partial r_1^{\nu}} \Phi_k((r_1 \cos \phi, r_1 \sin \phi, z)^T, \mathbf{y}) \big|_{r_1=\xi_r(\underline{v})}$ ,  $\underline{v} \in U$ ,  $\mathbf{y} \in \mathbf{R}^3$ ,  $\mathbf{y} \notin \partial\Omega_r$ , arguing as in [16] it can be shown that the solution of problem (39)-(44) can be reduced to the solution of the following set of systems of integral equations in the unknowns  $c_{k,\underline{\alpha},s}(\underline{v}') = g_c(\underline{v}') \tilde{c}_{k,\underline{\alpha},s}(\underline{v}') (\xi(\underline{v}') - \xi_r(\underline{v}'))^s$ ,  $\underline{v}' \in U'$ ,  $f_{k,\underline{\alpha},s}(\underline{v}') = g_c(\underline{v}') \tilde{f}_{k,\underline{\alpha},s}(\underline{v}') (\xi(\underline{v}') - \xi_r(\underline{v}'))^s$ ,  $\underline{v}' \in U'$ ,  $s = 0, 1, 2, \dots$ ,

$$\int_{U'} d\underline{v}' K_{\xi_r, \xi_c}(\underline{v}, \underline{v}') c_{k,\underline{\alpha},s}(\underline{v}') + \frac{(1+\chi)}{\zeta k} \int_{U'} d\underline{v}' \nu \overline{\Phi_k(\underline{x}_{\xi_r}(\underline{v}), \underline{y}_{\xi_c}(\underline{v}'))} f_{k,\underline{\alpha},s}(\underline{v}') = d_{1,k,\underline{\alpha},s}(\underline{v}), \quad \underline{v} = (v_1, v_2)^T \in U, \quad s = 0, 1, 2, \dots \quad (52)$$

$$-\frac{\lambda(1+\chi)}{k} \hat{f}_{\epsilon}(\underline{v}) \int_{U'} d\underline{v}' [\nu c_{k,\underline{\alpha},s}(\underline{v}') \cdot \Phi_k(\underline{p} + \xi_c(v_1, s(v_2)) \hat{\phi}(v_1) + \hat{z}(v_2), \underline{y}_{\xi_c}(\underline{v}'))] + \mu \int_{U'} d\underline{v}' \overline{K_{\xi_r, \xi_c}(\underline{v}, \underline{v}')} f_{k,\underline{\alpha},s}(\underline{v}') = d_{2,k,\underline{\alpha},s}(\underline{v}), \quad \underline{v} = (v_1, v_2)^T \in U, \quad s = 0, 1, 2, \dots, \quad (53)$$

where  $K_{\xi_r, \xi_c}$  is given by,

$$K_{\xi_r, \xi_c}(\underline{v}, \underline{v}') = \left[ \Phi_k(\underline{x}_{\xi_r}(\underline{v}), \underline{y}_{\xi_c}(\underline{v}')) + \left( \frac{\hat{\phi}_k(\underline{v})}{\zeta k}, (\nabla_{\mathbf{x}} \Phi_k)(\underline{x}_{\xi_r}(\underline{v}), \underline{y}_{\xi_c}(\underline{v}')) \right) \right], \quad \underline{v} \in U, \underline{v}' \in U' \quad (54)$$

and  $\hat{f}_{\epsilon}(\underline{v}) = v_{\epsilon}(\underline{v})/v(\underline{v})$  (see formulae (30) and (31)), moreover we have,

$$d_{1,k,\underline{\alpha},0}(\underline{v}) = -e^{i k(\underline{x}_{\xi}(\underline{v}), \underline{\alpha})} [1 + \chi (n(\underline{x}_{\xi}(\underline{v})), \underline{\alpha})] \quad (55)$$

$$d_{2,k,\underline{\alpha},0}(\underline{v}) = -\nu(1+\chi) \frac{\lambda}{k} \hat{f}_{\epsilon}(\underline{v}) \cdot u_{G,\omega,\underline{\alpha}}^s(\underline{p} + \xi_c(v_1 + s(v_2)) \hat{\phi}(v_1) + \hat{z}(v_2)), \quad \underline{v} = (v_1, v_2)^T \in U, \quad (56)$$

and for  $s = 1, 2, \dots$ , we have,

$$d_{1,k,\underline{\alpha},s}(\underline{v}) = - \sum_{\nu=0}^{s-1} \frac{(\xi(\underline{v}) - \xi_r(\underline{v}))^{s-\nu}}{(s-\nu)!} \int_{U'} d\underline{v}' \cdot \left\{ \left[ \left( \phi_k(\underline{v}), Q^{s-\nu}(\underline{v}, \underline{y}_{\xi_c}(\underline{v}')) \right) + \mathcal{L}^{s-\nu}(\underline{v}, \underline{y}_{\xi_c}(\underline{v}')) \right] \cdot c_{k,\underline{\alpha},\nu}(\underline{v}') + \frac{(1+\chi)}{\tilde{\zeta}k} \mathcal{L}^{s-\nu}(\underline{v}, \underline{y}_{\xi_c}(\underline{v}')) f_{k,\underline{\alpha},s}(\underline{v}') \right\} \quad \underline{v} \in U, \quad (57)$$

$$d_{2,k,\underline{\alpha},s}(\underline{v}) = -\mu \sum_{\nu=0}^{s-1} \frac{(\xi(\underline{v}) - \xi_r(\underline{v}))^{s-\nu}}{(s-\nu)!} \left\{ \left[ \left( \phi_k(\underline{v}), Q^{s-\nu}(\underline{v}, \underline{y}_{\xi_c}(\underline{v}')) \right) + \mathcal{L}^{s-\nu}(\underline{v}, \underline{y}_{\xi_c}(\underline{v}')) \right] \cdot f_{k,\underline{\alpha},\nu}(\underline{v}') \right\}, \quad \underline{v} \in U. \quad (58)$$

Figure 2 shows an example of the relation between the sets  $\Omega$ ,  $\Omega_r$ ,  $\Omega_c$ .

Roughly speaking assumptions  $(a_1)$ ,  $(b_1)$  say that the spatial parts of the time harmonic components of the scattered field and of the auxiliary variable can be represented as single layer potentials generated by suitable densities defined on a surface  $\partial\Omega_c$  contained in the interior of the smart obstacle  $(\Omega; \chi)$ . Assumptions  $(c_1)$ ,  $(d_1)$  say that the smart obstacle is not far from being a more regular obstacle  $(\Omega_r; \chi)$  and that the density functions of the single layer potentials of equations (48) and (49) can be expressed as a power series of the “distance” between the boundary of  $\Omega$  and the boundary of  $\Omega_r$ . Assumptions  $(a_1)$ ,  $(b_1)$  make possible to formulate the boundary conditions (41), (42) as Fredholm integral equations of the first kind avoiding singular kernels even when  $\partial\Omega$  is only Lipschitz continuous. These integral equations are ill posed so that to solve them numerically we try to take care of their ill-posedness using the perturbation series of equations (50) and (51) whose convergence is assumed in  $(c_1)$ ,  $(d_1)$ . Note that thank to these last assumptions, we have reduced the solution of the optimal control problem to the solution of a set of systems of integral equations of the first kind whose ill-posedness is controlled via the perturbation approach.

Since we want to solve these systems of integral equations when the smart obstacle has complex geometry and the wavelength of the incoming wave is small compared to the characteristic dimension of the obstacle

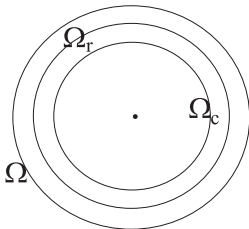


Fig. 2. An example of the relation between  $\Omega$ ,  $\Omega_r$ ,  $\Omega_c$ .

we need to discretize the integral equations using finite dimensional vector spaces of high dimension. The use of suitable wavelet bases to represent the unknown densities, the data and the integral kernels of the integral equations (52) and (53) allows us to approximate the integral equations in finite-dimensional vector spaces of high dimension with sparse systems of linear equations that can be solved with affordable computing resources even when they involve hundreds of thousands or millions of unknowns and equations. This is due to the “sparsifying properties” of the wavelet basis used.

Let us introduce the wavelet basis used in the experiments of Section IV. Let  $L^2(U')$  and  $L^2(U)$  be the Hilbert spaces of square integrable real functions with respect to the Lebesgue measure defined on  $U'$  and on  $U$  respectively. As shown in [14], we generate a wavelet basis of  $L^2(U')$  and of  $L^2(U)$  using the tensor product and a suitable affine transformation of a wavelet basis of  $L^2((0, 1))$ . The wavelet basis of  $L^2((0, 1))$  used in the numerical experiments presented in Section IV is generated via the multi-resolution analysis [18, 14] starting from the following orthonormal piecewise polynomial functions of  $L^2((0, 1))$ . Let us define three functions that are known as wavelet “mother” functions. Let  $c_1 = 0.44721359549996$ ,  $c_2 = 1.3416407864998$  and let  $V$  be the real matrix given by,

$$V = ((v_{i,j})_{i=1,2,3, j=1,2,3,4}) = \begin{pmatrix} 1 & -1 & -1 & 1 \\ -c_1 & c_2 & -c_2 & c_1 \\ -c_2 & -c_1 & c_1 & c_2 \end{pmatrix} \quad (59)$$

we define the following piecewise polynomial functions defined in the interval  $(0, 1)$ ,

$$\Psi_i(z) = \begin{cases} v_{i,1}, & 0 < x < 1/4, \\ v_{i,2}, & 1/4 \leq x < 1/2, \\ v_{i,3}, & 1/2 \leq x < 3/4, \\ v_{i,4}, & 3/4 \leq x < 1, \end{cases} \quad i = 1, 2, 3, \quad (60)$$

and let  $\psi_{i,m,\nu}(z)$ ,  $z \in (0, 1)$ ,  $i = 1, 2, 3$ ,  $m = 0, 1, \dots$ ,  $\nu = 0, 1, 2, \dots, 4^m - 1$  be the function defined by,

$$\psi_{i,m,\nu}(z) = \begin{cases} 4^{m/2} \psi_i(4^m z - \nu), & z \in (\nu 4^{-m}, (\nu + 1) 4^{-m}), \\ 0, & z \in (0, 1) \setminus (\nu 4^{-m}, (\nu + 1) 4^{-m}). \end{cases} \quad (61)$$

As shown in [14] the set  $W_{a,b}$  defined as follows,

$$W_{a,b} = \left\{ \hat{\psi}_{j,m,\nu}(y) = \frac{1}{\sqrt{b-a}} \psi_{j,m,\nu} \left( \frac{y-a}{(b-a)} \right), y \in (a, b), \right. \\ \left. j = 1, 2, 3, m = 0, 1, \dots, \nu = 0, 1, 2, \dots, 4^m - 1 \right\} \\ \cup \left\{ L_0(y) = \frac{1}{\sqrt{b-a}}, y \in (a, b) \right\} \quad (62)$$

is an orthonormal basis of  $L^2((a, b))$ ,  $a < b$ ,  $a, b \in \mathbf{R}$ . We note that in equation (62) we have used the announced affine transformation to go from  $L^2((0, 1))$  to  $L^2((a, b))$  and that the wavelet mother functions  $\Psi_1$  and  $\Psi_2$  of  $L^2((0, 1))$  defined in equation (60) have two vanishing

moments, that is:  $\int_0^1 dx x^m \Psi_i(x) = 0$ ,  $m = 0, 1$ ,  $i = 1, 2$  while the wavelet mother function  $\Psi_3$  of  $L^2((0, 1))$  has only one vanishing moment, that is  $\int_0^1 dx \Psi_3(x) = 0$  and we have  $\int_0^1 dx x \Psi_3(x) \neq 0$ .

Starting from this wavelet basis and using the tensor product the integral equations (52) and (53) can be reduced to a system of infinitely many linear equations whose unknowns are the coefficients of the representation on the wavelet basis of the density functions  $c_{k,\underline{\alpha},s}$ ,  $f_{k,\underline{\alpha},s}$ ,  $s = 0, 1, \dots$ . Truncating the wavelet expansions we reduce the approximate solution of the integral equations (52) and (53) to the solution of (eventually high dimensional) linear systems. These linear systems are approximated with sparse linear systems using a simple procedure that consists in setting to zero the elements of the matrices representing the integral kernels smaller in absolute value than a given threshold, in this way we obtain very sparse matrices. In fact, thank to the sparsifying properties of the wavelet basis introduced above, the kernels of the integral equations are approximated satisfactorily by the sparse matrices obtained with the procedure described above. Using these sparse matrices as coefficient matrices of linear systems that approximate those considered above we obtain sparse linear systems that approximate the integral equations (52) and (53). Finally, these sparse linear systems are solved with a suitable parallelization of the conjugate gradient method (see [16, 17]).

#### IV. NUMERICAL EXPERIENCE

In this Section we assume that the smart obstacle  $(\Omega; \chi)$  and the ghost obstacle  $(\Omega_G; \chi_G)$  have boundary acoustic impedance equal to infinity, i.e.:  $\chi = \chi_G = +\infty$ , that is the smart obstacle and the ghost obstacle are acoustically hard obstacles. This implies that the equations written in the previous Sections must be slightly changed to be adapted to deal with hard obstacles.

In the numerical experiments we consider as incoming acoustic fields time harmonic plane waves whose space dependent part is given by equation (46) or wave packets of the form,

$$u^i(\mathbf{x}, t) = e^{-[\underline{\gamma}, \mathbf{x}] - ct]^2 / 4\zeta^2}, \quad (\mathbf{x}, t) \in \mathbf{R}^3 \times \mathbf{R} \quad (63)$$

where  $\underline{\gamma} \in \partial B$  and  $\zeta \in \mathbf{R}$ ,  $\zeta \neq 0$ . The obstacle  $\Omega$  in all the experiments is given by a smart simplified model of the NASA space shuttle. The original model of the NASA space shuttle (see Fig. 3(a)) has been modified obtaining the simplified NASA space shuttle (see Fig. 3(b)) in order to have an obstacle whose boundary can be represented with a single valued function in a suitable cylindrical coordinate system, that is in order to have the representation of equation (11) of the boundary of the obstacle for a suitable choice of the cylindrical coordinate system and of the function  $\xi$ . The data relative to the original obstacle (see Fig. 3(a)) are

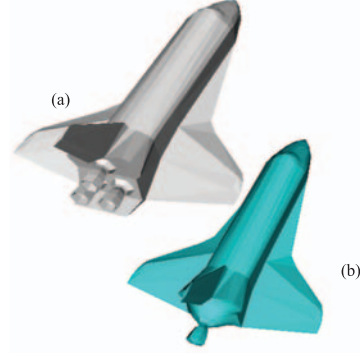


Fig. 3. (a) The NASA space shuttle, and (b) Simplified version of the NASA space shuttle.

available in the website <http://avalon.viewpoint.com/>. The physical dimensions of the shuttle are expressed in *units* where  $1 \text{ unit} = 56.14/14 \text{ meters}$ . The maximum length of the shuttle in the direction of the symmetry axis of its main body corresponds to  $14 \text{ units}$ . The space shuttle is an acoustically hard obstacle, this justifies the choice  $\chi = +\infty$  made previously and the sound speed in the air at sea level is  $331.45 \text{ meters/seconds}$  so that we choose  $c = 331.45 \text{ meters/seconds}$  that corresponds to  $c \simeq 82.65 \text{ units/seconds}$ . Note that the  $z$ -axis of the cylindrical coordinate system used to represent the obstacle is chosen to be the “symmetry” axis of the main body of the simplified NASA space shuttle (see Fig. 4). In the following the lengths are expressed in *units*.

We choose  $\underline{p} = (0, 7.5, 1.5)^T \in \mathbf{R}^3$ ,  $z_i = z_i^G = -7$ ,  $z_f = z_f^G = 7$ ,  $z_i^* = -16$ ,  $z_f^* = 16$ ,  $s(z) = z_i^* + (z_f^* - z_i^*)(z - z_i)/(z_f - z_i)$ ,  $z \in [z_i, z_f]$ ,  $\xi_\epsilon(\phi, s(z)) = \sqrt{(1 - s(z)^2/d^2)/(\cos^2 \phi/a^2 + \sin^2 \phi/b^2)}$  where  $a = 8$ ,  $b = 13$ ,  $d = 16$  and the ghost  $\Omega_G$  is the translation of  $\underline{p}_G = (0, 15, 2)^T \in \mathbf{R}^3$  of  $\Omega$ . In Fig. 4 we show the setting of the ghost obstacle experiment studied here and the coordinate axes. The surfaces  $\xi_r$ ,  $\xi_c$  have been chosen such that the kernels  $K_{\xi_r, \xi_c}$  and  $\Phi(\mathbf{x}_{\xi_r}, \mathbf{y}_{\xi_c})$  defined in Section III are continuous with their first partial derivatives and we use always the first two terms of the operator expansion series given in formulae (50) and (51).

We remind that the website:

<http://www.econ.univpm.it/recchioni/scattering/w16> contains some auxiliary material that helps the understanding of the numerical experiments discussed here including stereographic and virtual reality applications.

The first experiment shows the effect due to the smart character of the obstacle for different incident time harmonic plane waves given in equation (46) and several values of the parameter  $\lambda$ ,  $0 \leq \lambda \leq 1$ . We remind that we are assuming  $\lambda \geq 0$ ,  $\mu \geq 0$ ,  $\mu + \lambda = 1$ ,  $\varsigma = 1$  and that the smart obstacle reproduces exactly the field generated by the ghost obstacle on  $\partial\Omega_\epsilon$  when  $\mu = 0$ ,

$\lambda = 1$ . The  $L^2$  norm of the space dependent part of the pressure current  $\tilde{\Psi}$  employed to obtain the ghost effect can be considered as a measure of the price paid in order to have the smart behaviour of the obstacle. We note that the quantity  $n_{\psi,k}^\lambda$  defined in (equation (66)) is proportional to the  $L^2$  norm of the space dependent part of the pressure current mentioned above.

Let us introduce the following quantities,

$$d_{s,G,\omega}^\lambda = \sqrt{\frac{\int_{\partial\Omega_\epsilon} |u_{\omega,\alpha}^s(\mathbf{x}) - u_{G,\omega,\alpha}^s(\mathbf{x})|^2 ds_{\partial\Omega_\epsilon}(\mathbf{x})}{\int_{\partial\Omega_\epsilon} |u_{G,\omega,\alpha}^s(\mathbf{x})|^2 ds_{\partial\Omega_\epsilon}(\mathbf{x})}} \quad (64)$$

$$d_{p,G,\omega} = \sqrt{\frac{\int_{\partial\Omega_\epsilon} |u_{p,\omega,\alpha}^s(\mathbf{x}) - u_{G,\omega,\alpha}^s(\mathbf{x})|^2 ds_{\partial\Omega_\epsilon}(\mathbf{x})}{\int_{\partial\Omega_\epsilon} |u_{G,\omega,\alpha}^s(\mathbf{x})|^2 ds_{\partial\Omega_\epsilon}(\mathbf{x})}}, \quad (65)$$

$$n_{\psi,\omega}^\lambda = \sqrt{\int_{\partial\Omega} |\varphi_{\omega,\alpha}(\mathbf{x})|^2 ds_{\partial\Omega}(\mathbf{x})}, \quad (66)$$

$$e_\omega^\lambda = \frac{d_{s,G,\omega}^\lambda}{d_{p,G,\omega}}, \quad (67)$$

where  $u_{p,\omega,\alpha}^s$  is the field scattered by  $(\Omega; \chi)$  as a passive obstacle, and  $n_{\psi,\omega}^\lambda$ , as said above, is the quantity that measures the ‘‘size’’ of the pressure current required to get the smart effect.

Table 1 shows the ghost effect obtained. In fact from left to right Table 1 shows the value of  $\frac{\mu}{\lambda}$ ,  $\omega/c$ ,  $R_T$ ,  $d_{p,G,\omega}$ ,  $d_{s,G,\omega}^\lambda$ ,  $e_\omega^\lambda$  and  $n_{\psi,\omega}^\lambda$ , where  $R_T$  is defined as the ratio between the characteristic dimension of the obstacle and the wavelength of the incident time harmonic plane wave. The propagation direction of the incident plane wave in the coordinate system shown in Fig. 4 is  $\underline{\alpha} = (\sin(\pi/4) \cos(\pi/4), \sin(\pi/4) \sin(\pi/4), \cos(\pi/4))^T$ .

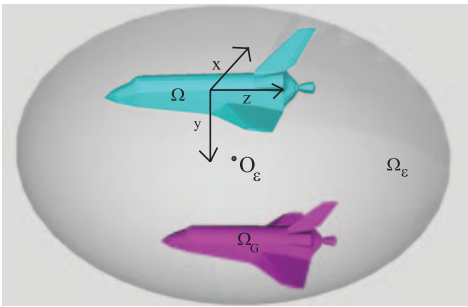


Fig. 4. The ghost obstacle experiment: the sets  $\Omega$  (smart obstacle),  $\Omega_G$  (ghost obstacle),  $\Omega_\epsilon$  (auxiliary set) and the coordinate system.

Note that when the ratio  $\frac{\mu}{\lambda}$  goes to zero, i.e., when  $\lambda$  goes to 1, the ghost effect increases, in fact  $d_{s,G,\omega}^\lambda$  decreases to zero and the price paid to obtain the ghost effect increases that is  $n_{\psi,\omega}^\lambda$  increases. Moreover, the column  $e_\omega^\lambda$  of Table 1 shows that when the wave number,  $\frac{\omega}{c}$ , increases the smart effect slightly deteriorates. In this experiment to solve the ghost obstacle scattering problem

Table 1. The ghost effect.

$\frac{\mu}{\lambda}$	$\frac{\omega}{c}$	$R_T$	$d_{p,G,\omega}$	$d_{s,G,\omega}^\lambda$	$e_\omega^\lambda$	$n_{\psi,\omega}^\lambda$
0.1	$10\pi$	60	2.2791	0.5150	0.2259	31.632
0.001	$10\pi$	60	2.2791	1.872e-02	8.2112e-03	333.04
0.1	$2\pi$	14	0.5325	0.1003	0.1883	48.464
0.001	$2\pi$	14	0.5325	2.9671e-03	5.5718e-03	88.168
0.1	1	$\frac{7}{\pi}$	1.9351	0.2068	0.1068	28.290
0.001	1	$\frac{7}{\pi}$	1.9351	3.2942e-03	1.7036e-03	35.841

we have used 262144 real variables when  $R_T = 60$ , 16384 real variables when  $R_T = 14$  and 1024 real variables when  $R_T = 7/\pi$ , that is we have solved linear systems of 262144, 16384, 1024 equations and unknowns when  $R_T$  is equal to 60, 14,  $7/\pi$  respectively. The CPU time required to compute the field scattered by the smart obstacle when  $R_T = 60$  is about 800 hours on the SP5 machine of the CASPUR center. This execution time includes the computation of the field scattered by the ghost obstacle as passive obstacle. Note that using a parallel implementation of the numerical code the clock time needed to compute the field scattered by the smart obstacle reduces to about six hours when we use 128 processors. We note that due to their condition number the linear systems used in this experiment must be solved with some care.

The second experiment shows the scattering from a time dependent incoming wave of equation (63) where  $\underline{\gamma} = (\sin(\pi/4) \cos(\pi/4), \sin(\pi/4) \sin(\pi/4), \cos(\pi/2))^T$  (in the coordinate system of Fig. 4, and  $\zeta = 1/(2\pi)$ ). In formulae (35), (36), (37), and (38) we have used the Gauss-Hermite quadrature rule with 400 nodes to approximate the Fourier transforms in the conjugate variable of time  $\omega$  that give  $u^i$ ,  $\tilde{u}^s$ ,  $u_G^s$ ,  $\tilde{\varphi}$  respectively. Indeed with the choice  $\zeta = 1/(2\pi)$  to get a satisfactory approximation of the incoming wave of equation (63) on an adequate compact set of the time axis only 30 wave numbers are needed. In order to compute the time harmonic components of the scattered waves of equations (36) and (37) we have used 1024, 4096, 16384 real variables to solve the time harmonic problems according with the value of  $R_T$  considered. We have chosen  $\frac{\mu}{\lambda} = 10^{-10}$ ,  $\mu = 1 - \lambda$ ,  $\varsigma = 1$ . We note that a rough estimate of the CPU time required to carry out this experiment is 5000 hours on the SP5 machine of the CASPUR center.

Let  $\partial D$  be the boundary of the sphere  $D$  having center in  $O_\epsilon = O + \underline{p}$  (see Figs. 1 and 4) and radius 15.

Note that  $D$  contains  $\Omega$ ,  $\Omega_G$ . Using the canonical spherical coordinate system  $(r, \theta, \rho)$  we have,

$$\partial D = \{ \underline{x} = \underline{p} + (15 \sin \theta \cos \rho, 15 \sin \theta \sin \rho, 15 \cos \theta)^T \in \mathbf{R}^3, \theta \in [0, \pi], \rho \in [0, 2\pi) \}. \quad (68)$$

Figure 5 shows the field scattered by the passive obstacle  $u_p^s(\mathbf{x}, t)$ , by the smart obstacle  $u^s(\mathbf{x}, t)$  and by

the ghost obstacle  $u_G^s(\mathbf{x}, t)$  when  $\mathbf{x} \in \partial D$ , for three time values:  $t = t_1 = -0.08, t_2 = 0.03, t_3 = 0.24$ .

Figure 6 shows the value of the incoming field when  $\mathbf{x} \in \partial D$  and  $t = t_1, t = t_2, t = t_3$ .

We note that in Fig. 5 the second and third columns are very similar, that is, the field scattered by the smart obstacle (second column) behaves as the field scattered by the ghost obstacle (third column). Furthermore, when the incoming acoustic field goes through  $D$  we can see that the passive and the ghost obstacle reacts in a different way when  $t = t_1$  and  $t = t_3$  and in a similar way when  $t = t_2$ . This effect is due to the fact that  $D$  is sphere with center in  $O_\epsilon = O + \underline{p}$  that is in a point lying between the passive and ghost obstacles. So that, when  $t = t_1$  the front of the incoming acoustic field (see Fig. 6 first column) is on the side where the passive obstacle is ( $\phi \in [\pi/4, \pi]$ ) and when  $t = t_3$  (see Fig. (6) third column) the incoming acoustic field is on the side where the ghost obstacle is ( $\phi \in [3\pi/2, 2\pi]$ ).

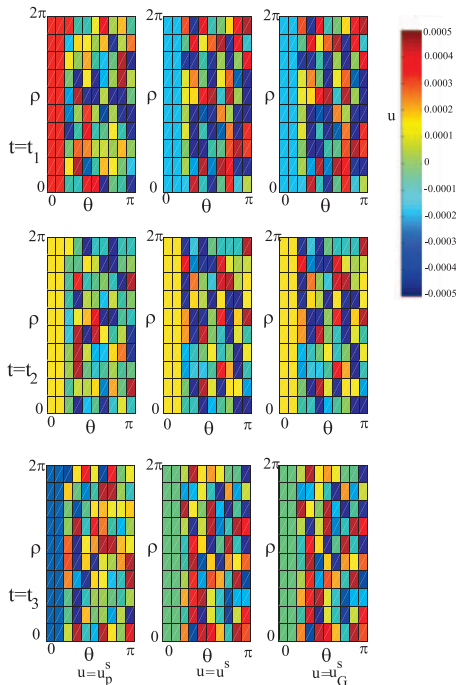


Fig. 5. The ghost effect on the surface of a sphere.

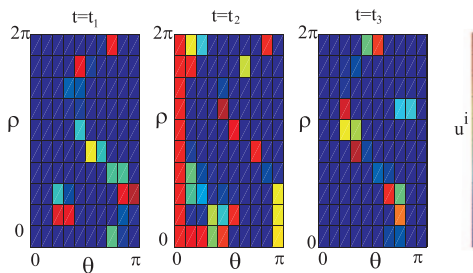


Fig. 6. The effect of the incoming field on the surface of a sphere.

**ACKNOWLEDGEMENT**

It is a pleasure to thank Mrs. Claudia Truini and Mr. Piero Lanucara of CASPUR (Roma, Italy) for the helpful assistance in the realization of the stereographic applications shown in the website associated with this paper.

**REFERENCES**

- [1] E. Mecocci, L. Misici, M. C. Recchioni, and F. Zirilli, "A new formalism for time dependent wave scattering from a bounded obstacle," *Journal of the Acoustical Society of America*, vol. 107, pp. 1825–1840, 2000.
- [2] F. Mariani, M. C. Recchioni, and F. Zirilli, "The use of the Pontryagin maximum principle in a furtivity problem in time-dependent acoustic obstacle scattering," *Waves in Random Media*, vol. 11, pp. 549-575, 2001.
- [3] L. Fatone, M. C. Recchioni, and F. Zirilli, "Some control problems for the Maxwell equations related to furtivity and masking problems in electromagnetic obstacle scattering," in *Mathematical and Numerical Aspects of Wave Propagation. Waves 2003*, G. C. Cohen, E. Heikkola, P. Joly, and P. Neittaanmaki Editors, Springer Verlag, Berlin, pp. 189-194, 2003.
- [4] L. Fatone, M. C. Recchioni, and F. Zirilli, "A masking problem in time dependent acoustic obstacle scattering," *Acoustics Research Letters Online*, vol. 5, no 2, pp. 25-30, 2004.
- [5] L. Fatone, M. C. Recchioni, and F. Zirilli, "Furtivity and masking problems in time dependent electromagnetic obstacle scattering," *Journal of Optimization Theory and Applications*, vol. 121, pp. 223-257, 2004.
- [6] L. Fatone, M. C. Recchioni, and F. Zirilli, "Mathematical models of "active" obstacles in acoustic scattering," in *Control and Boundary Analysis*, J. Cagnol and J. P. Zolesio Editors, Lecture Notes in Pure and Applied Mathematics, vol. 240, Marcel Dekker/CRC Press, Boca Raton, Fl. USA, pp. 119-129, 2005.
- [7] L. Fatone, G. Pacelli, M. C. Recchioni, and F. Zirilli, "Optimal control methods for two new classes of smart obstacles in time dependent acoustic scattering," *Journal of Engineering Mathematics*, vol. 56, pp. 385–413, 2006.
- [8] L. Fatone, M. C. Recchioni, A. Scoccia, and F. Zirilli, "Direct and inverse scattering problems involving smart obstacles," *Journal of Inverse and Ill-Posed Problems*, vol. 13, pp. 247-257, 2005.
- [9] L. Fatone, M. C. Recchioni, and F. Zirilli, "A method to solve an acoustic inverse scattering problem involving smart obstacles," *Waves in Random and Complex Media*, vol. 16, pp. 433-455, 2006.

- [10] L. Fatone, M. C. Recchioni, and F. Zirilli, “New scattering problems and numerical methods in acoustics,” in *Recent Research Developments in Acoustics*, S.G. Pandalai Managing Editor. Transworld Research Network, Kerala India, Vol. II, pp. 39-69, 2005.
- [11] L. Fatone, M. C. Recchioni, and F. Zirilli, “A numerical method to solve an acoustic inverse scattering problem involving ghost obstacles”, *Journal of Inverse and Ill-Posed Problems*, vol. 15, pp. 57–82, 2007.
- [12] D. Colton and R. Kress, *Integral Equation Methods in Scattering Theory*, John Wiley, New York, 1983.
- [13] L. Fatone, M. C. Recchioni, and F. Zirilli, “A numerical method for time dependent acoustic scattering problems involving smart obstacles and incoming waves of small wavelengths,” in *Mathematical Modeling of Wave Phenomena*, B. Nilsson, L. Fishman Editors, AIP Conference Proceedings, vol. 834, AIP Publ.,Melville, New York, pp. 109-121, 2006.
- [14] L. Fatone, M. C. Recchioni, and F. Zirilli, “New wavelet bases made of piecewise polynomial functions: approximation theory, quadrature rules and applications to kernel sparsification and image compression,” submitted to *SIAM Journal on Scientific Computing*.
- [15] J. Nečas, *Les Méthodes Directes en Théorie des Équations Elliptiques*. Masson & Cie. Publ., Paris, 1967.
- [16] M. C. Recchioni and F. Zirilli, “The use of wavelets in the operator expansion method for time dependent acoustic obstacle scattering,” *SIAM Journal on Scientific Computing*, vol. 25, pp. 1158–1186, 2003.
- [17] L. Fatone, G. Rao, M. C. Recchioni, and F. Zirilli, “High performance algorithms based on a new wavelet expansion for time dependent acoustic obstacle scattering,” *Communications in Computational Physics*, vol. 2, pp. 1139–1173, 2007.
- [18] S. Mallat, “Multiresolution approximation and wavelets”, *Transactions of the American Mathematical Society*, vol. 315, pp. 69–88, 1989.



nance.

**Lorella Fatone** Laurea in Mathematics cum Laude (1995), Ph. D. in Computational Mathematics and Operational Research, research fellow Universit di Modena and Reggio Emilia since 2001. Main research interests: numerical approximation of partial differential equations, mathematical finance.



tion, direct and inverse acoustic and electromagnetic scattering, mathematical finance.

**Cristina Recchioni** Laurea in Mathematics cum Laude (1987), associate professor of Mathematics at the Universit Politecnica delle Marche, Ancona since 2001. Author of over 60 scientific publications most of them published in international journals. Main research interests: numerical optimization, direct and inverse acoustic and electromagnetic scattering, mathematical finance.



and inverse acoustic and electromagnetic scattering, processing of remotely sensed data, mathematical finance.

**Francesco Zirilli** Laurea in Mathematics (1971), Laurea in Physics (1972), professor Universit di Roma “La Sapienza” since 1985. Author of over 150 scientific publications most of them published in international journals. Main research interests: partial differential equations, quantum field theory, nonlinear analysis, numerical optimization, direct and inverse acoustic and electromagnetic scattering, processing of remotely sensed data, mathematical finance.

# Analysis of a Singly-Fed Circularly Polarized Electromagnetically Coupled Patch Antenna

A. Hajiaboli and M. Popović

Department of Electrical and Computer Engineering, McGill University, Montreal, Canada  
 amir.hajiaboli@mail.mcgill.ca, milica.popovich@mcgill.ca

**Abstract** – This paper presents analysis of a broadband circularly polarized electromagnetically coupled patch antenna (EMCP) fed through a coaxial probe. The bandwidth of the antenna is investigated numerically and through experiment. The 12% bandwidth of the measured return loss implies broadband behavior considering the operating frequency and the inherent bandwidth limitations of the microstrip antenna structure. The changes in circular polarization bandwidth were investigated using finite-element-method (FEM) software and the results suggest that the increase in separation between the patches causes a decrease of 3dB bandwidth and the degradation of the minimum value of axial ratio. An axial ratio bandwidth of 2% is achieved in the reported EMCP structure. In addition, we have applied a modal analysis using finite-difference time-domain (FDTD) simulation to reveal the simultaneous excitation of  $TM_{01}$  and  $TM_{10}$  modes at around 1.3GHz. These results help explain the broadband and circular polarization characteristics of the EMCP structure under investigation.

**Keywords:** EMCP antenna, bandwidth, return loss, axial ratio, and modal analysis.

## I. INTRODUCTION

Achieving circular polarization is a challenge in microstrip antennas. Various single-layer microstrip structures have been tested for this purpose. In general, these can be categorized into two groups: multiple-feed (e.g. dual-feed) and single-feed structures (with modified patches, e.g. corner truncated structures). The noted configurations share the same principle for obtaining the circular polarization: the dimensions are modified in order to provide propagation of two orthogonal modes with close resonant frequencies. Then, the antenna is excited at a frequency between the two resonant frequencies to ensure approximately equal amplitude for both modes. The location of the feed point is chosen strategically to result in a phase difference of  $90^\circ$  between the two modes [1].

It has been shown [1] that, in the case of dual-feed microstrip structures, the trade-off for the wide axial ratio (AR) bandwidth is the narrow bandwidth in the return loss. Further, the feeding structure would increase the

complexity and overall size of the antenna. On the other hand, in the single-feed structures, such as corner truncated configurations, the AR bandwidth is narrow (on the order of 1%) [1]. The noted antennas, therefore, cannot be considered as desirable structures with optimized characteristics for applications that necessitate circular polarization.

Electromagnetically coupled patch (EMCP) antenna was first introduced in 1983 [2] for broadband applications. Further investigations revealed its circularly polarized characteristics [3-5] and the possibility of an EMCP design with acceptable AR and return loss bandwidth. Most of the previously tested EMCP structures are based on the dual-feed excitation, truncated corner patches or the circular patch. A recent work [3] reports a systematic design procedure that results in the return loss bandwidth of 43% and the AR bandwidth of 8% in C-band. The reported procedure is based on optimizing the dimensions of a corner truncated square patch microstrip antenna.

In this paper, we analyze and report on a singly-fed coaxially fed EMCP antenna for circular polarization applications. Unlike the structure discussed in [3], the patches are not corner truncated and the feed is a coaxial probe placed off the patch diagonal. The bandwidth of the antenna has been measured and compared with the simulation results. The axial ratio has been calculated with finite-element tool (HFSS, Ansoft Co.). Our parametric study reports on the variation of the return loss and axial ratio with the change in separation between two antenna patches. Finally, a modal analysis using finite-difference time-domain (FDTD) simulation explains the broadband and the polarization characteristics of the antenna by revealing the simultaneous excitation of standing modes on the patches around 1.3GHz.

## II. THE ANTENNA STRUCTURE

Three-dimensional view of the antenna is shown in Fig. 1. The antenna consists of two patches separated by a foam spacer of thickness  $d$  and  $\epsilon_r=1$ . The patches are etched on a substrate with relative permittivity  $\epsilon_r=3.38$ , loss tangent  $\tan \delta = 0.0027$ , and thickness of 0.81mm. Another substrate of identical specifications is stacked to the fed-patch. These two substrates for the fed-patch are



glued together to result in total thickness of 1.74 mm. We assume a 0.12 mm thick air gap between the two glued substrates. The dimensions of interest are as follows: fed-patch 64mm×68mm, upper patch 72 mm × 72 mm, and the finite ground plane 100mm×100mm. The coaxial probe feed, modeled according to work reported in [6, 7], is placed at x=10mm and y=4mm from the lower left hand corner of the lower patch as shown in Fig. 2. To reduce the computational burden, in both the FDTD and the FEM simulations, an infinite ground plane was considered.

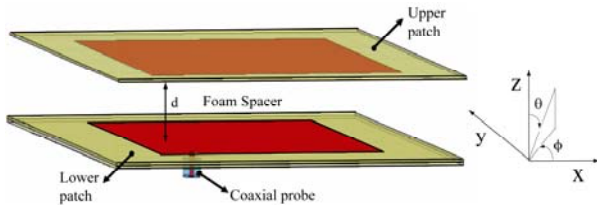


Fig. 1. Three-dimensional view of the electromagnetically coupled patch antenna.

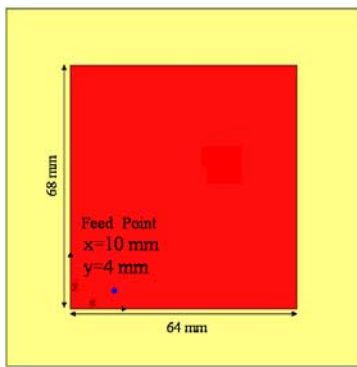


Fig. 2. Lower patch of antenna in Fig. 1, with indicated coordinates of the feed point.

III. RETURN LOSS

The measured and simulated results of input return loss ( $S_{11}$ ) are shown in Figs. 3 and 4 for  $d = 22$  mm and  $d=18$ mm. The center frequency of the -10dB return loss in simulation and experimental results is close to 1.3GHz and varies slightly with the change in  $d$ .

Figure 5 shows the percentage of -10dB bandwidth versus  $d$  both for simulation and measurement. Figure 5 demonstrates that the measured bandwidth of the antenna linearly increases with  $d$ , with the maximum measured bandwidth of 12% for  $d=30$ mm.

As it can be observed in Figs. 3, 4, and 5, around 1.25 GHz and for the foam spacer less than 24 mm, which is the focus of this paper, there is a good agreement between the simulated and the measured results. A discrepancy between the measurement and simulation can be noted around 1.8 GHz, and we suspect

that it was caused by several approximations of the simulation model. First, we assume  $\epsilon_r=1$  for the foam layer. Second, we neglect the electromagnetic properties of the glue. Finally, the scattering effects of the finite ground plane in the constructed antenna (not modeled in the simulations to reduce the computational burden) also represent a possible source of error. As discussed in [8], by considering the finite ground plane it is possible to obtain better simulated results around 1.8 GHz.

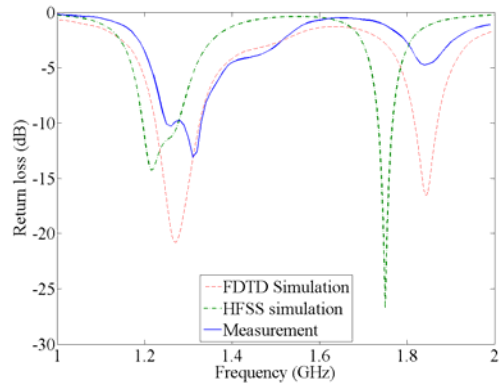


Fig. 3. Return loss for the separation between the lower and upper patch (foam spacer)  $d=18$  mm.

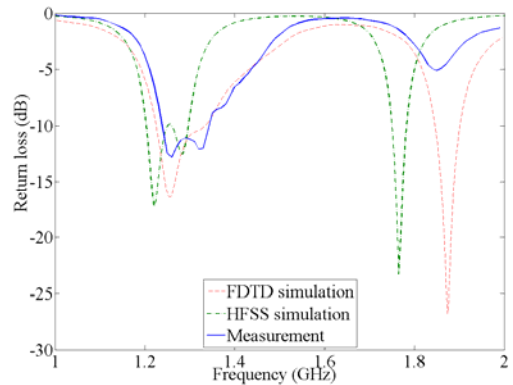


Fig. 4. Return loss for the separation between the lower and upper patch (foam spacer)  $d=22$ mm.

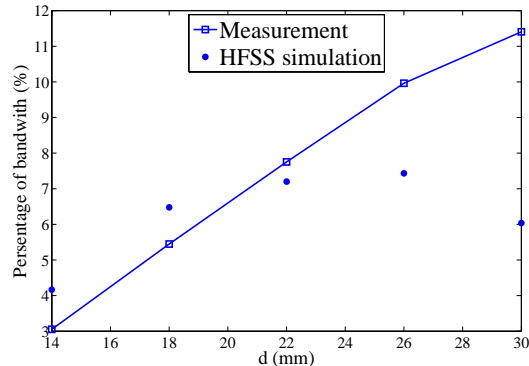


Fig. 5. Percentage of bandwidth for different values of the separation between the lower and upper patch (thickness of the foam spacer),  $d$ .

IV. AXIAL RATIO

Figure 6 shows the axial ratio computed with the finite-element software (HFSS) for different values of  $d$ . The minimum value of the axial ratio (0.7dB) occurs at  $d=20\text{mm}$  at the center frequency of 1.24 GHz. Considering the frequency shift between the measurements and simulations noted in Figs. 3 and 4, we can anticipate the minimum axial ratio value in the measurement at approximately 1.3GHz – the center frequency of the -10dB bandwidth. The maximum 3dB bandwidth of axial ratio (2.2 %) is observed for  $d=16\text{mm}$  at the center frequency of 1.23GHz. We note that the increase in  $d$  results in a reduction of the axial ratio bandwidth. As can be seen in Fig. 6, for  $d > 22 \text{ mm}$ , the 3dB bandwidth of axial ratio is negligible.

Figure 7 shows the variation of axial ratio at  $\varphi=0^\circ$  ( $x$ - $z$  plane) versus  $\theta$  at 1.24 GHz and for different values  $d$ . Again, it can be observed that, as  $d$  increases, the beamwidth of the 3-dB axial ratio decreases. For  $d>22\text{mm}$ , the beamwidth is zero. At  $d = 20 \text{ mm}$  the 3-dB beamwidth of axial ratio is around  $90^\circ$ .

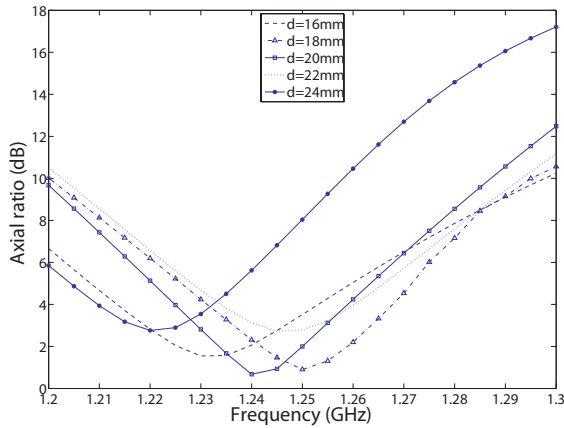


Fig. 6. Axial ratio vs. frequency for different values of  $d$ .

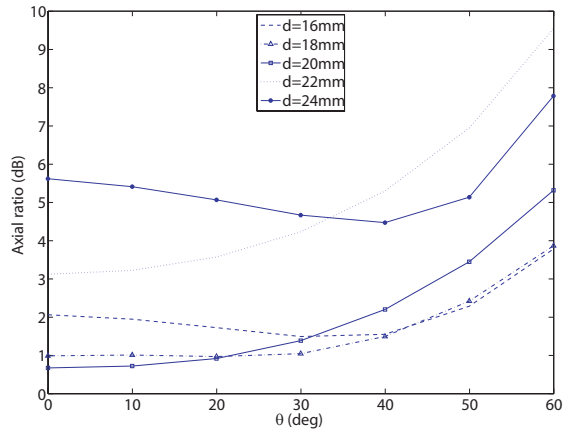
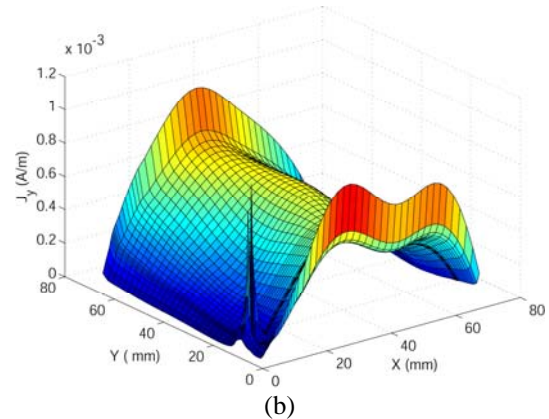
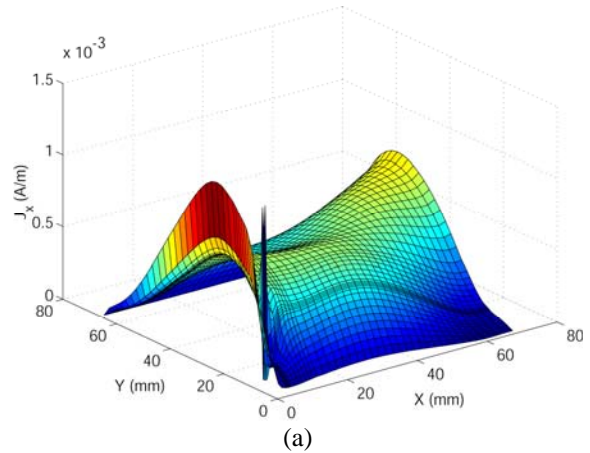


Fig. 7. Axial ratio at  $\varphi=0^\circ$  vs.  $\theta$  for different values of  $d$  at 1.24 GHz.

V. MODAL ANALYSIS

In order to explain the broadband behavior and the circular polarization of the antenna, we simulate the antenna using finite-difference time domain (FDTD) method and extract the current distribution on the patches at different frequencies.

In order to obtain the current distribution ( $J_x$  and  $J_y$ ) we calculate and save the current distribution at each time step at the late simulation time to ensure the higher order modes have diminished and only dominant modes exist on the patches. We then apply the Fourier transform on these current distributions to obtain the current distributions in frequency domain. As depicted in Fig. 8, the current distribution at 1.3 GHz corresponds to  $TM_{01}$  and  $TM_{10}$  modes with approximately the same amplitude. The simultaneous excitation of these two modes matches with the broadband behavior obtained at around 1.3GHz. Further, since the modes are spatially orthogonal this explains the circular polarization when the appropriate phase conditions are met for certain values of the separation between the upper and lower patch ( $d$ ).



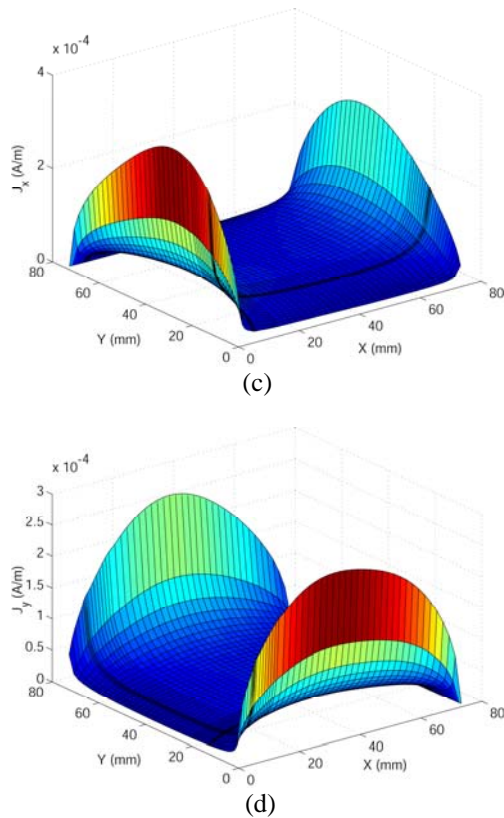


Fig. 8. Current distribution at 1.3GHz on the patches, illustrating the simultaneous excitation of  $TM_{01}$  and  $TM_{10}$  modes. a)  $J_x$  on lower patch, b)  $J_y$  on lower patch, c)  $J_x$  on upper patch, and d)  $J_y$  on upper patch.

## VI. CONCLUSION

A broadband circularly polarized electromagnetically coupled patch antenna has been analyzed through simulation and experiment. The measured bandwidth increases linearly with the separation between the lower and upper antenna patch  $d$  and the maximum bandwidth of 12% around 1.33 GHz (L-band) occurs for  $d=30$ mm. Through FEM simulations, we observed that the 3-dB bandwidth of axial ratio degrades as the separation of the patches  $d$  increases. A modal analysis on the current distributions obtained by FDTD has been performed and the current distribution at 1.3GHz clearly shows the simultaneous excitation of  $TM_{01}$  and  $TM_{10}$  modes with nearly equal amplitudes on the lower and upper patches. This can explain the broadband and circular polarization characteristics of the antenna.

## ACKNOWLEDGEMENTS

This work was funded by Natural Science and Engineering Research Council (NSERC) of Canada Discovery Grant and by the Le Fonds Québécois de la Recherche sur la Nature et les Technologies Nouveaux Chercheurs Grant.

## REFERENCES

- [1] G. Kumar and K. P. Ray, *Broadband microstrip antennas*, Artech House Inc. 2003.
- [2] A. Sabban "A new broadband stacked two layer microstrip antenna," *IEEE AP-S Int. Symp. Digest*, pp. 63-66, 1983.
- [3] K. L. Chung and A. S. Mohan, "A systematic design method to obtain broadband characteristics for singly-fed electromagnetically coupled patch antennas for circular polarization," *IEEE Trans. Antennas and Propagation*, vol. 51, no. 12, 2003.
- [4] K. T. V. Reddy and G. Kumar, "Stacked microstrip antennas for broadband circular polarization," *IEEE AP-S Int. Symp. Digest*, vol. 3, pp. 420-423, July 2001.
- [5] R. Q. Lee, T. Talty, and K. F. Lee, "Circular polarization characteristics of parasitic microstrip antennas," *IEEE AP-S Int. Symp. Digest*, vol. 1, pp. 310-313, June 1991.
- [6] A. Hajiaboli and M. Popovic, "FDTD Sub-Cell Modeling of the Inner Conductor of the Coaxial Feed: Accuracy and Convergence Analysis," *IEEE Trans. On Magnetics*, vol. 43, no. 4, pp. 1361 – 1364, Apr. 2007.
- [7] A. Hajiaboli and M. Popovic, "Comparison of Three FDTD Modeling Techniques for Coaxial Feed," *IEEE AP-S Int. Symp. Digest*, pp. 3432-3435, 2006.
- [8] FEKO website, knowledge base, application notes, "Singly- Fed Circularly Polarized Electromagnetically Coupled Patch Antenna," <http://www.feko.info/knowledge-base/application-notes/folder070329-1415/document070329-1416>



**Amir Hajiaboli** obtained a B.Sc. from University of Tehran in 2000 and a M.Sc. with highest distinction from Iran University of Science and Technology in 2003 both in electrical engineering. Since Sept. 2004, he has been pursuing his Ph.D. in electrical engineering with the Department of Electrical and Computer Engineering at

McGill University in Montreal Canada. His research interests focus on computational electromagnetic for antennas and bio-medical applications. He has extensive experiences in developing computational electromagnetic codes using finite-difference time-domain method. He is currently working on simulating the electrodynamics of light interaction with retinal photoreceptors.



**Milica Popović** received her B.Sc. (1994) from University of Colorado (Boulder, Colorado, USA) and M.Sc. (1997) and Ph.D. (2001) degrees from Northwestern University (Evanston, Illinois, USA), all in electrical engineering. She is currently an associate professor with the Department of Electrical and Computer Engineering at

McGill University in Montréal, Canada. Her research interests focus on numerical methods in computational electromagnetics for bio-medical applications, in particular: breast cancer screening with microwaves, wireless implants for physiological research and light interaction with retinal photoreceptor cells. On the teaching side, her efforts include improvement methods for instruction of introductory electromagnetics courses.

# Stochastic Optimization of a Patch Antenna

<sup>1</sup>S. Alfonzetti, <sup>2</sup>G. Borzi, <sup>1</sup>E. Dilettoso, and <sup>1</sup>N. Salerno

<sup>1</sup> Dipartimento di Ingegneria Elettrica, Elettronica e dei Sistemi (DIEES)  
Università di Catania, 95125 Catania, Italy

<sup>2</sup> Dipartimento di Ingegneria Civile  
Università di Messina, 98166 Messina, Italy

**Abstract** – The paper describes an efficient technique for optimizing the shape of the patch of a multi-band antenna by means of Genetic Algorithms and the hybrid FEM–RBCI method, for the analysis of open-boundary scattering and radiation electromagnetic field problems. The admissible rectangular patch area is logically and regularly subdivided into rectangular sub-areas, coinciding with the trace of the tetrahedral edge element mesh on the patch surface. In this way the relevant matrices of the finite element algebraic system, computed at the beginning of the optimization, remain unchanged, even if the patch is changed by inserting some metallic sub-areas. Moreover, in order to reduce the computing time of the iterative solver, the solution of a similar patch configuration is used as the initial guess for the solver.

**Keywords:** finite element method, genetic algorithms, micro-strip antennas, and optimization methods.

## I. INTRODUCTION

Antenna design is a topic of great importance in electromagnetics. It involves the selection of physical parameters to achieve optimal gain, pattern performance, bandwidth, and so on, subject to some specified constraints. If a trial and error process is used for antenna design, the designer is required to have great experience and intuition; so innovative design methods are required [1,2]. In addition to producing results with excellent performance, this also gives unconventional and non-intuitive physical realizations.

For personal communications, multi-band antennas are of particular interest. In this paper, we consider a dual-band antenna design using a single patch, and an efficient technique is described for optimizing the shape of the patch by means of Genetic Algorithms (GA) [3,4] and the hybrid FEM/RBCI (Finite Element Method – Robin Boundary Condition Iteration) method, for the analysis of open-boundary electromagnetic scattering [5] and radiation [6] problems.

In Section 2 the FEM/RBCI method for the microstrip antenna is briefly recalled. In Section 3 the

optimization procedure is outlined. Results are shown in Section 4 and our conclusions follow in Section 5.

## II. THE FEM-RBCI METHOD FOR A MICROSTRIP ANTENNA

Consider a patch antenna recessed in a perfectly conducting (PEC) plane; the cavity is filled with a homogeneous lossless material with relative dielectric permittivity  $\epsilon_r$  and relative magnetic permeability  $\mu_r$ . The original microstrip antenna is a rectangle of size  $w_a \times l_a$ , residing on top of a parallelepipedal cavity the dimensions of which are  $w_c \times l_c$  and depth  $h$  (see Fig.1).

A relevant application is to design a patch that operates at the two frequencies of the Global Positioning System (GPS): 1227 and 1572 MHz.

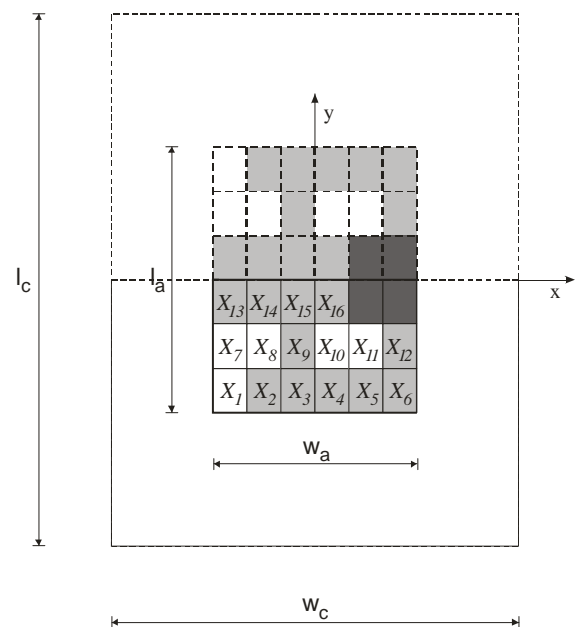


Fig. 1. Top view of cavity and patch.

In order to make the antenna work in this frequency range, the following values are selected:  $w_a=5.36$  cm,

$l_a=7.02$  cm,  $w_c=2w_a$ ,  $l_c=2l_a$ ,  $h=0.48$  cm,  $\epsilon_r=4.4$ ,  $\mu_r=1$ . The same device was optimized in [2]. The antenna is powered by a coaxial cable and irradiates in an unbounded vacuum medium.

To apply FEM-RBCI, the vacuum region is truncated, at a distance of  $d=5$  cm from the cavity, to a bounded one by means of a fictitious boundary  $B_F$ , which encloses the antenna aperture and the patch (see Fig.2). In the bounded domain thus obtained, the vector Helmholtz equation holds for the electric field,

$$\nabla \times (\mu_r^{-1} \nabla \times \bar{\mathbf{E}}) - k_0^2 \epsilon_r \bar{\mathbf{E}} = 0 \quad (1)$$

where  $\mu_r$  and  $\epsilon_r$  are the relative magnetic permeability and electrical permittivity, respectively, and  $k_0$  is the free-space wave number, given by  $k_0^2 = \omega^2 \epsilon_0 \mu_0$ , with  $\omega$  being the angular frequency and  $\mu_0$  and  $\epsilon_0$  the free-space permeability and permittivity, respectively. Homogeneous Dirichlet ( $\hat{\mathbf{n}} \times \bar{\mathbf{E}} = 0$ ) conditions hold on the PEC surfaces of the cavity, the PEC plane, and the patch surface.

A Robin (mixed) boundary condition is assumed on  $B_F$ ,

$$\hat{\mathbf{n}} \times \nabla \times \bar{\mathbf{E}} + jk_0 \hat{\mathbf{n}} \times (\hat{\mathbf{n}} \times \bar{\mathbf{E}}) = \bar{\mathbf{U}} \quad (2)$$

where  $\hat{\mathbf{n}}$  is the outward normal to  $B_F$  and  $\bar{\mathbf{U}}$  is an unknown vector tangent to  $B_F$ .

The internal conductor of the coaxial cable is assumed to carry an impressed density current  $\mathbf{J}_{\text{int}}$  which represents the known source of the antenna. Since the source is electrically short and small, it can be modeled as a current filament [7]. The source can be expressed as,

$$\bar{\mathbf{J}}_{\text{int}} = I_{\text{int}} \delta(x - x_f) \delta(y - y_f) \hat{\mathbf{z}} \quad (3)$$

where  $x_f=w_a/3$  and  $y_f=0$  specifies the feed position,  $I_{\text{int}}$  denotes the electric current magnitude, and  $\delta(x)$  is the Dirac delta function.

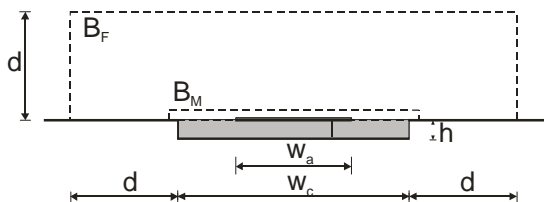


Fig. 2. Cross section of the FEM domain at plane  $y=0$ .

Discretizing the domain by tetrahedral edge elements, the FEM leads to the matrix equation,

$$\mathbf{A}\mathbf{E} = \mathbf{B}_0 + \mathbf{B}\mathbf{U} \quad (4)$$

where  $\mathbf{A}$  is a complex and symmetric matrix,  $\mathbf{B}_0$  is due to the source,  $\mathbf{B}$  is a rectangular matrix, whose entries are the Kronecker delta, and links the vector  $\bar{\mathbf{U}}$  with the right hand side of the FEM system,  $\mathbf{E}$  is the array of the expansion coefficients for the electric field and  $\mathbf{U}$  is the array whose generic entry is given by,

$$U_j = \int_{B_F} \bar{\mathbf{U}} \cdot \bar{\mathbf{w}}_j dS \quad (5)$$

in which  $\bar{\mathbf{w}}_j$  is the generic edge form function. Since  $\mathbf{U}$  is unknown, system of equation (4) cannot be solved.

Let us now consider another surface,  $B_M$ , lying between the antenna and the fictitious boundary (see Fig. 2). At minimum,  $B_M$  can be selected as coinciding with the antenna aperture itself. The total field outside  $B_M$  can be expressed as [8],

$$\begin{aligned} \bar{\mathbf{E}}(\bar{\mathbf{r}}) = \int_{B_M} & \left[ \bar{\bar{\mathbf{G}}}(\bar{\mathbf{r}}, \bar{\mathbf{r}}') \cdot (\hat{\mathbf{n}}' \times \nabla' \times \bar{\mathbf{E}}(\bar{\mathbf{r}}')) + \right. \\ & \left. + \nabla \times \bar{\bar{\mathbf{G}}}(\bar{\mathbf{r}}, \bar{\mathbf{r}}') \cdot (\hat{\mathbf{n}}' \times \bar{\mathbf{E}}(\bar{\mathbf{r}}')) \right] ds' \end{aligned} \quad (6)$$

where the dyadic Green's function, which takes account of the presence of the ground plane, is given by [8],

$$\begin{aligned} \bar{\bar{\mathbf{G}}}(\bar{\mathbf{r}}, \bar{\mathbf{r}}') = & \bar{\bar{\mathbf{G}}}_0(\bar{\mathbf{r}}, \bar{\mathbf{r}}') - \bar{\bar{\mathbf{G}}}_0(\bar{\mathbf{r}}, \bar{\mathbf{r}}'') + \\ & + 2\hat{\mathbf{z}}\hat{\mathbf{z}}g_0(\bar{\mathbf{r}}, \bar{\mathbf{r}}'') \end{aligned} \quad (7)$$

where  $\bar{\mathbf{r}}''$  is the symmetrical of  $\bar{\mathbf{r}}'$  with respect to the ground plane and,

$$\bar{\bar{\mathbf{G}}}_0(\bar{\mathbf{r}}, \bar{\mathbf{r}}') = \left( \bar{\bar{\mathbf{I}}} + k_0^{-2} \nabla \nabla \right) g_0(\bar{\mathbf{r}}, \bar{\mathbf{r}}') \quad (8)$$

$$g_0(\bar{\mathbf{r}}, \bar{\mathbf{r}}') = \frac{1}{4\pi|\bar{\mathbf{r}} - \bar{\mathbf{r}}'|} e^{-jk_0|\bar{\mathbf{r}} - \bar{\mathbf{r}}'|} \quad (9)$$

A similar expression for  $\nabla \times \bar{\mathbf{E}}$  is easily obtained from equation (6), so that an integral equation is derived which links  $\bar{\mathbf{U}}$  to  $\bar{\mathbf{E}}$  [5]. Note that, since  $B_M$  and  $B_F$  do not intersect with each other, singularities are avoided in this integral equation. The discrete form of the equation reads [5],

$$\mathbf{U} = \mathbf{M}\mathbf{E} \quad (10)$$

where  $\mathbf{M}$  is a rectangular matrix in which null columns appear for the internal edges not involved in the computation.

Equations (4) and (10) together form the global algebraic system of the FEM-RBCI method, which can be conveniently solved by an iterative scheme as follows:

1) Select an arbitrary first guess for  $\mathbf{U}$ ; 2) solve equation (4) for  $\mathbf{E}$ , by means of a standard conjugate gradient solver (COCG); 3) obtain an improved guess for  $\mathbf{U}$  by means of equation (10); 4) if the procedure has converged, i.e. a user-selected end iteration tolerance  $\tau$  is satisfied, stop; otherwise go to 2). This scheme can be seen as a two-block Gauss-Seidel iterative method,

$$\begin{aligned} \mathbf{E}^{(n)} &= \mathbf{A}^{-1}\mathbf{B}_0 + \mathbf{A}^{-1}\mathbf{B}\mathbf{U}^{(n)} \\ \mathbf{U}^{(n+1)} &= \mathbf{M}\mathbf{E}^{(n)} \end{aligned} \quad (11)$$

In this way the symmetry of matrix  $\mathbf{A}$  is fully exploited. Moreover, since this procedure converges in a few iterations, it also minimizes the number of multiplications of the dense matrix  $\mathbf{M}$  by a vector.

### III. OPTIMIZATION OF THE ANTENNA

The admissible rectangular patch surface is regularly subdivided into rectangular sub-areas, which coincide with the trace of the tetrahedral finite element mesh on the patch itself (see Fig. 1). The optimal antenna is designed by making these sub-areas metallic or not, although the four sub-areas at the middle-right part of the patch are always filled with metal in order to fix the feed point.

Only the lower half of the domain is considered, in order to exploit the symmetry of the problem. We can therefore operate on 16 sub areas, the design variables  $x_1, x_2, \dots, x_{16}$ , each of which can assume the value 1 (filled) or 0 (empty): they form the GA chromosome, giving a total of  $2^{16}$  different configurations.

The objective function to minimize is chosen as [2],

$$f = |S_{11}|_1 + 0.1|S_{11}|_2 + |S_{11}|_3 \quad (12)$$

where  $|S_{11}|_k$ , for  $k=1, 2, 3$ , refers to the return loss at the three frequency points: 1227, 1400 and 1572 MHz, respectively. The return loss is defined as,

$$|S_{11}| = \left| \frac{Z_{in} - Z_0}{Z_{in} + Z_0} \right| \quad (13)$$

where  $Z_{in}$  is the input impedance at the feed and  $Z_0=50\Omega$ . After the electric field  $\mathbf{E}$  along the source edges has been obtained, by means of the FEM/RBCI, the voltage drop along the current filament can be calculated. Thus, the input impedance  $Z_{in}$  can be obtained.

We point out that the objective function  $f$  as given in equation (12) will not guarantee pattern integrity for the shaped patch: the objective function can be combined with a penalization criterion to drive the GA search towards topologies for which pattern connectivity is maintained [9]. However, at this point our focus will only be on optimization of the return loss.

During the optimization procedure, the mesh remains unchanged, hence the domain discretization is only performed once at the beginning of the optimization. Moreover, before starting the evaluation of  $|S_{11}|_k$  for a GA population, a structure in which all the sub-areas are non-metallic, except for those near the source current filament (dark gray areas in Fig.1), is selected. For this first patch configuration, corresponding to the null chromosome, the FEM/RBCI matrices  $\mathbf{A}$ ,  $\mathbf{B}$  and  $\mathbf{M}$  are computed and stored. Adding a sub-area of metal to the patch is equivalent to forcing a homogeneous Dirichlet boundary condition for the electric field on the edges lying on that sub-area. This, in turn, is equivalent to dropping the corresponding rows and columns in matrix  $\mathbf{A}$ . Matrices  $\mathbf{B}$  and  $\mathbf{M}$  remain unchanged for a fixed  $k$ . In this way, the whole preprocessing phase consists of modifying some Dirichlet boundary conditions, and the FEM matrices are recomputed only three times, once for each frequency in equation (12), during the fitness evaluation of a GA population, thus saving a great amount of computing time.

Moreover, before evaluating the objective functions, the GA population is ordered, taking into account the Hamming distance between the chromosomes, starting from the chromosome with more bits equal to 0, in such a way that configurations having similar patch shapes will be contiguously ordered. Then, by using solution  $\mathbf{E}$  for the electric field of the previous configuration as the initial guess for the iterative conjugate gradient solver (COCG) in the FEM/RBCI analysis of the next configuration, the number of iterations, of both the solver and the FEM/RBCI, is reduced and a further saving in the overall computing time is obtained.

Fig.3 is a flowchart of the whole optimization procedure as described above.

### IV. RESULTS

The formulation described in Section 2 was implemented in ELFIN [10], a finite element code developed by the authors for electromagnetic CAD research, which employs zero-order tetrahedral edge elements to solve three-dimensional electromagnetic scattering and radiation problems.

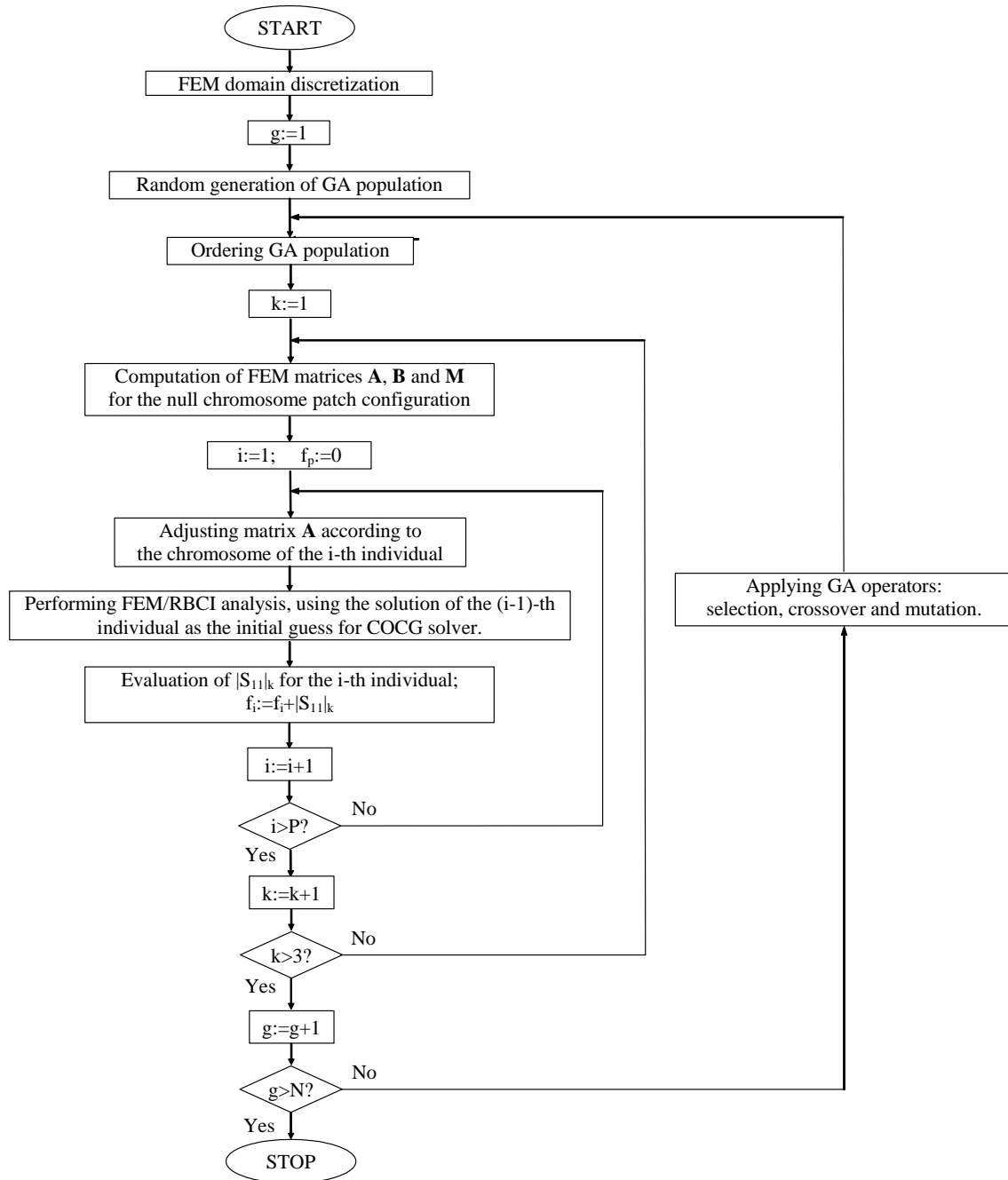


Fig. 3. A flowchart of the whole optimization procedure.

The finite element mesh used was made up of 10133 elements with 12661 edges. The feed source was discretized with four edges.

The COCG solver [11] with diagonal preconditioning was used for the solution of the various FEM complex symmetric systems of equation (4); the stopping criterion used for COCG was that described as criterion 2 in [12], with an end-iteration tolerance of

$\delta=0.05\%$ . The RBCI end-iteration tolerance was set to  $\tau = 1\%$  and convergence was reached, on average, in about five iterations.

In order to calculate the return loss, for a fixed configuration and a single frequency, the following computing times are required: a time  $T_p$  for the preprocessing phase, a time  $T_c$  for the construction of the FEM/RBCI system, i.e. for the computation of matrices

**A**, **B** and **M**, and a time  $T_s$  for the solution of the global algebraic system, equations (4) and (10). The whole computing time  $T$  for a single evaluation of  $|S_{11}|$  is thus,

$$T = T_p + T_c + T_s \quad (14)$$

with the above data,  $T_p$  is about 35% of  $T$ ,  $T_c$  is about 10% of  $T$  and  $T_s$  is about 55% of  $T$ . Solving the problem on a 3.2 GHz Pentium IV workstation with 4Gb RAM,  $T$  is about 50 s.

In optimization by GAs, the population size was set to  $P=30$  individuals: each individual is a 16-bit chromosome relating to a patch configuration. The reproduction process, which randomly creates a new generation from the old one, was chosen by tournament selection with a shuffling technique, to choose random pairs for mating, and elitism was also used. The crossover process, by means of which individuals exchange portions of chromosomes from one generation to the other, was 2-point crossover with a probability  $p_c$  varying from 0.3 to 0.7 as the optimization proceeds. The mutation process, by means of which some random flips in the chromosomes of an individual are made, was employed with a probability  $p_m$  varying from 0.05 to 0.01 as the optimization proceeds. This choice of GA parameters is the same as discussed in [4]. The evolution was halted after  $N=30$  generations.

The whole computing time  $T_{CPU}$  to find the optimum for a standard optimization procedure is therefore,

$$T_{CPU} = N \times P \times 3 \times T \quad (15)$$

Using the strategy described in Section 3, the preprocessor is called only once at the start of the optimization so  $T_p$  is added just once. Moreover, the computation of matrices **A**, **B** and **M**, occurs only three times for each GA generation. Finally, using the solution of the previous configuration as the initial guess for the iterative conjugate gradient solver (COCG) in the FEM/RBCI analysis, the time  $T_s$  is, on average, reduced by about 7% (about 15% when the new configuration is very similar to the previous one). Hence the whole computing time  $T_{cpu}$  is reduced to,

$$T_{CPU} = T_p + N \times 3 \times T_c + N \times P \times 3 \times 0.93 \times T_s \quad (16)$$

Further computing time is saved when an individual has a null Hamming distance from the previous one: in this case, in fact, the fitness is not recalculated but is simply allotted the same value as its twin.

Implementing all these tricks, the time required to carry out the optimization,  $T_{CPU}$ , is reduced from  $1.35 \cdot 10^5$  s to  $0.6 \cdot 10^5$  s.

The optimum configuration for the patch antenna was that with sub areas  $x_1$ ,  $x_7$ ,  $x_8$ ,  $x_{10}$ ,  $x_{11}$  empty and all the others filled with metal (light gray areas in Fig. 1). The objective function value calculated for the optimal configuration is  $f_{min} = 0.94$ . The return loss of the optimized patch is shown in Fig. 4, as per design; the resonant frequencies occur at 1.23 and 1.57 GHz. The best (minimum) and average values of the objective function  $f$ , through the various generations, are plotted in Fig. 5. The history of GA optimization shows a good convergence by the algorithm.

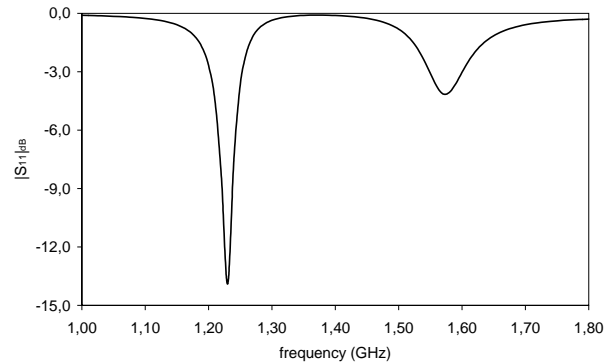


Fig. 4. Return loss of the optimal patch.

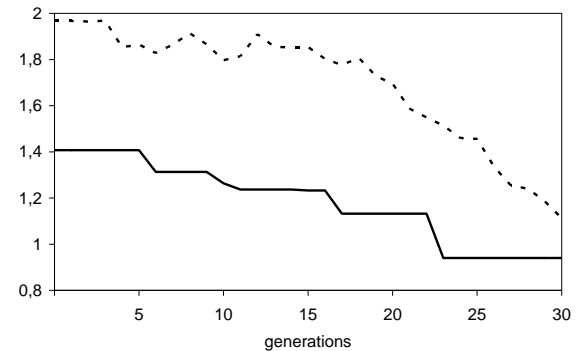


Fig. 5. Best (solid) and average (dotted) objective function  $f$  over GA generations.

## V. CONCLUSIONS

In this paper optimization of a microstrip antenna has been performed by means of Genetic Algorithms and a hybrid Finite Element – Robin Boundary Condition Iteration method. The goal was to design a patch antenna for personal communications that operates at the two GPS frequencies.

A strategy to make the optimization procedure more efficient has been outlined. The optimum was reached in about half the time required by the standard procedure.



The optimized patch performs well in the design frequency bandwidth and has an unconventional and non-intuitive shape.

## REFERENCES

- [1] J. M. Johnson and Y. Rahmat-Samii, "Genetic algorithms and method of moments (GA/MoM) for the design of integrated antennas," *IEEE Trans. Antennas Propagat.*, vol. 47, pp. 1606–1614, Oct. 1999.
- [2] Z. Li, Y. E. Erdemli, J. L. Volakis, and P. Y. Papalambros, "Design Optimization of Conformal Antennas by Integrating Stochastic Algorithms With the Hybrid Finite-Element Method," *IEEE Trans. Antennas Propagat.*, vol. 50, pp. 676–684, May 2002.
- [3] D. E. Goldberg, *Genetic Algorithms in Search, Optimization & Machine Learning*, Addison Wesley, Massachusetts, 1989.
- [4] S. Alfonzetti, E. Dilettoso, N. Salerno, "A Proposal for a Universal Parameter Configuration for Genetic Algorithm Optimization of Electromagnetic Devices," *IEEE Trans. Magn.*, vol. 37, pp. 3208–3211, Sept. 2001.
- [5] S. Alfonzetti and G. Borzi, "Finite Element Solution to Electromagnetic Scattering Problems by means of the Robin Boundary Condition Iteration Method," *IEEE Trans. Antennas Propagat.*, vol. 50, pp. 132–140, Feb. 2002.
- [6] S. Alfonzetti, B. Azzeroni, and G. Borzi, "Numerical Computation of Antenna Parameters by means of RBCI," *Electromagnetics*, vol. 22, pp. 381–392, 2002.
- [7] J. Jin, *The Finite Element Method in Electromagnetics*. New York, John Wiley & Sons Inc, 1993.
- [8] J. A. Kong, *Electromagnetic Wave Theory*. New York, J. Wiley & Sons Inc, 1986.
- [9] S. Y. Wang and K. Tai, "A Constraint Handling Strategy for Bit-Array Representation GA in Structural Topology Optimization," *4<sup>th</sup> SMA Annual Symp.*, Singapore, pp. 19–20, Jan. 2004
- [10] G. Aiello, S. Alfonzetti, G. Borzi, and N. Salerno, "An overview of the ELFIN code for finite element research in electrical engineering," in *Software for Electrical Engineering Analysis and Design IV*, A. Konrad and C. A. Brebbia (editors), WIT Press, Southampton, 1999.
- [11] H. A. van der Vorst and J. Melissen, "A Petrov-Galerkin type method for solving  $Ax=b$  where  $A$  is symmetric complex," *IEEE Trans. Magn.*, vol. 26, pp. 706–708, 1990.
- [12] R. Barrett et al., *Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods*, SIAM, Philadelphia, 1994.



**Salvatore Alfonzetti** was born in Catania, Italy, in 1952. He received his degree in Electrical Engineering (summa cum laude) from the University of Catania in 1976. Since 1977 he has held various posts at the same university and is presently Full Professor of Electrical Engineering. From 1992 to 1998 he was the Director of the Computing Centre of the Engineering Faculty. Since 2005 he has been the Director of the DIEES (Electrical, Electronic and

Systems Department) of the University of Catania. His main research work is concerned with the finite element analysis of electromagnetic systems. He is a member of the editorial boards of a number of international conferences (Compumag, CEFC, Eurosoft). He is the author of more than 160 technical papers. Prof. Alfonzetti is a member of IEEE (The Institute of Electrical and Electronic Engineers) and a founder member of ICS (International Compumag Society).

Prof. Alfonzetti's email address is [alfo@diees.unict.it](mailto:alfo@diees.unict.it)



**Giuseppe Borzi** was born in Catania, Italy, in 1969. He received his degree in Electronic Engineering (summa cum laude) from the University of Catania in 1993. In 1998 he received his PhD degree in Electromagnetism from the same University. Currently he is a University Researcher in Electrical Engineering at the University of Messina, Italy. His research

interests are concerned with the numerical computation of unbounded electromagnetic fields, both at high and low frequencies.

Dr. Borzi's email address is [gborzi@ieee.org](mailto:gborzi@ieee.org).



**Emanuele Dilettoso** was born in Catania, Italy, in 1973. He received his degree in Electronic Engineering from the University of Catania in 1999. In 2003 he received his PhD degree in Electrical Engineering from the same University. Currently he works as a Lecturer in Electrotechnics at the

University of Catania. His research mainly concerns the numerical computation of unbounded electromagnetic fields and the optimization of electromagnetic devices.

Dr. Dilettoso's email address is [edilettoso@diees.unict.it](mailto:edilettoso@diees.unict.it).



**Nunzio Salerno** was born in Catania, Italy, in 1966. He received his degree in Electronic Engineering from the University of Catania in 1992. In 1997 he received his PhD degree in Electrical Engineering from the same University. Since 1998 he is a University Researcher in Electrical Engineering at the University of Catania, Italy. His research interests are concerned

with the numerical computation of unbounded electromagnetic fields and with optimization of electromagnetic devices. He is the author of more than 70 technical papers.

Dr. Salerno's email address is [nsalerno@diees.unict.it](mailto:nsalerno@diees.unict.it).

# A New Broadband Microstrip Leaky – Wave Antenna

O. Losito

ITEL Telecommunications Research Laboratory on EM  
Via Labriola, 39 - 70037 Ruvo di Puglia, Italy  
E-mail: o.losito@itelte.it

**Abstract** – This paper presents a novel design of a curve tapered leaky-wave antenna (LWA). An FDTD code was used to extract the propagation constant of a microstrip LWA and to run a simple algorithm to design its layout. A physical grounding structure along the length of the antenna allows the adoption of a simple feeding planar line and the reduction of sidelobes. Moreover, the proposed design of LWA has interesting performance both for its bandwidth, (up to 33% for VSWR < 2) and for its gain (more than 12 dBi peak power gain at 10.5 GHz), compared with conventional planar microstrip LWAs which work in the same frequency range but with narrower bandwidth (20% VSWR < 2) and peak power gain less than 10 dBi. A prototype of a LWA proposed, was made showing a good agreement between experimental and theoretical results.

**Keywords:** Leaky wave, broadband antennas, FDTD, and tapered antenna.

## I. INTRODUCTION

Substantial enhancements were achieved since pioneering studies [1,2] on microstrip leaky-wave antennas (LWA), and they are now very popular and widely used in applications thanks to their advantages of low-profile, easy matching, narrow beamwidth, fabrication simplicity, and frequency/electrically scanning capability. In some applications the mainbeam variation of LWA should be as low possible. In these cases it is possible to use a tapered microstrip LWA in which each section of the antenna, irradiates in specific ranges of frequency, obtaining therefore a fixed mainbeam. It's equivalent to a broadband antenna.

Unfortunately in these structures the impedance mismatch between the different sections of LWA, and the fundamental mode perturbation, that is a bound mode, reduce the bandwidth of LWA. Slots in the microstrip conductor are possible solutions to eliminate the fundamental mode [1]. Suitable metal walls down the centerline connecting the conductor strip and the ground plane can be considered alternatively, as shown in Fig. 1 determining the possibility to reduce the width of the antenna simplifying its feeding structure [3].

We have studied and designed a broadband tapered LWA, with a simple algorithm, as discussed in the

following sections, showing the experimental results of a LWA prototype made using the proposed design, which is in a good agreement with theoretical results.

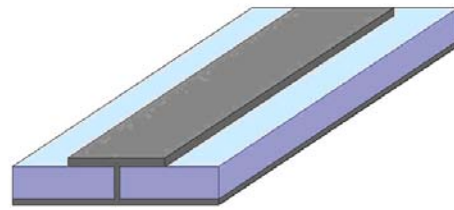


Fig. 1. Travelling half antennas.

## II. CHARACTERISTICS OF MICROSTRIP LWA

The radiation mechanism of higher order modes on microstrip LWA is attributed to a traveling wave instead of the standing wave as in patch antennas and above cutoff frequency, where the phase constant equals the attenuation constant ( $\alpha_c = \beta_c$ ), it is possible to observe three different range of propagation: bound wave, surface wave and leaky wave. While at low frequency, below the cutoff frequency, we have the reactive region due to evanescent property of LWA. We can explain the character of microstrip LWA trough the complex propagation constant  $k = \beta - j\alpha$ , where  $\beta$  is the phase constant of the first higher mode, and  $\alpha$  is the leakage constant of the guided mode.

The main-beam radiation angle of LWA can be approximated by,

$$\theta = \cos^{-1} \left( \frac{\beta}{K_0} \right) \quad (1)$$

where  $\theta$  is the angle measured from the endfire direction and  $K_0$  is the free space wavenumber.

From equation (1) we can observe that the leaky mode leaks away in the form of space wave when  $\beta < K_0$ , therefore we can define the radiation leaky region from the cutoff frequency to the frequency at which the phase constant equals the free-space wavenumber ( $\beta = K_0$ ). For ( $\beta > K_0$ ) we have the bound

mode region and for  $K_0 < \beta < K_s$ , exists a narrow frequency range in which we can have surface-wave leakage.

### II. DESIGN OF BROADBAND LWA

The dispersion equation is the main instrument of analysis to determine the range of propagation of the antenna. Generally the solution of that equation can be obtained through a full-wave analysis such as spectral domain analysis (SDA) [4], or with a transverse resonance approximation according to [5]. Recently an FDTD code which uses a PML boundary condition has been introduced in [3] to extract the propagation constant of a microstrip LWA. Using this code we can obtain in a simple way the radiation leaky region for a different geometry of LWAs. Moreover, the idea to replace slots with a physical grounding structure along the length of the antenna connecting the conductor strip and the ground plate, allow using only half of the structure due to image theory, strongly reducing the computational costs to run a FDTD code.

A transverse electric field  $E_z$  was used in 3D FDTD code, to excite the antenna using a sinusoidal source with a cubic ramp over the first three periods according to [3], located inside the substrate. A PEC was used to modeled ground plane and all conductors, while PML were applied to all other boundaries directly in contact with the antenna to suppress the reflection of travelling wave. A recursive least-squares procedure was used to determine the constants  $\alpha$  and  $\beta$  by matching a known exponential curve to the transverse electric field amplitude, retrieved from 3D FDTD data (see Fig. 2), along the length of the antenna. The symbol  $\alpha$  was found from the peak values, while  $\beta$  was found from the zero crossings, as shown in Fig. 3. The dispersion characteristics curve of Fig. 4, obtained by FDTD code, shows a good agreement with transverse resonance approximation derived by Kuestner [5].

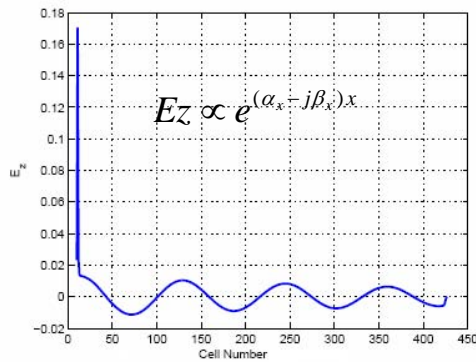


Fig. 2. A transverse electric field  $E_z$  data retrieved by FDTD code.

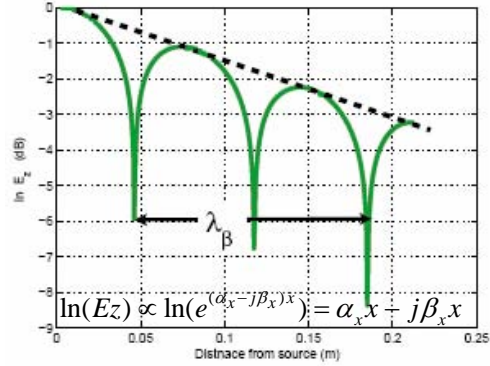


Fig. 3. The logarithmic curve of transverse electric field  $E_z$  data used to determine the propagation constant.

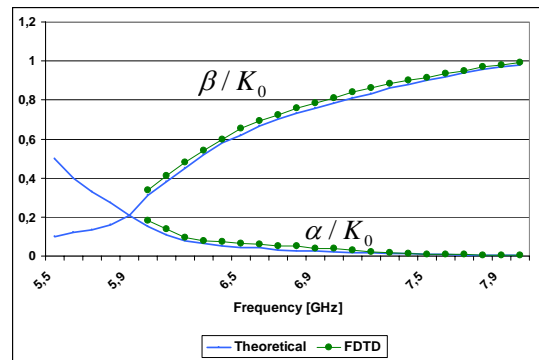


Fig. 4. Theoretical and FDTD dispersion characteristics of leaky wave first high mode.

Moreover, to validate the FDTD code we have calculated and plotted the relationship between the relative permittivity of the substrate and the propagation constant, and the relationship between the thickness of the substrate and the propagation constant. The good agreement of these curves, with the theoretical transverse resonance approximation [5], as shown in Figs. 5 and 6, confirms the validity of the numerical FDTD code used.

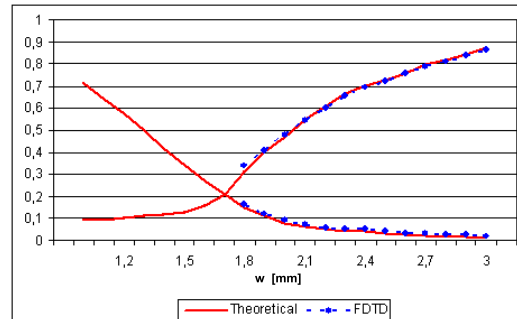


Fig. 5. Theoretical and FDTD dependence between the relative permittivity of the substrate and the propagation constant at  $f = 6.7$  GHz,  $w = 15$  mm,  $h = 0.787$  mm.

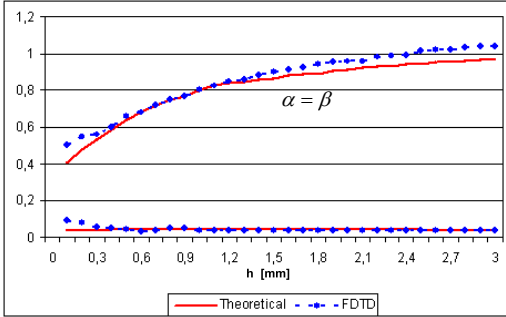


Fig. 6. Theoretical and FDTD dependence between the thickness of the substrate and the propagation constant at  $f = 6.7$  GHz,  $\epsilon_r = 2.32$ ,  $h = 0.787$  mm.

Through the dispersion characteristic equation, evaluated with FDTD code, we can obtain the radiation region of the leaky waves indicated in the more useful way for the design of our antenna,

$$\frac{c}{2w_{\text{eff}}\sqrt{\epsilon_r}} = f_c < f < \frac{f_c\sqrt{\epsilon_r}}{\sqrt{\epsilon_r - 1}}. \quad (2)$$

From equation (2) we can observe that the cutoff frequency increases when the width of the antenna decrease, shift toward high frequencies, the beginning of the radiation region as shown in Fig. 7. Therefore it is possible to design a multisection microstrip antenna [as Type I antenna in Fig. 4], in which each section able to radiate at a desired frequency range, can be superimposed, obtaining an antenna with the bandwidth more than an uniform microstrip antenna. In this way every section should be into bound region, radiation region or reactive region, permitting the power, to uniformly radiated at different frequencies.

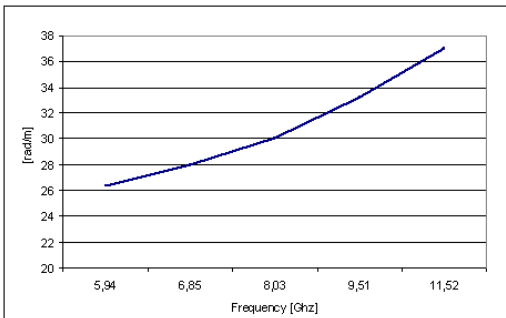


Fig. 7. Cutoff freq. of multisections microstrip LWA .

Using the same start width and substrate of Menzel travelling microstrip antenna (TMA) [1], and total length of 120 mm, we have started the iterative procedure mentioned in [6] to obtain the number, the width and the

length of each microstrip section. From Menzel TMA width, we have calculated the  $f_{\text{START}}$  (onset cutoff frequency) of the curve tapered LWA, than, choosing the survival power ratio ( $\tau = e^{-2\alpha_i L_i}$ ) opportunely, at the end of the first section, we have obtained the length of this section. The cutoff frequency of subsequent section ( $f_i$ ), was determined by FDTD code, while the length of this section was determined, repeating the process described previously. This iterative procedure was repeated, until the upper cutoff frequency of the last microstrip section.

The presence of ripples in return loss curve and the presence of spurious sidelobes shows the impedance mismatch and discontinuity effect of this multisection LWA that reduce the bandwidth. A simple way to reduce these effects is to design a tapered antenna in which the beginning and the ending, respectively, of the first and the last sections are linearly connected together (as the Type II antenna in Fig. 8).

Alternatively, this idea was to design a LWA using a physical grounding structure along the length of the antenna, with the same contour of the cutoff phase constant or attenuation constant curve ( $\alpha_c = \beta_c$ ), obtained varying the frequency, for different width and length of each microstrip section as shown in Fig. 3, employing the following simple equation (3),

$$\beta_c = c_1 f^2 + c_2 f + c_3 \quad (3)$$

obtained from linear polynomials interpolation, where  $c_1 = 0.0016$ ,  $c_2 = 0.03$ ,  $c_3 = -15.56$ .

The antenna layout (as the Type III antenna in Fig. 8), was optimized through a 3D electromagnetic simulator, and the return loss and the radiation pattern was compared with Type I antenna and Type II antenna.

### III. SIMULATION RESULTS

An asymmetrical planar  $50 \Omega$  feeding line was used to excite the first higher-order mode while a metal wall down the centerline connecting the conductor strip and the ground plate was used to suppress the dominant mode for Type I - III. The chosen substrate had a dielectric constant of 2.32 and a thickness of 0.787 mm, while the total length of the leaky wave antenna was chosen to be 120 mm.

The leaky multisection tapered antenna Type I was open-circuited, with a 15 mm start width, and 8.9 mm of final width obtained according to [6]. For LWA layout Type I, we used four microstrip steps, for layout Type II we tapered the steps linearly, while the curve contour of the LWA layout Type III, was designed through equation (3).

Figure 9 shows the simulated return loss of three layouts. We can see that the return loss (S11) of Type I is

below -5 dB from 6 to 10.3 GHz, but only three short-range frequencies are below -10 dB. S11 of Type II is below -5 dB from 6.1 to 9.1 GHz, and below -10 dB from 6.8 to 8.6 GHz. At last, S11 of Type III is below -5 dB from 6.8 to 11.8 GHz, and below -10 dB from 8.0 to 11.2 GHz. In Fig. 10 are shows the mainlobe direction at 9.5 GHz for the different Type I to Type III. We can see a reduction of sidelobe and only few degrees of mainlobe variation between Types I to III. Moreover, in Fig. 11 is shown the variation of mainlobe of antenna Type III, for different frequency, while in Fig. 12 is shown the trend of gain versus frequency of the same antenna. It is clear that, the peak power gain is more than 12 dBi, which is almost 3 dBi higher than uniform LWAs.

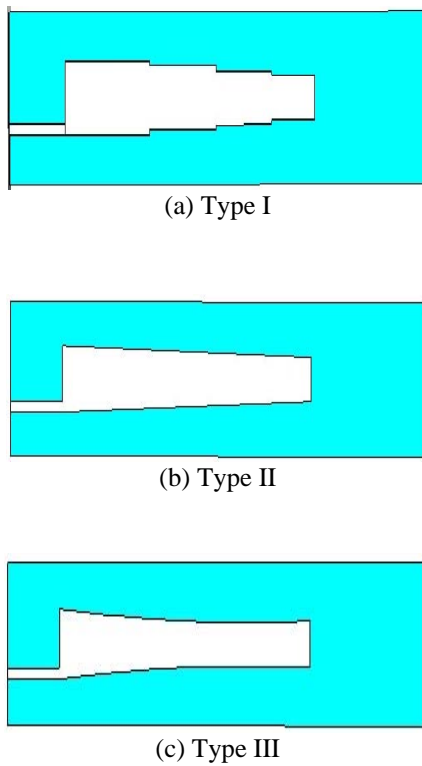


Fig. 8. Layout of leaky wave antennas types I-III. A physical grounding structure was used to connect the conductor strip and the ground plane.

Finally, the simulated VSWR is less than 2 and between 8.01 and 11.17 GHz (33%), yielding an interesting relative bandwidth of 1.39:1, as shown in Fig. 13, compared with uniform microstrip LWAs (20% for VSWR < 2) as mentioned in [7].

These results indicate a high performance of Type III LWA: high efficiency excitation of the leaky mode, increases of the bandwidth, improves the return loss and reduction of 19% of metallic surface with respect to uniform LWA. Moreover, these results are in a good

agreement with the experimental results of return loss and radiation pattern of a prototype (shown in Fig. 14) made using a RT/Duroid 5880 substrate with thickness of 0.787 mm and relative dielectric constant of 2.32, as shown in Figs.15 and 16.

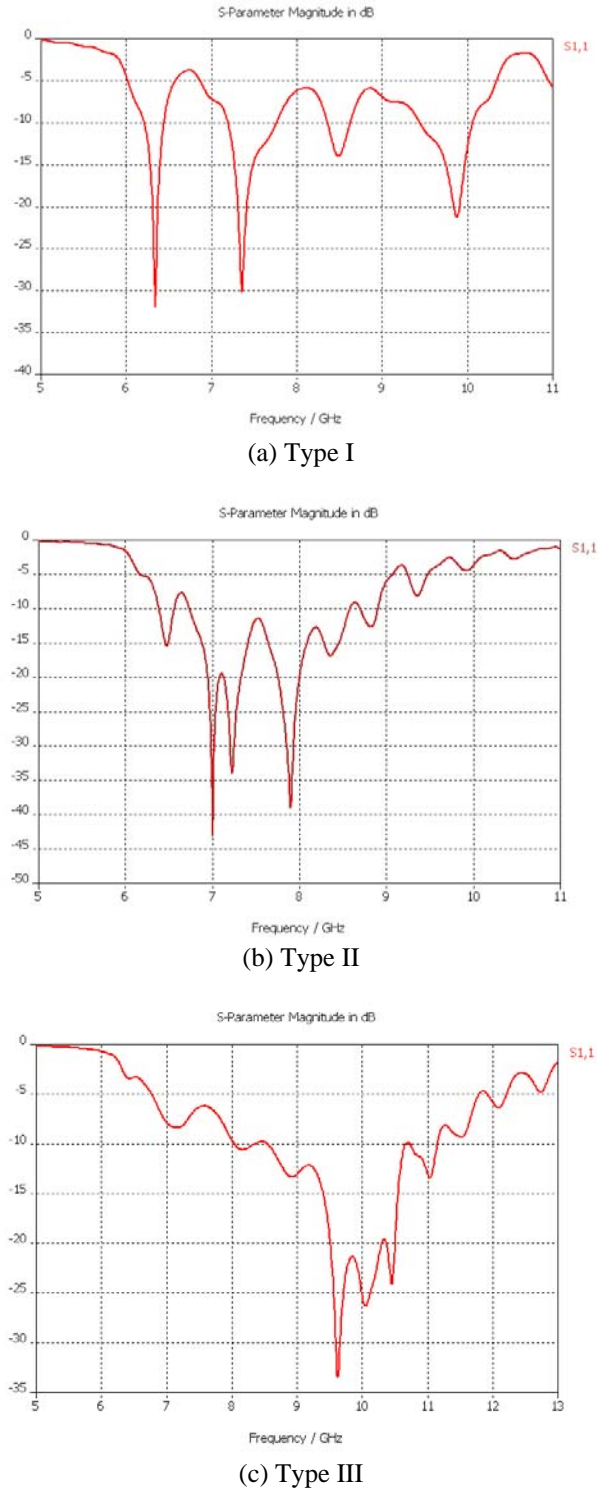
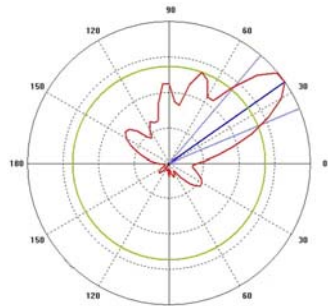
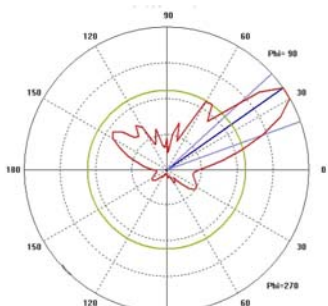


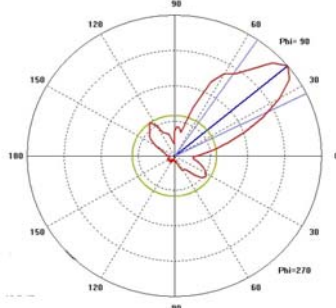
Fig. 9. Simulated return loss of types I-III LWA.



(a) Type I



(b) Type II



(c) Type III

Fig. 10. Radiation patterns of electric field (H plane) of types I-III, LWA at 9.5 GHz.

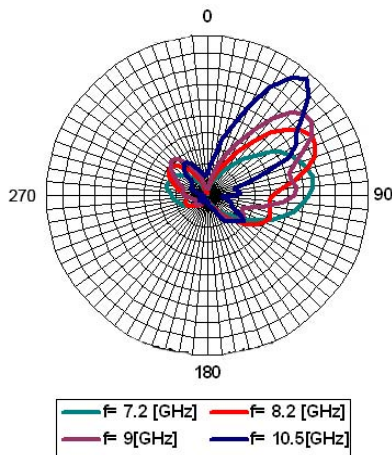


Fig. 11. Simulated radiation patterns of E field of LWA type III for different frequencies.

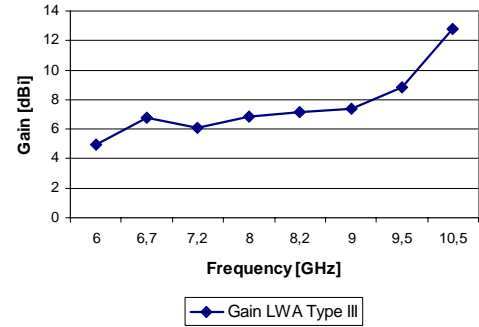


Fig. 12. The gain versus frequency of the LWA type III.

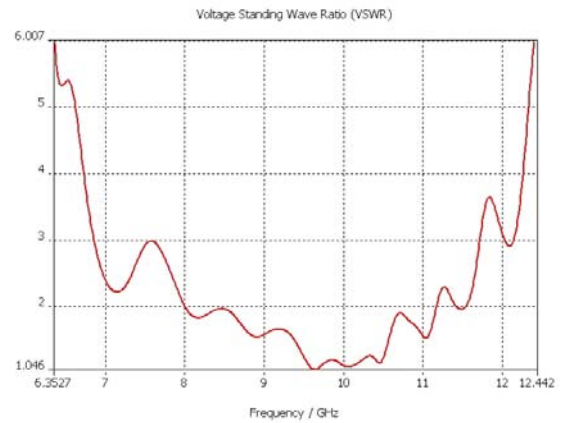


Fig. 13. Simulated radiation patterns of E field of LWA type III for different frequency.

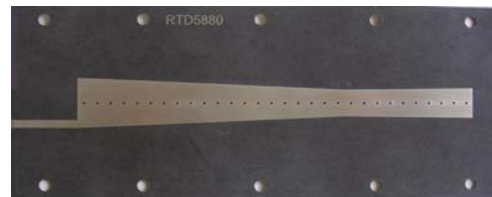


Fig. 14. A prototype of tapered LWA with holes made in the center line of the antenna.

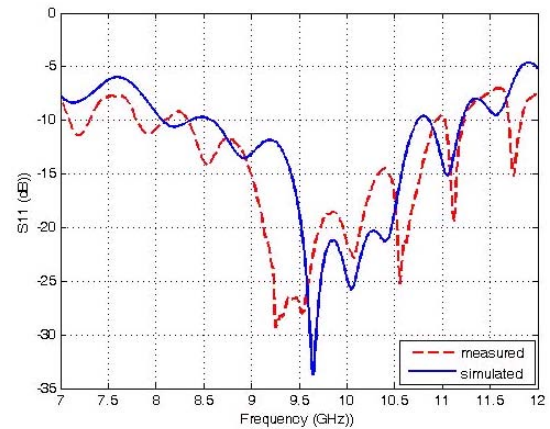


Fig. 15. Experimental and simulated return loss of LWA type III.

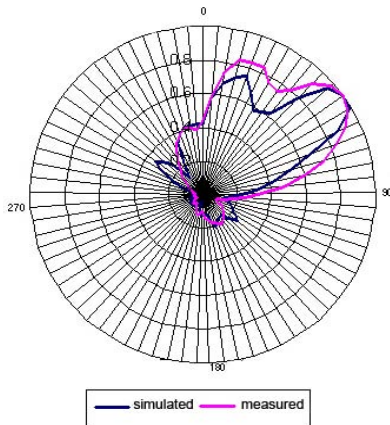


Fig. 16. Measured and simulated radiation patterns of E field of LWA type III (at 8 GHz on the right).

#### IV. CONCLUSIONS

A novel technique which provides broadband tapered microstrip leaky-wave antennas with high added value has been introduced in this study, from 8 to 11 GHz. The propagation constant evaluated with FDTD code was used to design the layout, with the same contour obtained from the interpolation of the cutoff frequency points, for different widths and lengths of each microstrip section, and with a physical grounding structure along the length of the antenna.

The experimental and simulation results shown the good performance of a curve tapered microstrip leaky wave antenna, with reference to conventional uniform microstrip LWAs (wider band and higher gain), and indicate that this structure is attractive for the design of high performance microstrip leaky-wave antennas for microwave and millimeter wave applications.

#### REFERENCES

- [1] W. Menzel, "A new traveling-wave antenna in microstrip," *Arch. Elek. Uebertragung.*, vol. 33, pp. 137-140, 1979.
- [2] A. Oliner, "Leakage from higher modes on microstrip line with application to antennas," *Radio Science*, vol. 22, pp. 907-912, 1987.
- [3] G. M. Zelinski, M. L. Hastriter, M. J. Havrilla, J. S. Radcliffe, A. J. Terzuoli, and G. A. Thiele, "FDTD analysis of a new leaky traveling wave antenna," *Wireless Communications and Applied Computational IEEE/ACES 2005 Intl. Conf.*, pp.152-155, Apr. 2005.
- [4] T. Itoh and R. Mittra, "Spectral domain approach for calculating the dispersion characteristics of microstrip lines," *IEEE Trans. Microwave Theory Techn.*, vol. MTT-21, pp. 496-499, 1973.
- [5] E. F. Kuester, R. T. Johnk, and D. C. Chang, "The thin-substrate approximation for reflection from the end of a slab-loaded parallel-plate waveguide with application to microstrip patch antennas," *IEEE Trans. Antennas and Propagation*, vol. AP-30, no. 5, Sep. 1982.
- [6] W. Hong, T. L. Chen, C. Y. Chang, J. W. Sheen, and Y. D. Lin, "Broadband tapered microstrip leaky-wave antenna," *IEEE Trans. Antennas and Propagation*, vol. 51, pp. 1922-1928, 2003.
- [7] Y. Qian, B. C. C. Chang, T. Itoh, K. C. Chen, and C. K. C. Tzuang, "High efficiency and broadband excitation of leaky mode in microstrip structures," *IEEE MTT-S Microwave Symposium Digest*, vol. 4, pp. 1419-1422, 1999.



**Onofrio Losito** was born in Molfetta – Ba (Italy) on the 21th of August 1969. He took his degree in Electronic Engineering, in 2002 from Polytechnic of Bari, and in the September 2007, a PhD degree in Information Engineering at the University of

Lecce. Since 1999 he has cooperated as an outside expert with the I.P.S.I.A.M. of Molfetta (BA) and Bari - Palese, as regards activities post diploma on the techniques of installation for systems of telecommunications (network of mobile phone) and measures of electromagnetic fields. From 1999 to 2002 he practised the free profession of an engineer as an external adviser for various societies operating in the field of the electromagnetic monitoring sub-contractors of the mobile telephony company: Tim, Vodafone, Wind, H3G. Since 2007 he has joined at the ITEL telecommunication s.r.l., as researcher in the ITEL Laboratory of electromagnetic compatibility, and now he has joined the Electromagnetics Fields Group at the Department of electrotechnics and electronic (DEE) of Polytechnic of Bari, focusing the research activity, to the analysis and design of one dimensional and periodic leaky wave antennas, microwave filters, and electromagnetic shielding for RF.

# Two-Element T-Array for Cross-Polarized Breast Tumor Detection

H. Kanj and M. Popović

Department of Electrical and Computer Engineering  
 McGill University, 3480 University Street  
 Montreal, Quebec Canada H3A 2A7  
 hussam.kanj@mail.mcgill.ca, milica.popovich@mcgill.ca

**Abstract** – This paper reports on a computational study of a 2-element cross-polarized antenna array for breast cancer detection. The “Dark Eyes” antenna is used in a T-arrangement to form the array. The antenna return loss is below -10dB in the range 2.3-10.3 GHz and the cross-polarized mutual coupling is less than -30dB for frequencies over 2.4 GHz. Using the finite-difference time-domain (FDTD) method, an ultra-wideband (UWB) pulse is transmitted in the numerical breast phantom and the co- and cross-polarized back-scatter response is recorded from tumors at different locations. The breast phantom was modeled as a simple layered medium of skin and fat. The dispersive properties of layers were introduced through a Debye model. Simulation results for a spherical and a cylindrical tumor of different orientations are presented and discussed.

**Keywords:** Resistively loaded antennas, microwave imaging, cross-polarized array, breast cancer detection.

## I. INTRODUCTION

Microwave imaging techniques are currently being studied as an effective low-cost screening tool for breast cancer detection systems [1, 2]. Furthermore, polarimetric radar-base imaging is thought to improve the imaging technique. For this application, several antennas have been suggested [3, 4]. However, most of these antennas are either large or nonplanar and thus difficult for use in an antenna array.

In a previous work, we have proposed an ultra-wideband (UWB) compact planar antenna design [5] that is easily fabricated using the standard printed circuit board (PCB) process with embedded resistive technology. The antenna can be manufactured on Rogers Duroid 6010 high frequency substrates [6] laminated with a resistive conductive material (RCM) available from Ohmega Technologies [7]. The key advantage of this antenna is its forward-region radiation pattern. This makes the antenna a perfect candidate for a cross-polarized card-array arrangement as suggested in [8]. In our work presented here, we extend our study [9] of the T-arrangement to be used as a

sub-array unit for radar-based polarimetric breast cancer detection system.

The paper is organized as follows. Section II focuses on the geometry and characterization of T-arrangement for two “Dark Eyes” antennas in a homogeneous lossless medium. In Section III, we present a layered tissue model to assess near-field radiation characteristics of the antenna arrangement of Section II. Section IV offers a detailed study on the tumor detection levels as a function of the tumor shape and orientation (spherical, cylindrical – oriented in the cross- or co-polarized manner) and discusses the reported results. Finally, we make the concluding remarks and comment on our near-future work in Section V.

## II. GEOMETRY AND CHARACTERIZATION OF THE TWO-ANTENNA T-ARRANGEMENT

Figure 1(a) shows the miniaturized “Dark Eyes” antenna reported in [5]. Figure 1(b) illustrates the proposed T-arrangement of two of these antennas in a cross-polarized configuration [8]. As the antenna is aimed to serve for microwave breast cancer detection, in our study, it is immersed in a lossless medium of relative permittivity  $\epsilon_r = 10.2$  which is dielectrically close to that of the fatty breast tissue at the center of the frequency range under investigation [10]. The structure was simulated with SEMCAD [11] (three-dimensional finite-difference time-domain (FDTD) solver). The return loss  $S_{11}$  and mutual coupling  $S_{12}$  results were also verified with HFSS [12] (three-dimensional finite-element (FEM) solver).

We here report on the return loss and mutual coupling of the T-arrangement unit. The structure exhibits asymmetry and should be described by all three S-parameter values –  $S_{11}$ ,  $S_{12}$ , and  $S_{22}$ , where the indices correlate with antenna numeration indicated in Fig. 1(b). Figure 2(a) shows that the Antenna-1 return loss  $S_{11}$  in the 2.3-10.3GHz range and does not exceed -10dB. Nearly identical results (not shown here) are obtained for  $S_{22}$  (Antenna-2 return loss). Another important parameter for array design is the mutual coupling. This is quantified by the  $S_{12}$  parameter graphed in Fig. 2(b). As can be observed, although the antennas were placed with only



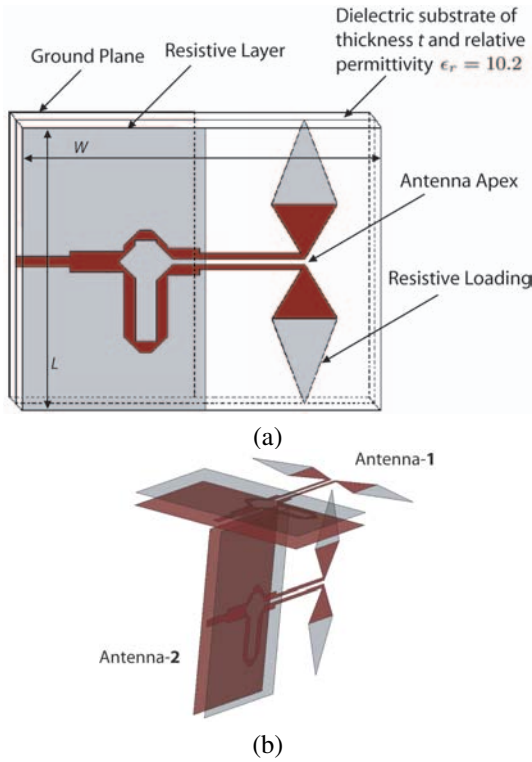


Fig. 1. (a) Geometry of the microstrip-fed resistively loaded "Dark Eyes" antenna.  $L = 20$  mm,  $W = 25$  mm, and  $t = 0.65$  mm. Full description and analysis of the antenna can be found in [5]. (b) Proposed cross-polarized T-arrangement of the microwave sensing array unit, with "Dark Eyes" antenna from Fig. 1(a) as its main element [8].

11mm center-to-center spacing, the  $S_{12}$  is less than -30dB in the 2.4-11GHz range. These results suggest broadband behavior (low return loss and negligible mutual coupling) of the T-arrangement unit within the microwave range of interest for the intended application.

### III. LAYERED TISSUE MODEL AND NEAR-FIELD RADIATION OF THE T-ARRANGEMENT

Figure 3 presents the simple layered tissue model of the breast. It consists of a matching medium, a 1-mm skin layer, and a fat medium. The T-arrangement is immersed in the matching medium which has a dielectric constant of  $\epsilon_r = 10.2$ . The skin, fat, and tumor tissue are modeled with a single-pole Debye dispersive medium defined as follows, [13]

$$\epsilon_r^*(w) = \epsilon_r'(w) - j\epsilon_r''(w) = \epsilon_\infty + \frac{\epsilon_s - \epsilon_\infty}{1 + jw\tau} - j\frac{\sigma_s}{w\epsilon_0} \quad (1)$$

where  $\epsilon_\infty$  is the relative permittivity at infinite frequency,  $\epsilon_s$  is the static relative permittivity,  $\sigma_s$  is the static conductivity, and  $\tau$  is the relaxation time constant. Specific material properties for each tissue are presented in Table 1 [14]. These values are slightly higher than the

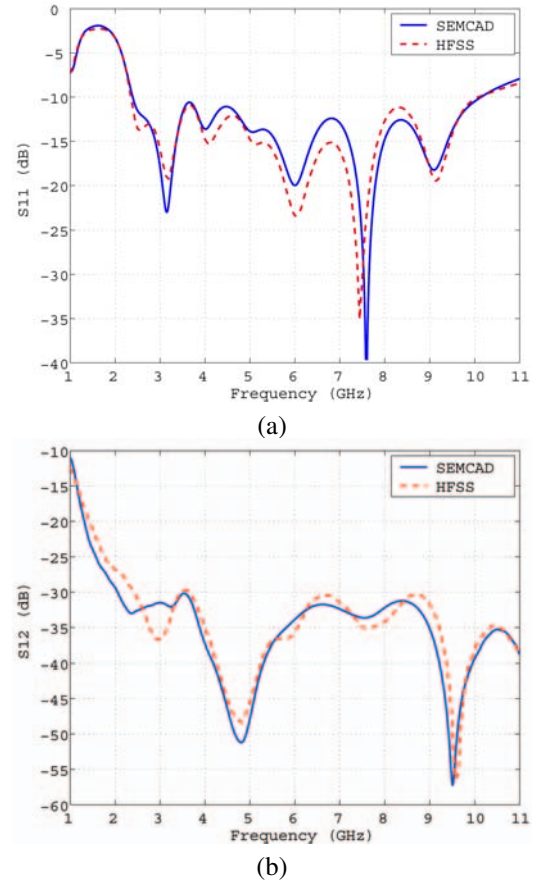


Fig. 2. Simulated S-parameters for the T-arrangement of Fig. 1(b). (a) Input return loss characteristic  $S_{11}$ . (b) Mutual coupling  $S_{12}$ . Results are shown for SEMCAD (FDTD-based) and HFSS (FEM-based) simulation tools for comparison and verification.

measured data for the high percentage adipose tissue recently reported in [15, 16].

Figure 3 shows 9 of the 25 simulated tumor locations alphabetically labeled 'a', 'b', ... 'i'. These locations, in addition to the other 16 not shown in the sketch, are evenly distributed on the portion of the sphere centered at Antenna-1 apex. Therefore, the 25 points find themselves distributed along a "bowl-like" surface beneath Antenna-1. The choice of sampling field points over such a surface (as opposed to, e.g., a plane) was motivated by the fact that we are investigating near-field radiation,

Table 1. Material properties of the Debye dispersive model.

Tissue	Parameters			
	$\epsilon_\infty$	$\epsilon_s$	$\sigma_s$ (S/m)	$\tau$ (ps)
Skin	4.00	37.00	1.10	7.23
Tumor	3.99	54.00	0.70	7.0
Fat	7.00	10.00	0.15	7.0

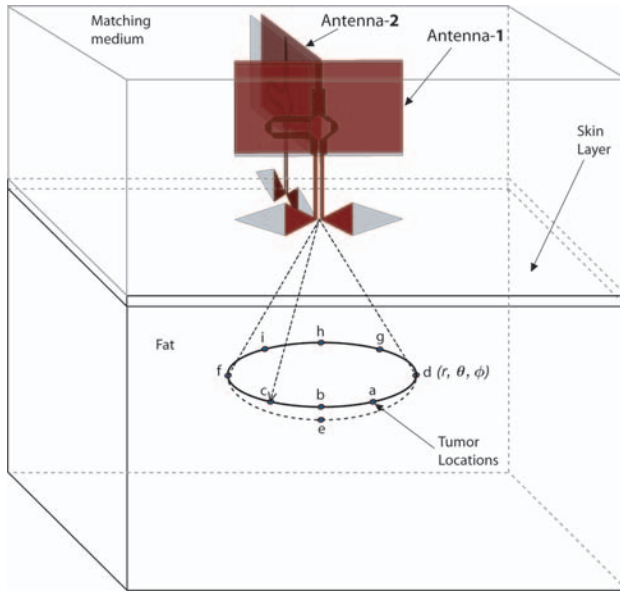


Fig. 3. Layered model of the breast showing the T-arrangement, the matching medium, skin layer, fat, and 9 of the 25 simulated tumor locations.

where the wave has not yet acquired the plane-front of propagation. In all the simulations, the distance between the antenna apex and the skin is 5 mm. The tumor locations are also measured radially from the antenna apex with a radial distance of 3 cm.

Figure 4 shows near-field plots on the portion of the sphere inside the layered model. The plots are computed with HFSS and shown for three frequencies: (a) 3 GHz, (b) 6 GHz, and (c) 9 GHz. In these simulations, both antennas are present, however, only Antenna-1 is active, while Antenna-2 is passive. As we can see from the plots, the maximum intensity of the radiated electric field shifts toward the left side as the frequency increases. However, the range of the field magnitudes at all three frequencies is similar, approximately (2.5-5.5 V/m). This is important to the detection process as the antenna is supposed to radiate the energy uniformly in the intended direction.

#### IV. TUMOR RESPONSE STUDY OF THE T-ARRANGEMENT

In this section, the antenna array is used to study the co-polarized and cross-polarized response of different tumor shapes and orientations at various locations, summarized in Table 2. The antenna array was simulated using SEMCAD X [11]. The tumor was considered to be either a sphere of diameter  $D = 5 \text{ mm}$  or a cylinder with a base diameter  $D = 2.75 \text{ mm}$  and height  $H = 5.5 \text{ mm}$ . The chosen cylinder has its height equal to twice of its diameter ( $H = 2D$ ), and it has the same volume as the chosen sphere. Finally, the cylinder is either oriented in parallel with Antenna-1, when we refer to it as Co-Cylinder, or it is oriented in parallel with Antenna-2, when it is called X-Cylinder. In this study, a tumor is

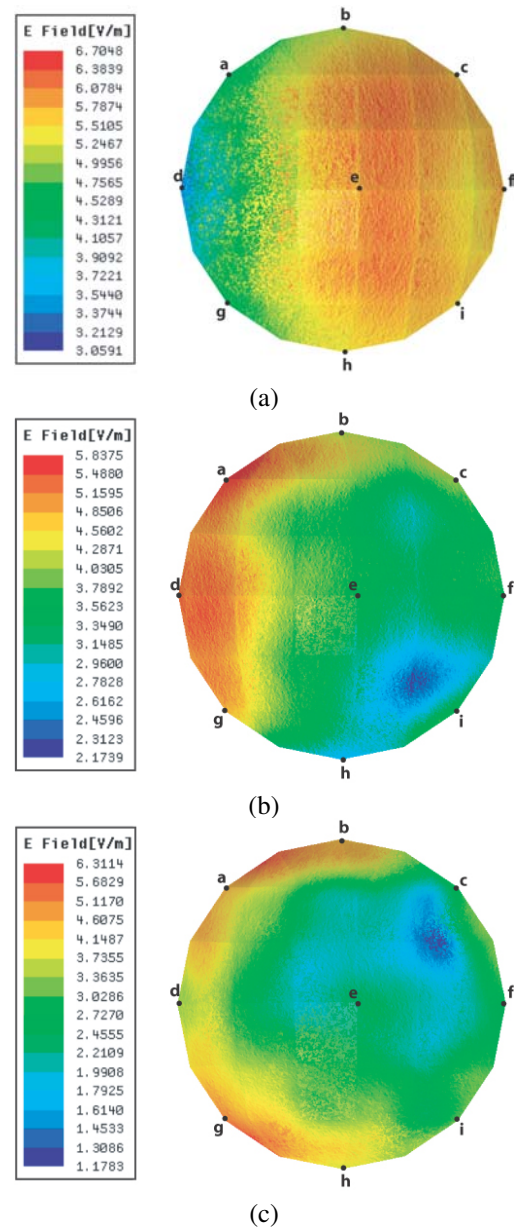


Fig. 4. Electric field magnitude at a radial distance of 30 mm from the antenna apex in the layered model of Fig. 3 at: (a) 3 GHz, (b) 6 GHz, (c) 9 GHz. In the simulations, the parameters of the fat layer containing the points for which the field is plotted are as follows:  $\epsilon_r(3\text{GHz}) = 9.95$  and  $\sigma(3\text{GHz}) = 0.21\text{S/m}$ ,  $\epsilon_r(6\text{GHz}) = 9.8$  and  $\sigma(6\text{GHz}) = 0.4\text{S/m}$ ,  $\epsilon_r(9\text{GHz}) = 9.6$  and  $\sigma(9\text{GHz}) = 0.66\text{S/m}$ .

placed in the layered model at one of the 25 locations evenly distributed on the portion of a sphere shown in Fig. 3, and SEMCAD is used to compute the tumor response. Antenna-1 is excited with a Gaussian modulated sinusoidal pulse described by,

$$\mathcal{V}(t) = \sin[2\pi f_0(t - t_0)] \exp\left[-\frac{(t - t_0)^2}{2\tau^2}\right], \quad (2)$$

with  $f_0 = 6\text{GHz}$ ,  $\tau = 80\text{ps}$ , and  $t_0 = 5\tau$ , while Antenna-2 is kept passive. To compute the tumor response

for each tumor location, two simulations are performed to obtain the voltage at the antenna feed: one with, and one without the tumor. By subtracting one simulated response from the other, skin reflection and early-time artifacts could be removed, providing the response of the tumor only. This is done for both antennas, Antenna-1 and Antenna-2, to compute the co-polarized and cross-polarized tumor response, respectively.

Figure 5 shows the co-polarized and cross-polarized tumor response for the different tumor shapes and orientations considered within the spherical surface defined in Fig. 3. The presented results are the linear interpolation of the computed tumor response at the 25 simulated locations for each case. First, from Fig. 5 (a) and the computed data, we can see that the co-polarized tumor response for the sphere ranges from  $-88.6\text{dB}$  to  $-81.5\text{dB}$  (7.1dB). Our study confirms that the tumor response plot in this specific case correlates with the near-field radiation pattern of the “Dark Eyes” antenna. For the Co-Cylinder case shown in Fig. 5 (c), the co-polarized tumor response ranges from  $-90.2\text{dB}$  to  $-83.2\text{dB}$  (7.0dB), while for the X-Cylinder case shown in Fig. 5 (e), the co-polarized tumor response ranges from  $-94.8\text{dB}$  to  $-86.5\text{dB}$  (8.3dB).

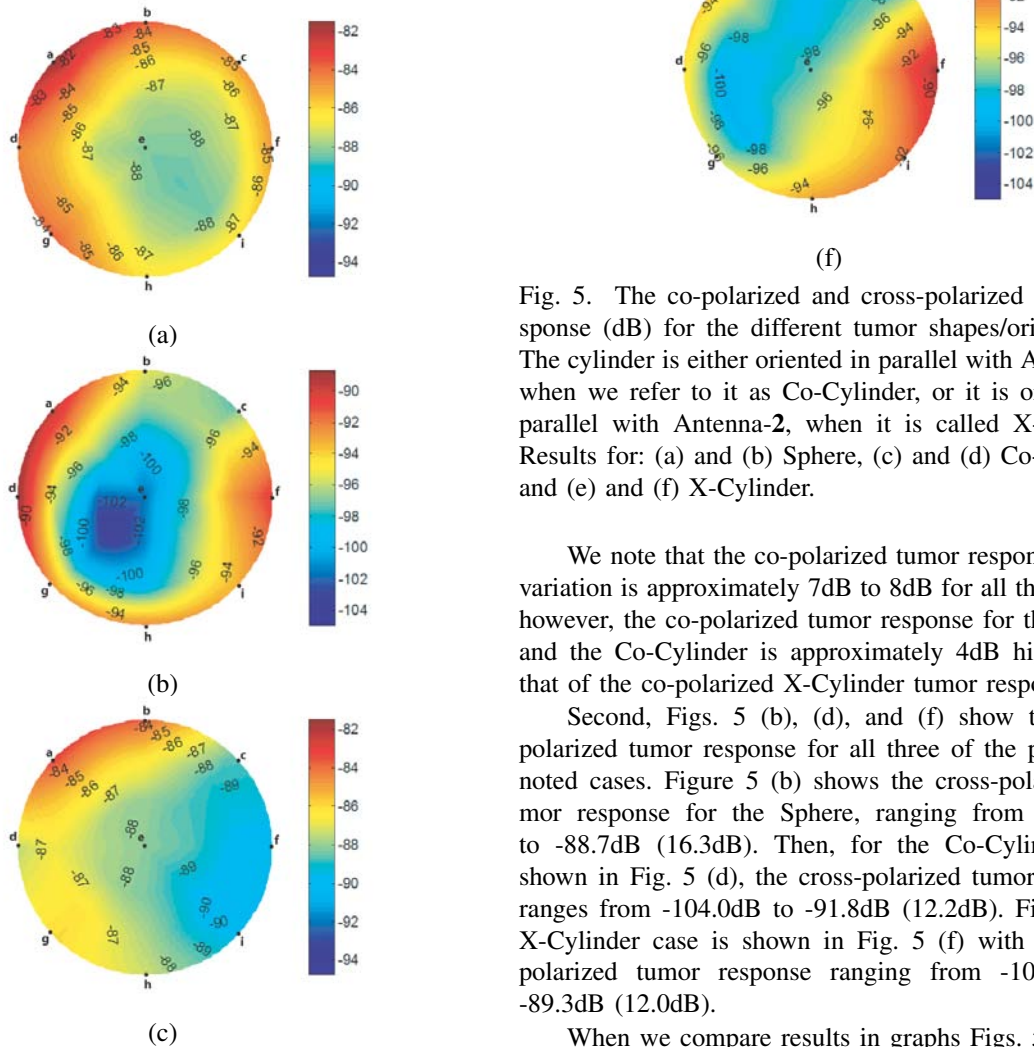


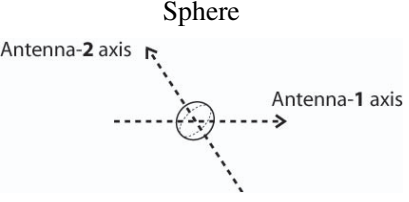
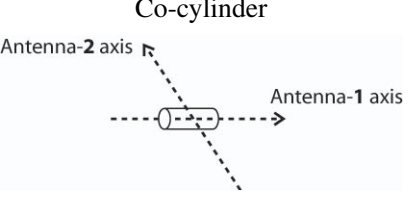
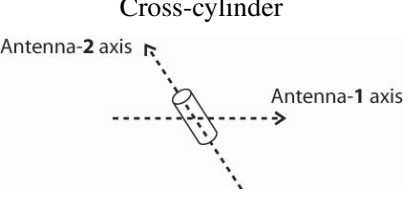
Fig. 5. The co-polarized and cross-polarized tumor response (dB) for the different tumor shapes/orientations. The cylinder is either oriented in parallel with Antenna-1, when we refer to it as Co-Cylinder, or it is oriented in parallel with Antenna-2, when it is called X-Cylinder. Results for: (a) and (b) Sphere, (c) and (d) Co-Cylinder, and (e) and (f) X-Cylinder.

We note that the co-polarized tumor response spatial variation is approximately 7dB to 8dB for all three cases, however, the co-polarized tumor response for the Sphere and the Co-Cylinder is approximately 4dB higher than that of the co-polarized X-Cylinder tumor response.

Second, Figs. 5 (b), (d), and (f) show the cross-polarized tumor response for all three of the previously noted cases. Figure 5 (b) shows the cross-polarized tumor response for the Sphere, ranging from  $-105.0\text{dB}$  to  $-88.7\text{dB}$  (16.3dB). Then, for the Co-Cylinder case shown in Fig. 5 (d), the cross-polarized tumor response ranges from  $-104.0\text{dB}$  to  $-91.8\text{dB}$  (12.2dB). Finally, the X-Cylinder case is shown in Fig. 5 (f) with its cross-polarized tumor response ranging from  $-101.3\text{dB}$  to  $-89.3\text{dB}$  (12.0dB).

When we compare results in graphs Figs. 5 (b), (d),

Table 2. Tumor response summary.

Tumor shape and orientation	Size (mm)	Co-pol response range (dB)	Cross-pol response range (dB)
<p style="text-align: center;">Sphere</p> 	Diameter $D = 5$	$max = -81.5$ $min = -88.6$ $delta = 7.1$	$max = -88.7$ $min = -105.0$ $delta = 16.3$
<p style="text-align: center;">Co-cylinder</p> 	Base Diameter $D = 2.75$ Height $H = 5.5$	$max = -83.2$ $min = -90.2$ $delta = 7.0$	$max = -91.8$ $min = -104.0$ $delta = 12.2$
<p style="text-align: center;">Cross-cylinder</p> 	Base Diameter $D = 2.75$ Height $H = 5.5$	$max = -86.5$ $min = -94.8$ $delta = 8.3$	$max = -89.3$ $min = -101.3$ $delta = 12.0$

and (f) with Figs. 5 (a), (c), and (e), we can note that the co-polarized tumor response is always higher than that of the cross-polarized tumor response for the case of the Sphere and the Co-Cylinder. However, this is not the case for the X-Cylinder where the approximate 5.5dB overlap in the range of the co-polarized tumor response [-94.8dB to -86.5dB] and that of the cross-polarized tumor response [-101.3dB to -89.3dB] implies that having both co- and cross-polarization measurement may be advantageous for the overall detection process. This overlap in the co-polarized and cross-polarized tumor response is mainly attributed to the polarimetric signatures of the cylinder. Table 2 presents a summary of the tumor response study for all three cases under investigation.

## V. CONCLUSIONS

This work presented an array arrangement of the “Dark Eyes” antenna for breast cancer detection. The antennas are arranged in a cross-polarized card array and exhibit low return loss (-10dB from 2.3GHz to 10.3GHz) and low mutual coupling (-30dB from 2.4GHz to 11GHz) even when placed in the very proximity of each other. The near-field radiation patterns at 3GHz, 6GHz, and 9GHz were also presented and the results show that the radiated energy is better directed to the forward region of the antenna at the lower frequency range. Further, a computational study of the tumor response for different tumor shapes and orientations was executed in a simple layered breast model. Both, the co-polarized and the cross-polarized backscatter was recorded from the tumor

placed at different locations in the forward region of the antenna array. The initial results indicate that when looking at the co-polarized response, it is the Sphere and the Co-Cylinder that provided higher-amplitude response. On the other hand, when looking at the cross-polarized response, it is the X-Cylinder that provided the higher-amplitude response. Thus, for extracting maximum information about the tumor, and since the tumor shape and orientation are unknown, there is an advantage in using the cross-polarized array arrangement.

## ACKNOWLEDGMENTS

The authors would like to thank Schmid & Partner Engineering AG for providing the SEMCAD package and Ansoft Co. for their help with the HFSS software. We are also grateful to Natural Science and Engineering Research Council (NSERC) of Canada and Le Fonds Québécois de la Recherche sur la Nature et les Technologies (FQRNT) for their funding support.

## REFERENCES

- [1] Y. Xie, B. Guo, L. Xu, J. Li, and P. Stoica, “Multistatic adaptive microwave imaging for early breast cancer detection,” *IEEE Transactions on Biomedical Engineering*, vol. 53, pp. 1647–1657, August 2006.
- [2] D. J. Kurrant, E. C. Fear, and D. T. Westwick, “Tumor estimation in tissue sensing adaptive radar (tsar) signals,” in *Canadian Conference on Electrical and Computer Engineering*, 2007.

- [3] S. C. Hagness, A. Taflove, and J. E. Bridges, "Three-dimensional FDTD analysis of a pulsed microwave confocal system for breast cancer detection: Design of an antenna-array element," *IEEE Trans. Antennas Propagat.*, vol. 47, pp. 783–791, May 1999.
- [4] X. Yun, E. C. Fear, and R. H. Johnston, "Compact antenna for radar-based breast cancer detection," *IEEE Trans. Antennas Propagat.*, vol. 53, pp. 2374–2380, August 2005.
- [5] H. Kanj and M. Popovic, "Miniaturized microstrip-fed "Dark Eyes" antenna for near-field microwave sensing," *IEEE Antennas Wireless Propagat. Lett.*, vol. 4, pp. 397–401, 2005.
- [6] (2007) Rogers website. [Online]. Available: <http://www.rogerscorporation.com>
- [7] (2007) Ohmega website. [Online]. Available: <http://www.ohmega.com/>
- [8] H. Kanj and M. Popovic, "T- and X-arrangement of "Dark Eyes" antennas for microwave sensing array," in *Proc. IEEE AP-S International Symposium*, Albuquerque, NM, pp. 1111–1114, July 2006.
- [9] —, "Two-element T-array for cross-polarized breast tumor detection," in *The 23rd Annual Review of Progress in Applied Computational Electromagnetics*, Verona, Italy, pp. 846–850, March 2007.
- [10] S. Gabriel, R. W. Lau, and C. Gabriel, "The dielectric properties of biological tissues-ii: Measurements on the frequency range 10 Hz to 20 GHz," *Physics in Medicine and Biology*, vol. 41, pp. 2251–2269, November 1996.
- [11] (2007) SEMCAD website. [Online]. Available: <http://www.semcad.com/>
- [12] (2007) Ansoft website. [Online]. Available: <http://www.ansoft.com/>
- [13] A. Taflove and S. Hagness, Eds., *Computational Electrodynamics: The Finite-Difference Time-Domain Method*. Boston, MA: Artech House, 2005.
- [14] M. Converse, E. J. Bond, B. D. Van Veen, and S. C. Hagness, "A computational study of ultrawideband versus narrowband microwave hyperthermia for breast cancer treatment," *IEEE Trans. Microwave Theory Tech.*, vol. 54, pp. 2169–2180, May 2006.
- [15] M. Lazebnik, L. McCartney, D. Popovic, C. B. Watkins, M. J. Lindstrom, J. Harter, S. Sewall, A. Magliocco, J. H. Booske, M. Okoniewski, and S. C. Hagness, "A large-scale study of the ultrawideband microwave dielectric properties of normal breast tissue obtained from reduction surgeries," *Physics in Medicine and Biology*, vol. 52, pp. 2637–2656, 2007.
- [16] M. Lazebnik, D. Popovic, L. McCartney, C. B. Watkins, M. J. Lindstrom, J. Harter, S. Sewall, T. Ogilvie, A. Magliocco, T. M. Breslin, W. Temple, D. Mew, J. H. Booske, M. Okoniewski, and S. C. Hagness, "A large-scale study of the ultrawideband microwave dielectric properties of normal, benign

and malignant breast tissues obtained from cancer surgeries," *Physics in Medicine and Biology*, vol. 52, pp. 6093–6115, 2007.

**Houssam Kanj** received his B.Sc. in Computer and Communication Engineering from The American University of Beirut in 1999, the M.Sc. in Electrical Engineering from North Carolina State University in 2003, and the Ph.D. degree in Electrical Engineering from McGill University in 2008. His Masters research was in the fields of analog circuit design and computer-aided modeling of nonlinear circuits including electro-thermal and lasers. His current research interests include numerical methods in electromagnetics for bio-medical applications and antenna design.

**Milica Popović** received her B.Sc. (1994) from University of Colorado (Boulder, Colorado, USA) and M.Sc. (1997) and Ph.D. (2001) degrees from Northwestern University (Evanston, Illinois, USA), all in electrical engineering. She is currently an associate professor with the Department of Electrical and Computer Engineering at McGill University in Montreal, Canada. Her research interests focus on numerical methods in computational electromagnetics for bio-medical applications, in particular: breast cancer screening with microwaves, wireless implants for physiological research and light interaction with retinal photoreceptor cells. On the teaching side, her efforts include improvement methods for instruction of introductory electromagnetics courses.

# Finite Array Analysis Through Combination of Macro Basis Functions and Array Scanning Methods

C. Craeye and R. Sarkis

Université catholique de Louvain, Laboratoire TELE  
Place du Levant, 2, 1348 Louvain-la-Neuve, Belgium  
*Email* : craeye@tele.ucl.ac.be

**Abstract** – The Macro Basis Functions (MBF) approach receives increasing attention for the evaluation of the effects of array truncation. In this paper, we show how physically based MBFs can be obtained from solutions for infinite arrays and for arrays of minimal size. The method is first explained and illustrated for the case of finite-by-infinite arrays of electrically connected tapered-slot antennas. It is then extended to the case of planar arrays. Very low error levels are achieved with a small number of MBFs, in terms of port currents as well as in terms of radiation patterns.

## I. INTRODUCTION

In the last few years, very efficient techniques appeared in the literature for the Method-of-Moments (MoM) analysis of large finite periodic antenna arrays. Some methods rely on the infinite-array solution, with corrections for the effects of array truncation [1–3]. Those methods are generally valid for very large arrays and entail several approximations. Other approaches involve fast iterative methods, in which matrix-vector products are accelerated with the help of multipole decompositions or Fast Fourier Transforms, combined with efficient preconditioners [4]. In [5], the FFT approach is advantageously combined with the concept of subentire-domain basis functions, which consists of assuming that the currents on a given antenna in the finite array can be decomposed in terms of a limited number of known current distributions, obtained through the solution of smaller problems. This underlying idea has been found in many publications, where the “macro basis functions” [6, 7] are also called “characteristic basis functions” [8]. Among recent works on this subject, we should note [9] and [10], where primary and secondary distributions are considered, in order to accurately catch the effects of mutual coupling. A very fast implementation of this method has been described in [11], where the interactions between macro basis and macro testing functions are computed with the help of a multipole approach. This method allows the efficient computation of all coupling coefficients of the finite array. Besides this, the patterns of macro basis functions, which are side-products of this method, are then used to rapidly compute the embedded element patterns and, in turn, the array pattern for any excitation law. A similar approach,

involving an iterative scheme, is presented in [12].

At first glance, finite and infinite-array approaches seem very distinct. For large arrays, an initial design can be obtained from the infinite-array simulations, potentially with approximate corrections for truncation, while the behavior of the finite array is verified with the help of an efficient iterative or multiscale approach [13]. The present paper consists of making one more step toward the reconciliation of infinite-array and finite-array approaches. A first set of macro basis functions is obtained from the infinite-array analysis in a very specific and physically-based way, which finds its justification in the Array Scanning Method [14]. A second set of MBFs is obtained from the solution of very small ( $2 \times 1$  or  $2 \times 2$ ) arrays. In [15], MBFs (named “standard distributions”) were also obtained from infinite array solutions, but this method was limited to finite-by-infinite arrays and, more importantly, to periodic excitation, thus requiring only one infinite-array solution (besides single-element and semi-infinite array solutions). The combination of ASM and MBF approaches, for the solution of the array excited at a single port, was first shown in [16], where only finite-by-infinite arrays are considered, while full-wave treatment of edge elements was needed. In the present paper, the method shown in [16] is extended to planar arrays, while special MBFs are included to represent currents on elements on the array periphery.

This paper is organized as follows. In Section 2, the phenomenology of current waves in infinite arrays excited at one element is recalled, and its quantitative evaluation is obtained with the help of the Array Scanning Method in Section 3. In Section 4, the method for extraction of edge MBFs is explained and the numerical examples are given in Section 5 for finite-by-infinite arrays of tapered-slot antennas. The method is extended to arrays finite in both directions in Section 6, where examples are shown in terms of port currents and radiation patterns. The method is summarized and discussed in Section 7.

## II. CURRENT DISTRIBUTIONS IN INFINITE ARRAYS

To properly assess the efficiency of the method explained below, it is interesting to consider worst-case situations in terms of coupling between elements. This is

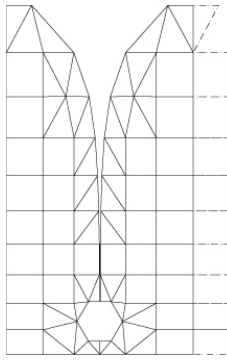


Fig. 1. Discretization of tapered-slot antenna with overlapping basis functions (dashed).

why arrays of electrically connected tapered-slot antennas [15, 17], will be taken as examples (see Figs. 1 and 2 for meshes of the antennas and of the array): first, near the lowest frequency end, the wavelength can be large compared to the array spacing, which often leads to stronger couplings and, second, the electrical connection between elements also leads to very strong couplings and to a quite different current distribution on elements located on the periphery of the array. A first possibility for determining the macro basis functions consists of considering a large spectrum of plane-wave excitations of the antenna, as well as source excitations. In practice, when successive elements are electrically connected, the currents can flow from one element to the next, with currents essentially concentrated on the edges of conductors. If a free-standing element is cut out of the array and excited, currents will also flow along the cutting lines. This will not lead to an acceptable representation of current distributions in array conditions. Hence, some authors [18] proposed to extend the subdomains to two or three antennas and to carefully avoid singular currents by a spatial windowing to the obtained current distributions. Despite the very good results obtained for certain configurations [18], it is difficult to say in how far the macro basis functions generated in this way form a complete set of solutions. Here, “complete” means that any current distribution that can appear on the array can also be described as a linear combination of the proposed set of macro basis functions.

### III. ARRAY SCANNING METHOD AND MBFS

While attempting to answer the latter question, it is useful to come back to the physical interpretation of fields in infinite and finite arrays. Let us assume a finite array excited at one given element. The currents excited on the whole structure can be regarded as those present in an infinite array, plus currents reflected (or “diffracted”) by the edges of the array. Let us first consider the case of excitation at one port in an infinite array. In this case, the currents can be obtained from infinite-array simulations, with the help of the Array Scanning Method (ASM) [14,

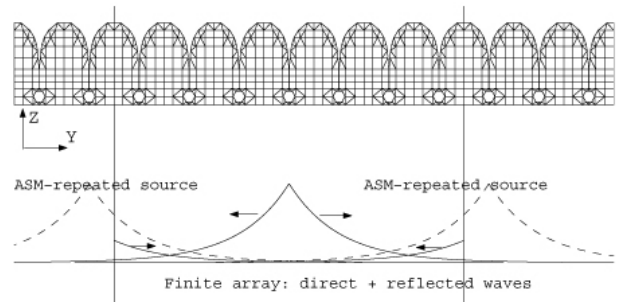


Fig. 2. Wave phenomenology in a finite-by-infinite array excited along one infinite row. The currents on a given point of successive antennas can be regarded as progressive waves launched by the excited element and reflected by the ends of the array (vertical lines). If the Array Scanning Method is implemented with the help of a finite summation, the source is repeated (cf. auxiliary peaks).

19, 20]. The method is recalled in [21] for the case of planar arrays, where examples are provided for wideband phased arrays. In this section, it will be illustrated for a simpler problem, involving arrays infinite along  $\hat{x}$ , and with  $M$  elements along  $\hat{y}$ . The complexity of the method is then reduced to that of a linear array (see Fig. 2). If element 0 (or row 0) is excited, the current on element  $m$  (or row  $m$ ) reads,

$$I(m) = \frac{1}{2\pi} \int_0^{2\pi} I^\infty(\psi) e^{-j m \psi} d\psi \quad (1)$$

where  $I^\infty(\psi)$  is the infinite-array current distribution obtained with inter-element phasing  $\psi$  along  $\hat{y}$ . As for the phasing along  $\hat{x}$ , the infinite-array direction, it is assumed identical for infinite-array and finite-by-infinite array solutions.

The Array Scanning Method assumes the integration of current distributions for all possible phase shifts, from 0 to  $2\pi$ , between successive elements. In practice, solutions will be computed for a finite set of phase shifts. The simplest approximation, computing integral of equation (1) with rectangles in the reciprocal (phase-shift) domain, then comes down to a DFT approach,

$$I(m) \simeq \frac{1}{N} \sum_{p=0}^{N-1} I^\infty(\psi_p) e^{-j m \psi_p} \quad (2)$$

with  $\psi_p = 2\pi p/N$ .

The effect of this approximation is that excited currents are obtained as if the source were repeated every  $N$  elements (see dashed lines for port currents in Fig. 2 for  $N = 4$ ). This poses a problem when the exact solution for a single excitation is looked for, but, as will be seen further, it is not really a drawback when we are just looking for characteristic current distributions, or “macro basis functions”. Indeed, in an infinite array,

and far from the source, the currents along successive antennas may be regarded as a “wave” with a given decay rate in terms of amplitude and with a given phase velocity [22]. Far enough from the source, say beyond  $N_o$  elements away from it, the successive current distributions may -within a constant factor- be very similar, while they can differ quite a lot near the source. Hence, the repetition of excitation every  $N_o$  antennas, as implicitly incurred by the use of equation (2), is not a problem when the number of linearly independent current distributions available remains sufficient, i.e., as long as  $N$  is not too small compared to  $N_o$ . Moreover, the repetition of the sources does not really “spoil” the generated current distributions, since the added current component is very similar to the characteristic current distributions looked for.

Besides this, when the array is finite, the difference with the solution considered above may be regarded as currents reflected by the edges of the array. This has been illustrated in [22] for linear arrays of broadband dipoles. In general, except for the elements right at the edges of the array (which will be treated below), these “reflected” current distributions are very similar to those found in an infinite array excited by a single source. This supports the use of the currents obtained from the ASM analysis as macro-basis functions for the finite array. It is also interesting to notice that, as a result of the physical ground for the choice of current distributions, the distributions obtained in this way will naturally exhibit continuous currents at the boundaries between electrically connected elements.

The explanation above provides a physically-based choice of macro basis functions for elements inside the array (i.e., for all elements, except those located right at the edges of the array). This means that, considering the excitation of any element in the array, the currents on a given antenna can be written as a linear combination of the current distributions obtained with the ASM method. In turn, through superposition of excitation at individual elements of the array, the ASM-based method should provide accurate results for any active excitation law of the array. Furthermore, since the ASM results are obtained as linear combinations of infinite-array results, we can also say that the currents for any excitation law should be a linear combination of currents obtained in the infinite-array case. Hence, the set of MBFs is simply made up of all  $I^\infty(\psi_p)$  current distributions. The only constraint, for the physical justification above to hold, is that the infinite-array solutions be obtained by sampling regularly in the reciprocal (i.e., phase-shift  $\psi$ ) domain, i.e.,  $\psi_p = 2\pi p/N$ , with  $p$  integer between 0 and  $N - 1$ . This includes solutions outside the visible space, for which the antenna active impedances are purely reactive.

#### IV. EDGE MBFS

As explained previously, the wave phenomenology described above may not hold for the elements on the

edges of the array. This is particularly true for arrays made of connected elements, since, in that case, the edge elements may support significantly different types of current distributions. Those are not well captured through the ASM procedure. In the following, they will be obtained by solving very small arrays, consisting of edge elements only. In the finite-by-infinite array example, such an array contains only two electrically connected elements; or more precisely, a two-by-infinite array. Current distributions may be different on left and on right edge elements. These elements may be fed directly, or they may be illuminated through the feeding of another element in the array. In the two-elements case, this leads to four possible current distributions. This requires the full-wave solution of a two-elements array, which, compared to other steps in the computation procedure, takes a relatively small computation time. The resulting four MBFs are simply added to the set of  $N$  distributions obtained from the ASM procedure. As already considered in [23], MBFs are then orthogonalized through the SVD (Singular Value Decomposition) procedure, in order to preserve a good conditioning for the reduced system of equations.

#### V. FINITE-BY-INFINITE ARRAY EXAMPLE

Simulation examples will be shown for an array made of metallic tapered slot antennas, with the discretization shown on Fig. 1. The surface is meshed with the help of 132 elementary basis functions, which are of rooftop and RWG (“Rao-Wilton-Glisson” [24]) types, with half basis functions electrically connecting the antenna to the infinite ground plane and overlapping basis functions (dashed) connecting antennas with each other in the E-plane. Based on comparisons with results obtained with finer meshes, the meshing used here has been found to provide a satisfactory representation of the main antenna characteristics, while still allowing the brute-force solution of intermediate-size finite arrays. For instance, with this discretization, as well as with finer ones, the standing wave ratio for the array scanned at broadside is below 2 from 0.42 GHz to 1.5 GHz.

The array analyzed here is infinite along  $\hat{x}$  (perpendicular to the figure), without a phase shift between elements in that direction, and has 32 elements along  $\hat{y}$  (from left to right). The element spacing is 12.7 cm, while the wavelength is 30 cm. The elements are excited successively (excitation of successive infinite rows) and the currents are compared with the “brute-force solution”, obtained through inversion (or LU decomposition) of the MoM impedance matrix. Results are shown in Fig. 3 for excitation at port 1 and at port 16. The upper plot shows the port currents, while the lower plot shows the error w.r.t. the brute-force solution, with a scale enlarged by a factor of 6000. It can be seen that the residual error is extremely small.

The remaining question is how many MBFs are necessary to obtain a sufficient accuracy; in other terms,



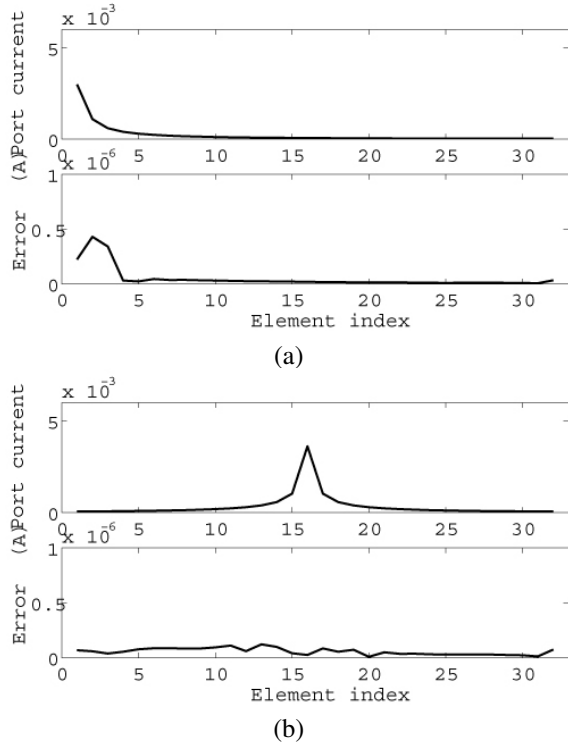


Fig. 3. Port currents in array of tapered slot antennas, with 32 elements in one direction and infinite in orthogonal direction, excited along one infinite row. Array spacing: 12.7 cm, wavelength: 30 cm. Top: port currents. Bottom: with scale enlarged by a factor 6000, errors between full-wave solution and MBF approach (a) first row of antennas excited, (b) row 16 excited.

how large should  $N$  be in the Array Scanning Method? In the case of weakly coupled arrays, for instance when elements are located far apart, it is expected that current distributions on antennas away from the excited element should be very similar. In that case,  $N$  can be very small. For strongly coupled arrays, and a fortiori, for electrically connected elements,  $N$  can become larger, because of the large variety in possible current distributions. The maximum port error, normalized relative to current at the excited port, is represented in Fig. 4 for increasing values of  $N$ . It can be seen that the error suddenly drops to very low values for  $N > 7$ . If the SVD procedure is not used, the error increases again after  $N = 10$  (crosses), because an ill-conditioned system of equations is obtained through the use of very similar basis functions. It can be seen that the use of the SVD procedure (circles) avoids this difficulty.

## VI. EXTENSION TO PLANAR ARRAYS

A similar procedure can be used for planar arrays, like finite arrays of tapered-slot antennas (Fig. 5). In this case, the 2-D version of the ASM is exploited. The current on antenna  $(m, n)$  for excitation of antennas with indices

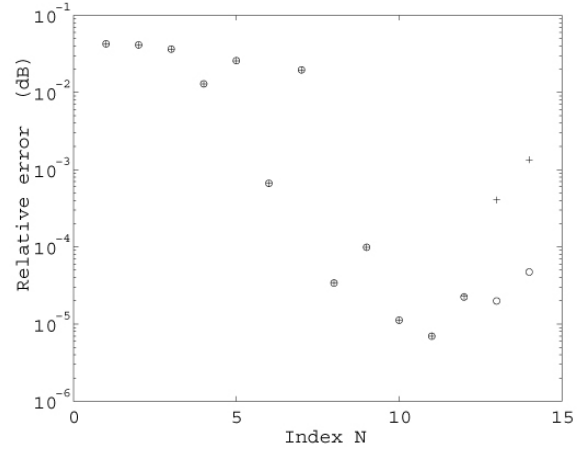


Fig. 4. For 32-by-infinite array, with row 16 excited, maximum relative error in port currents, versus value of  $N$  in ASM. Circles: with SVD procedure. Crosses: without SVD procedure.

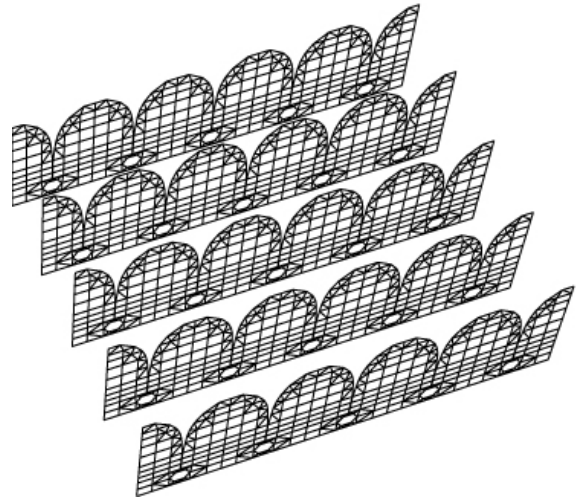


Fig. 5.  $5 \times 5$  array of tapered-slot antennas.

$(rN, sN)$ , with  $r$  and  $s$  integer, is given by,

$$I_{m,n} = \frac{1}{N^2} \sum_{p=1}^N \sum_{q=1}^N I^\infty(\psi_{x,p}, \psi_{y,q}) e^{-jm\psi_{x,p}} e^{-jn\psi_{y,q}} \quad (3)$$

with

$$\psi_{x,p} = 2\pi p/N \quad \text{and} \quad \psi_{y,q} = 2\pi q/N. \quad (4)$$

In equation (3),  $I^\infty(\psi_x, \psi_y)$  is the infinite-array current obtained with inter-element phasings along  $X$  and  $Y$  equal to  $\psi_x$  and  $\psi_y$ , respectively.

As in the finite-by-infinite array case, if  $N$  is large enough, this procedure provides a sufficiently complete set of MBFs for large arrays, except for elements on the outer edge, which may support significantly different current distributions. This time, the problem is solved by adding to the set of MBFs a few current distributions

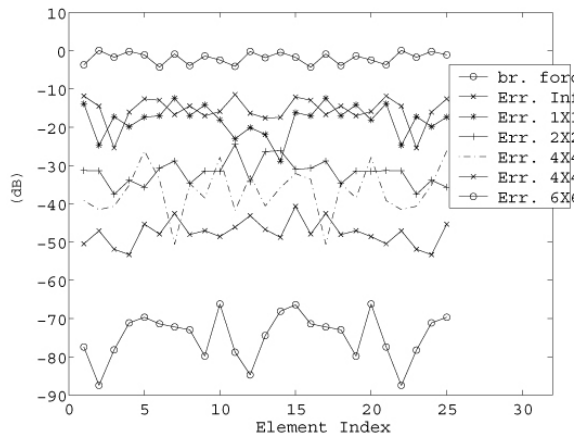
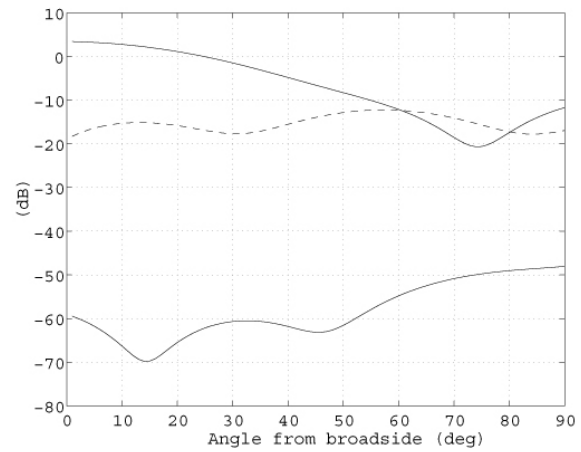


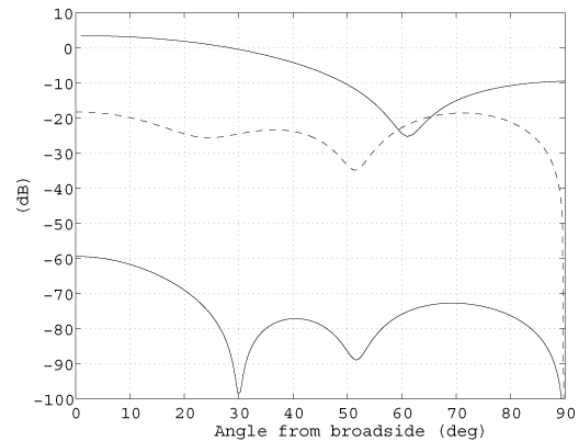
Fig. 6. Port currents for all elements of  $5 \times 5$  array excited (upper line) and errors produced by the infinite-array approximation (crosses around -15 dB) and by ASM+MBF with different values of  $N$  for the two-dimensional ASM and with 8 additional MBFs from  $2 \times 2$  array. The  $4 \times 4$  case denoted with dashed lines does not include the 8 additional MBFs.

obtained through the full-wave analysis of a small array, e.g. a  $2 \times 2$  array, which contains only edge elements. Four elements and four independent excitations lead to 16 possible MBFs. However, if some symmetry is present in the array, fewer distributions need to be considered. In the present example with tapered-slot antennas, since two parallel columns are perfectly symmetrical, this leads to 8 different current distributions. Adding these distributions to those stemming from the ASM procedure with  $N = 4$ , for example, leads to 24 MBFs, instead of the original 132 elementary basis functions per antenna. This reduction by a factor 5.5 of the number of unknowns leads to a reduction by a factor 166 in terms of direct solving time. Larger time saving factors are expected for more complex antennas.

Figure 6 shows results obtained for port currents in a  $5 \times 5$  array of electrically connected tapered-slot antennas (Fig. 5). Overlapping basis functions ensure the electrical connection between antennas. It should be noted that, for simplicity, these overlapping basis functions have been kept on the last elements in each row. In the first numerical example, all antennas are excited with a uniform voltage excitation, with  $100 \Omega$  series impedances. The 25 port currents are represented by the upper line in the right plot of Fig. 6. It can be seen that variations of the order of 3 dB appear, which underscores the strong effects of array truncation. Errors, defined as the magnitude of the complex difference between brute-force and approximate solutions, are presented by the lines below. The errors produced by the infinite-array solution are less than 10 dB below the brute-force solution. Slightly better results are obtained when considering only one MBF, that corresponds to the infinite-array solution for scanning at broadside. Except



(a)



(b)

Fig. 7. Pattern for a  $5 \times 5$  array of tapered-slot antennas, with uniform excitation (a) E-plane, (b) H-plane. Upper lines: brute-force solution. Lower lines: error from ASM+MBF approach. Dashed: error from use of infinite-array solution for surface currents.

for one ( $N = 4$ , dashed lines), the other examples include the 8 additional MBFs obtained from the  $2 \times 2$  finite-array case. It can be seen that, for increasing values of  $N$ , the errors are steadily going down. Basis functions obtained with  $N = 4$  (24 MBFs in total) lead half-way to single machine precision; hence  $N = 4$  will be considered sufficient. This involves the calculation of the infinite-array solutions for  $N^2$  different phase shifts. These can be computed very fast, because the domain is limited to a unit cell. Details about fast implementation, involving fast calculation and tabulation of the periodic Green's function, can be found in [3].

Corresponding patterns can be seen in Fig. 7, for cuts in E plane and in H plane. The upper line corresponds to the brute-force solution, the dashed line stands for the error produced by considering the infinite-array solution on all antennas, and the lower line corresponds to the error obtained with the  $N = 4$  ASM+MBF approach (24

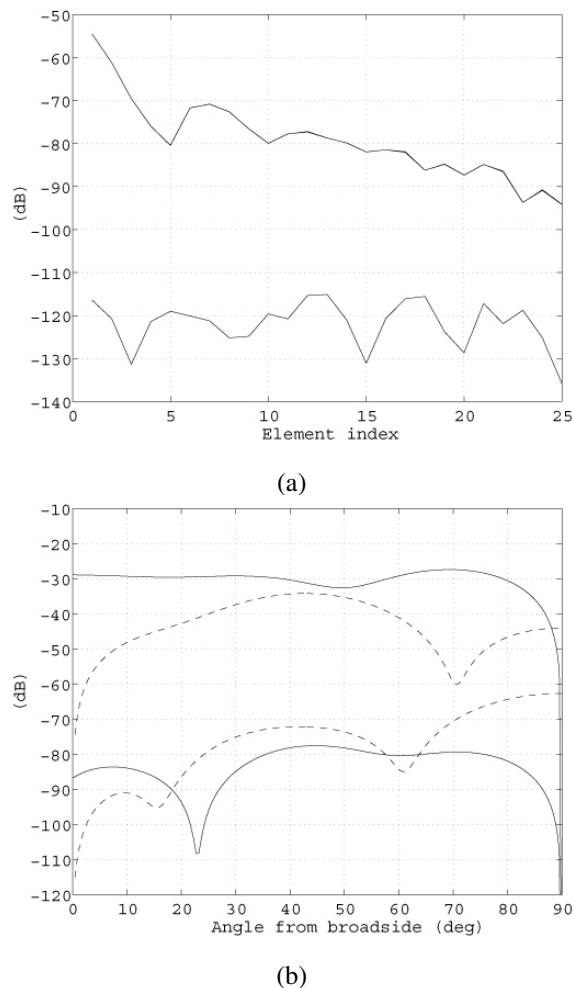


Fig. 8. For a  $5 \times 5$  array of tapered-slot antennas, excited at a corner element. (a): port currents: brute-force solution and error from ASM+MBF approach. (b): Embedded element pattern in E-plane. Solid: vertical polarisation. Dashed: horizontal polarisation. Upper curves: brute-force solution. Lower curves: error from ASM+MBF approach.

MBFs in total). Compared to the brute-force solution, the errors are located from 60 dB lower at broadside to 30 dB lower at grazing incidence.

The last example concerns excitation at a single port. A worst case has been considered, in terms of array truncation. It corresponds to excitation of a corner element. The accuracy of the result on the 25 antennas (taken along successive columns) is shown in Fig. 8(a) for the port currents, and in Fig. 8(b) for the embedded element pattern in the E plane. Upper curves correspond to the brute-force solution, lower curves represent the errors. As for the patterns, the dashed lines stand for the cross-polar component, due to the asymmetric excitation of the finite array. It can be seen that excellent accuracies are achieved, even for the cross-polar fields.

Finally, it is interesting to notice that, although lower frequency cases generally lead to stronger truncation ef-

fects, with the method presented here, significantly better accuracies were achieved for a wavelength of 60 cm (not shown here) instead of 30 cm (examples above).

## VII. CONCLUSION AND FURTHER PROSPECTS

A physically based choice of macro basis functions has been described for the full-wave simulation of finite antenna arrays. We showed that a good choice corresponds to infinite-array solutions computed on a regular grid in the reciprocal domain (domain of phase shifts between elements in both directions). This method allows capturing of current waves launched at one element, propagating over the passively terminated array, and bouncing back on the edges of the array. This method is also well suited to the particularly difficult case of electrically connected elements, without requiring the development of techniques devoted to tapering of current distributions obtained over domains defined over more than one unit cell.

The method has been demonstrated for metallic arrays of tapered-slot antennas. Since edge elements can exhibit quite different current distributions, a few more MBFs are obtained from very small ( $2 \times 2$ ) arrays. Excellent results have been obtained for both port currents and radiation patterns. It should be recalled that, once the reduced MoM impedance matrix has been obtained, solutions can be computed simultaneously at negligible computational cost for any excitation law. Future efforts will concentrate on the demonstration of this method to the case of arrays containing dielectric parts.

## REFERENCES

- [1] A. K. Skrivervik and J. R. Mosig, "Analysis of finite phased arrays of microstrip patches," *IEEE Trans. Antennas Propagat.*, vol. 41, no. 8, pp. 1105–1114, Aug. 1993.
- [2] A. Neto, S. Maci, G. Vecchi, and M. Sabbadini, "A truncated Floquet wave diffraction method for the full-wave analysis of large phased arrays. II. Generalization to 3-D cases," *IEEE Trans. Antennas Propagat.*, vol. 48, no. 4, pp. 601–611, Apr. 2000.
- [3] C. Craeye and X. Dardenne, "Element pattern analysis of wideband arrays with the help of a finite-by-infinite array approach," *IEEE Trans. Antennas Propagat.*, vol. 54, no. 2, pp. 519–526, Feb. 2006.
- [4] P. Janpugdee, P. H. Pathak, P. Mahachoklertwattana, and R. J. Burkholder, "An accelerated DFT-MoM for the analysis of large finite periodic antenna arrays," *IEEE Trans. Antennas Propagat.*, vol. 54, no. 1, pp. 279–283, Jan. 2006.
- [5] W. B. Lu, T. J. Cui, X. X. Yin, Z. G. Qian, and W. Hong, "Fast algorithms for large-scale periodic structures using subentire domain basis functions," *IEEE Trans. Antennas Propagat.*, vol. 53, no. 3, pp. 1154–1162, Mar. 2005.
- [6] E. Suter and J. R. Mosig, "A subdomain multilevel approach for the efficient MoM analysis of large planar antennas," *Microwave Optical Technol. Lett.*, vol. 26, no. 4, pp. 270–277, Mar. 2000.

- [7] L. Matekovits, G. Vecchi, G. Dassano, and M. Orefice, "Synthetic function analysis of large printed structures: The solution space sampling approach," *2001 IEEE Antennas Propagat. Soc. Int. Symp.*, Boston, pp. 568–571, Boston, June 2001.
- [8] V. Prakash and R. Mittra, "Characteristic basis function method: A new technique for efficient solution of method of moments matrix equations," *Microwave Optical Technol. Lett.*, vol. 36, no. 2, pp. 95–100, Jan. 2003.
- [9] J. Yeo, V. Prakash, and R. Mittra, "Efficient analysis of a class of microstrip antennas using the Characteristic basis function method (CBFM)," *Microwave Optical Technol. Lett.*, vol. 39, no. 6, pp. 456–464, Dec. 2003.
- [10] J. X. Wan, J. Lei, and C. H. Liang, "An efficient analysis of large-scale periodic microstrip antenna arrays using the characteristic basis function method," *Prog. Electromag. Res.*, vol. 50, pp. 61–81, Aug. 2005.
- [11] C. Craeye, "A fast impedance and pattern computation scheme for finite antenna arrays," *IEEE Trans. Antennas Propagat.*, vol. 54, no. 10, pp. 3030–3034, Oct. 2006.
- [12] W. B. Lu, T. J. Cui, and H. Zhao, "Acceleration of fast multipole method for large-scale periodic structures with finite sizes using sub-entire-domain basis functions," *IEEE Trans. Antennas Propagat.*, vol. 55, no. 2, pp. 414–421, Feb. 2007.
- [13] F. Vipiana, G. Vecchi, and P. Pirinoli, "A multiresolution system of Rao-Wilton-Glisson functions," *IEEE Trans. Antennas Propagat.*, vol. 55, no. 3, pp. 924–930, Mar. 2007.
- [14] B. A. Munk 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 half-space," *IEEE Trans. Antennas Propagat.*, vol. 27, no. 5, pp. 331–343, May 1979.
- [15] C. Craeye, A. G. Tijhuis, and D. H. Schaubert, "An efficient MoM formulation for finite-by-infinite arrays of two-dimensional antennas arranged in a three-dimensional structure," *IEEE Trans. Antennas Propagat.*, vol. 52, no. 1, pp. 271–282, Jan. 2004.
- [16] C. Craeye and R. Sarkis, "Finite array analysis through combination of macro basis functions and array scanning methods," *Proc. of 23rd Ann. Rev. of Prog. Applied Comput. Electromagnetics (ACES)*, Verona, Mar. 19–23, 2007.
- [17] J. Shin and D. H. Schaubert, "A parameter study of stripline-fed Vivaldi notch-antenna arrays," *IEEE Trans. Antennas Propagat.*, vol. 47, no. 5, pp. 879–886, May 1999.
- [18] R. Maaskant, R. Mittra, and A. G. Tijhuis, "Application of trapezoidal-shaped characteristic basis functions to arrays of electrically interconnected antenna elements," *Proc. of Int. Conf. Electromagnetics in Advanced Appl.*, Torino, Italy, Sept. 17–21, 2007.
- [19] F. Capolino, D. R. Jackson, and D. R. Wilton, "Fundamental properties of the field at the interface between air and a periodic artificial material excited by a line source," *IEEE Trans. Antennas Propagat.*, vol. 53, no. 1, pp. 3030–3034, Jan. 2005.
- [20] X. Dardenne and C. Craeye, "Application of the array scanning method with windowing to the analysis of finite rectangular periodic structures," *Proc. of the ICECOM 2005 Conference*, Dubrovnic, Croatia, May 2005.
- [21] C. Craeye, B. Parvais, and X. Dardenne, "MoM simulation of signal-to-noise patterns in infinite and finite receiving antenna arrays," *IEEE Trans. Antennas Propagat.*, vol. 52, no. 12, pp. 3245–3256, Dec. 2004.
- [22] C. Craeye, A. O. Boryszenko, and D. H. Schaubert, "Wave propagation and coupling in linear arrays with application

to the analysis of large arrays," *IEEE Trans. Antennas Propagat.*, vol. 54, no. 7, pp. 1971–1978, Jul. 2006.

- [23] E. Lucente, A. Monorchio, and R. Mittra, "Generation of characteristic basis functions by using sparse MoM impedance matrix to construct the solution of large scattering and radiation problems," *Proc. of 2006 IEEE Antennas Propagat. Symp.*, pp. 4091–4094, Albuquerque, NM, July 9–14, 2006.
- [24] S. Rao, D. Wilton, and A. Glisson, "Electromagnetic scattering by surfaces of arbitrary shape," *IEEE Trans. Antennas Propagat.*, vol. 30, no. 5, pp. 409–418, May 1982.



**Christophe Craeye** was born in Belgium in 1971. He received the Electrical Engineer and bachelor in Philosophy degrees in 1994, from the Universit catholique de Louvain (UCL). He received the Ph.D. degree in Applied Sciences from the same university in 1998. From 1994 to 1999, he was a teaching assistant at UCL and carried out research on the radar signature of the sea surface perturbed by rain, in collaboration with ESA and the Rain-sea interaction facility of NASA, Wallops Island (VA). From 1999 to 2001, he stayed as a post-doc researcher at the Eindhoven University of Technology, the University of Massachusetts and the Netherlands Institute for Research in Astronomy, where he worked on the numerical analysis of wideband phased arrays devoted to new generation radio telescopes (SKA project). Since 2002, C. Craeye is an Associate Professor (Chargé de Cours) at the Universit catholique de Louvain. His research interests are finite antenna arrays, multiple antenna systems and numerical methods for fields in periodic media. He currently is an Associate Editor of the IEEE Transactions on Antennas and Propagation.



**Rémi Sarkis** was born in Lebanon in 1982. He received in 2005 a Diploma in Electrical, Electronics, Telecommunications and Computer Engineering, from the Lebanese University - Faculty of Engineering. In 2006, he accomplished a Masters in Science at the University of Gvle-Sweden, with a specialisation in microwaves and telecommunications. In the spring of 2006, he carried out an internship at the Universit catholique de Louvain (UCL), Belgium, in collaboration with the Dominion Radio Astrophysical Observatory of NRC in Canada, on antenna arrays for the Square Kilometer Array. Since October 2006, R. Sarkis is a PhD Student at UCL. His research interests are fast numerical methods for electromagnetic fields in finite periodic structures and multi-band antenna array analysis and design.

# Robust Adaptive Beamforming Using Least Mean Mixed Norm Algorithm

<sup>1</sup>R. M. Shubair, <sup>1</sup>S. A. Jimaa, and <sup>2</sup>A. A. Omar

<sup>1</sup> Communication Engineering Dept., Khalifa University of Science, Technology and Research, Sharjah, U.A.E.

<sup>2</sup> Telecommunication Engineering Dept., Yarmouk University, Irbid, Jordan

**Abstract** – This paper proposes an accurate and rapidly-convergent algorithm for enhanced adaptive beamforming based on the combination of the least mean mixed norm (LMMN) algorithm with initialization using sample matrix inversion (SMI). The algorithm uses a mixing parameter  $\delta$  which controls the proportions of the error norms and offers an extra degree of freedom within the adaptation. Monte Carlo simulations show that the misadjustment curve has a minimum at  $\delta = 0.40$  which means that the proposed algorithm has an optimum steady-state performance at this mixing parameter value. The convergence of the algorithm is further improved by employing SMI to initialize the weights vector in the LMMN update equation. This makes the proposed SMI-initialized LMMN algorithm have a better steady state performance when compared to the least mean squares (LMS) algorithm and better stability properties when compared to the least mean fourth (LMF) algorithm. Simulation results obtained show that the developed SMI-initialized LMMN algorithm outperforms other algorithms in terms of computational efficiency, numerical accuracy, and convergence rate.

**Keywords:** Smart antennas, adaptive beamforming, and least mean squares.

## I. INTRODUCTION

Smart antennas have emerged as one of the leading innovations for achieving highly efficient networks that maximize capacity and improve quality and coverage. Smart antennas provide greater capacity and performance benefits than standard antennas because they can be used to customize and fine-tune antenna coverage patterns to the changing traffic or radio frequency (RF) conditions in a wireless network [1]. Figure 1 shows a smart antenna system which consists of a uniform linear array (ULA) for which the current amplitudes are adjusted by a set of complex weights using an adaptive beamforming algorithm. The adaptive beamforming algorithm optimizes the array output beampattern such that maximum radiated power is produced in the directions of desired mobile users and deep nulls are generated in the directions of undesired signals representing co-channel interference from mobile users in adjacent cells. Prior to adaptive beamforming, the

directions of users and interferers must be obtained using a direction-of-arrival (DOA) estimation algorithm [2].

Recent research efforts into smart antennas have varied over a wide range of methods and applications including array pattern synthesis based on null steering and multi-user beamforming using a phase control method [3], circular and hexagonal array geometries for smart antenna systems [4], adaptive and a switched beam smart antenna systems for wireless communications [5], tapered beamforming method for uniform circular arrays [6], beam steering with null and excitation constraints for linear antenna arrays [7], displaced sensor array for improved signal detection [8], and finally robust adaptive beamforming algorithms [9, 10]. The emphasis of this paper is on the development of enhanced adaptive beamforming algorithms for robust interference suppression.

Adaptive beamforming is achieved using adaptive antenna array for which the weights of the array element currents are adjusted in order to filter out the interfering signals from undesired sources, while enhancing the signal of interest from the desired source. Adaptive beamforming algorithms are typically characterized in terms of their convergence properties and computational complexity. One practical adaptive algorithm is the Least Mean Squares (LMS) which is simple to implement. It does not require measurements of the pertinent correlation functions, nor does it require matrix inversion. However, the LMS algorithm converges slowly when compared with other complicated algorithms such as the Recursive Least Square (RLS) [11]. On the other hand, Sample Matrix Inversion (SMI) algorithm has a fast convergence behavior. However, because its speedy convergence is achieved through the use of matrix inversion, the SMI algorithm is computationally intensive. Moreover, the SMI algorithm has a block adaptive approach for which it is required that the signal environment does not undergo significant change during the course of block acquisition.

Various adaptive MMSE receivers are based on the standard quadratic cost function. So far, the LMS algorithm has proved popular for many applications because of its simplicity and ease of implementation [2, 12]. However, many alternatives can also be defined to improve the adaptation performance in specific statistical environments including the Least Mean Mixed Norm (LMMN) algorithm [13, 14]. This algorithm has been used to

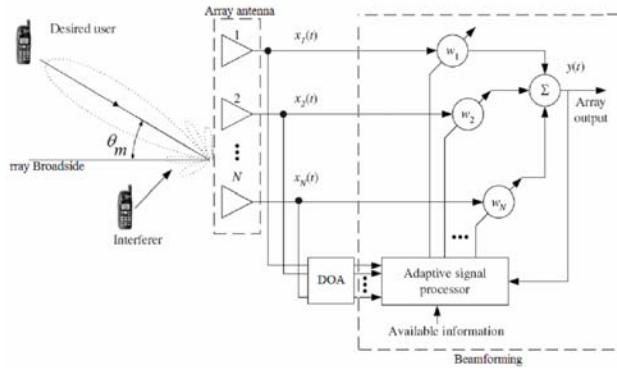


Fig. 1. A functional block diagram of a smart antenna system.

update the tap coefficients of the feedforward and feedback filters for the adaptation of the non-linear receiver, coupled with a second order phase tracking subsystem, for asynchronous DS-CDMA communication system impaired by double-spread multipath channel and Gaussian mixture impulsive noise [15]. The purpose of this paper is to develop an enhanced adaptive beamforming algorithm based on LMMN but with SMI initialization to ensure faster convergence. It is shown that judicious choice of the LMMN algorithm mixing parameter provides an algorithm with intermediate performance between the two special cases of least mean squares (LMS) and least mean fourth (LMF) algorithms. It is shown that the developed LMMN algorithm along with SMI initialization provides better steady state performance than the LMS algorithm and better stability properties than the LMF algorithm.

The rest of the paper is organized as follows: Section II describes the signal model for an adaptive beamformer based on the ULA configuration. Section III presents the theory of adaptive beamforming using the LMS algorithm, the LMF algorithm, and the proposed SMI-initialized LMMN algorithm. Simulation results are presented in Section IV showing that the proposed SMI-initialized algorithm outperforms the other algorithms. Finally, conclusions are given in Section V.

## II. SIGNAL MODEL

The standard array geometry that has been used for smart antenna systems is the uniform linear array (ULA) depicted in Fig. 2. A uniform linear array consists of  $N$  elements that are spaced apart by half wavelength ( $d = \lambda/2$ ). The inter-element spacing in a ULA is chosen to be  $\lambda/2$  in order to reduce mutual coupling effects which deteriorate the performance of the DOA estimation algorithm as demonstrated in [16-21]. If the inter-element spacing is chosen to be smaller than  $\lambda/2$ , mutual coupling effects then cannot be ignored and the DOA estimation algorithm fails to produce the desired peaks in the angular spectrum. On the other hand, increasing the inter-element spacing beyond  $\lambda/2$  results in spatial aliasing which takes the form of unwanted or

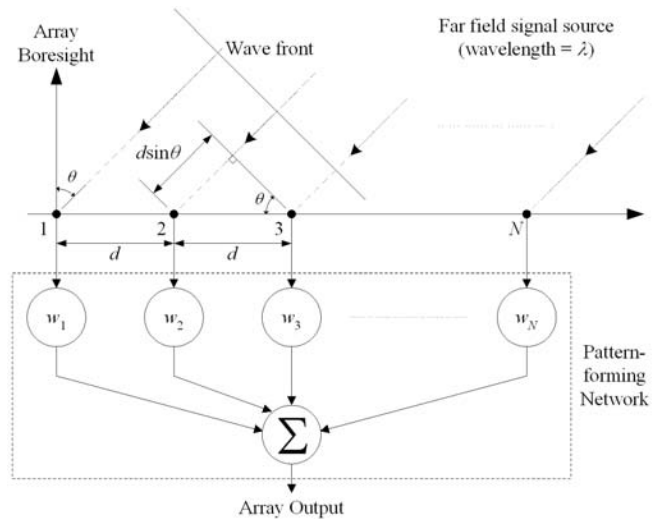


Fig. 2. Geometry of a Uniform Linear Array (ULA) of  $N$  sensors that are equally spaced apart a distance  $d = \lambda/2$ .

misplaced peaks in the angular spectrum. It is therefore concluded that  $d = \lambda/2$  represents the optimum value for the inter-element spacing in a ULA.

The main advantage of using a ULA is that it has the simplest geometry, an excellent directivity, and produces the narrowest main-lobe in a given direction in comparison to other array geometries. The ULA consists of  $N$  linear equispaced omnidirectional sensors with inter-element spacing  $d = \lambda/2$  and is positioned along the  $x$  axis with an azimuth angle  $\theta_m$  measured with respect to the  $z$  axis. It is assumed that the ULA receives  $M$  narrowband source signals  $s_m(t)$  from incidence directions  $\theta_1, \theta_2, \dots, \theta_M$ , as shown in Fig. 2. The array also receives  $I$  narrowband source signals  $s_i(t)$  from undesired (or interference) users arriving at directions  $\theta_1, \theta_2, \dots, \theta_I$ . At a particular instant of time  $t = 1, 2, \dots, K$ , where  $K$  is the total number of snapshots taken, the desired users signal vector  $\mathbf{x}_S(t)$  can be defined as,

$$\mathbf{x}_S(t) = \sum_{m=1}^M \mathbf{a}(\theta_m) s_m(t) \quad (1)$$

where  $\mathbf{a}(\theta_m)$  is the  $N \times 1$  array steering vector which represents the array response at direction  $\theta_m$  and is given by,

$$\mathbf{a}(\theta_m) = [\exp[j(n-1)\psi_m]]^T; \quad 1 \leq n \leq N \quad (2)$$

where  $[(\cdot)]^T$  is the transposition operator, and  $\psi_m$  represents the electrical phase shift from element to element along the array defined by  $\psi_m = 2\pi(d/\lambda) \sin \theta_m$  where  $d$  is the inter-element spacing and  $\lambda$  is the wavelength of the received signal. The desired users signal vector  $\mathbf{x}_S(t)$  of equation (1) can be written as,

$$\mathbf{x}_S(t) = \mathbf{A}_S s(t) \quad (3)$$

where  $\mathbf{A}_S$  is the  $N \times M$  matrix of the desired users signal

direction vectors and is given by,

$$\mathbf{A}_S = [\mathbf{a}(\theta_1), \mathbf{a}(\theta_2), \dots, \mathbf{a}(\theta_M)] \quad (4)$$

and  $\mathbf{s}(t)$  is the  $M \times 1$  desired users source waveform vector defined as,

$$\mathbf{s}(t) = [s_1(t) \quad s_2(t) \quad \dots \quad s_M(t)]^T. \quad (5)$$

We also define the undesired (or interference) users signal vector  $\mathbf{x}_I(t)$  as,

$$\mathbf{x}_I(t) = \mathbf{A}_I \mathbf{i}(t) \quad (6)$$

where  $\mathbf{A}_I$  is the  $N \times I$  matrix of the undesired users signal direction vectors and is given by,

$$\mathbf{A}_I = [\mathbf{a}(\theta_1), \mathbf{a}(\theta_2), \dots, \mathbf{a}(\theta_I)] \quad (7)$$

and  $\mathbf{i}(t)$  is the  $I \times 1$  undesired (or interference) users source waveform vector defined as,

$$\mathbf{i}(t) = [i_1(t) \quad i_2(t) \quad \dots \quad i_I(t)]^T. \quad (8)$$

The overall received signal vector  $\mathbf{x}(t)$  is given by the superposition of the desired users signal vector  $\mathbf{x}_S(t)$ , undesired (or interference) users signal vector  $\mathbf{x}_I(t)$ , and an  $N \times 1$  vector  $\mathbf{n}(t)$  which represents white Gaussian sensor noise. Hence,  $\mathbf{x}(t)$  can be written as,

$$\mathbf{x}(t) = \mathbf{x}_S(t) + \mathbf{n}(t) + \mathbf{x}_I(t). \quad (9)$$

The conventional (forward-only) estimate of the covariance matrix defined as,

$$\mathbf{R} = E\{\mathbf{x}(t)\mathbf{x}^H(t)\} \quad (10)$$

where  $E\{\cdot\}$  represents the ensemble average; and  $(\cdot)^H$  is the Hermitian transposition operator. By applying temporal averaging over  $K$  observation snapshots taken from the signals incident on the sensor array,  $\mathbf{R}$  can be approximated as, [22]

$$\mathbf{R} = \frac{1}{K} \sum_{k=1}^K \mathbf{x}(k)\mathbf{x}^H(k) \quad (11)$$

Substituting for  $\mathbf{x}(t)$  from equation (9) in equation (11) yields,

$$\begin{aligned} \mathbf{R} &= \frac{1}{K} \sum_{t=1}^K \mathbf{A}_S [\mathbf{s}(k)\mathbf{s}(k)^H] \mathbf{A}_S^H + \mathbf{n}(k) \mathbf{n}(k)^H \\ &+ \frac{1}{K} \sum_{t=1}^K \mathbf{A}_I [\mathbf{i}(k)\mathbf{i}(k)^H] \mathbf{A}_I^H. \end{aligned} \quad (12)$$

Finally, equation (12) can be written in compact form as,

$$\mathbf{R} = \mathbf{A}_S \mathbf{R}_{ss} \mathbf{A}_S^H + \sigma_n^2 \mathbf{I} + \mathbf{A}_I \mathbf{R}_{ii} \mathbf{A}_I^H \quad (13)$$

where  $\mathbf{R}_{ss} = E\{\mathbf{s}(t)\mathbf{s}^H(t)\}$  is an  $M \times M$  desired users source waveform covariance matrix,  $\mathbf{R}_{ii} = E\{\mathbf{i}(t)\mathbf{i}^H(t)\}$  is an  $I \times I$  undesired users source waveform covariance matrix,  $\sigma_n^2$  is the noise variance, and  $\mathbf{I}$  is an identity matrix of size  $N \times N$ . In general the array correlation matrix obtained in equation (13) is referred as the covariance

matrix only when the mean values of the signals and noise are zero. The arriving signals mean value must be necessarily zero because antennas can not receive d.c. signals.

### III. ADAPTIVE BEAMFORMING ALGORITHM

An adaptive beamformer, which is shown in Fig. 2, consists of multiple antennas; complex weights, the function of which is to amplify (or attenuate) and delay the signals from each antenna element; and a summer to add all of the processed signals, in order to tune out the signals not of interest, while enhancing the signal of interest. Hence, beamforming is sometimes referred to as spatial filtering, since some incoming signals from certain spatial directions are filtered out, while others are amplified. The output response of the uniform linear array is given by,

$$y(k) = \mathbf{w}^H \mathbf{x}(k), \quad (14)$$

where  $\mathbf{w}$  is the complex weights vector and  $\mathbf{x}$  is the received signal vector given in equation (9). If  $d(k)$  denotes the sequence of reference or training symbols known a priori at the receiver at time  $n$ , an error,  $e(k)$  is formed as,

$$e(k) = d(k) - \mathbf{w}(k)^H \mathbf{x}(k). \quad (15)$$

This error signal  $e$  is used by the beamformer to adaptively adjust the complex weights vector  $\mathbf{w}$  so that the mean-squared error (MSE) is minimized. It is intuitively reasonable that successive corrections to the weights vector in the direction of the negative of the gradient of the MSE function should eventually lead to minimum mean square error, at which point the weights vector assumes its optimum value. Recursive estimates for the unknown weight vector can be obtained adaptively via, [23]

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \mu \mathbf{x}^*(k) f_e(k) \quad (16)$$

where  $\mathbf{w}(k+1)$  denotes the weights vector to be computed at iteration  $n+1$ ,  $\mu$  is the algorithm step size, and  $f_e(k)$  is a scalar function of the estimation error  $e(k)$ . The step size  $\mu$  is related to the rate of convergence: in other words, how fast the algorithm reaches steady state. The smaller the step size the longer it takes the algorithm to converge. This means that a longer reference or training sequence is needed, which would reduce the payload and, hence, the bandwidth available for transmitting data.

The most popular variant of equation (16) is the least mean squares (LMS) algorithm for which the cost function to be minimized is given by,

$$J_2(k) = E\{e^2(k)\} \quad (17)$$

where  $E\{\cdot\}$  which results in an estimation error function  $f_e^{LMS}(k)$  given as,

$$f_e^{LMS}(k) = e(k). \quad (18)$$

Hence, the weights update equation (16) for LMS becomes,

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \mu \mathbf{x}^*(k)e(k). \quad (19)$$

The cost function that is minimized for the least mean fourth (LMF) algorithm is given by,

$$J_4(k) = \frac{1}{4}E\{e^4(k)\} \quad (20)$$

which results in an estimation error function  $f_e^{LMF}(k)$  given by,

$$f_e^{LMF}(k) = e^3(k) \quad (21)$$

In this case, the weights update equation (16) for LMF becomes,

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \mu \mathbf{x}^*(k)e^3(k). \quad (22)$$

Among other variants is the least mean mixed norm (LMMN) algorithm for which the cost function to be minimized is a linear mixture of the cost functions  $J_2(k)$  and  $J_4(k)$  for the LMS and LMF algorithms, respectively. It is given by,

$$J(k) = \frac{\delta}{2}J_2(k) + \frac{1-\delta}{4}J_4(k) \quad (23)$$

where the parameter  $\delta$  is called the norm mixing parameter such that  $\delta \in [0, 1]$ . This results in an estimation error function given by,

$$f_e^{LMMN}(k) = \delta e(k) + (1-\delta)e^3(k). \quad (24)$$

The weights update equation (16) for LMMN becomes,

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \mu \mathbf{x}^*(k) [\delta e(k) + (1-\delta)e^3(k)]. \quad (25)$$

It is to be noted from equation (25) that the LMMN algorithm requires at each iteration only two more multiplications and one more addition than the LMS algorithm. Moreover, when  $\delta = 1$  equation (25) becomes the weights update equation for the LMS algorithm. On the other hand, when  $\delta = 0$  equation (25) becomes the weights update equation for the LMF algorithm. Judicious choice of  $\delta$  thereby provides an algorithm with intermediate performance between that of the LMS and LMF, and a mechanism to mitigate the problem of instability within the LMF algorithm. Moreover, for operation in a statistically non-stationary environment, the mixing parameter  $\delta$  may be adapted to match appropriately the properties of measured signals.

In order to ensure the stability and convergence of the LMMN algorithm, the adaptive step size parameter  $\mu$  should be chosen within the range specified as, [13]

$$0 < \mu < \frac{1}{N.E\{x^2(k)\}} \left[ \delta + (1-\delta) \frac{1}{6.E\{n^2(k)\}} \right] \quad (26)$$

where  $E\{x^2(k)\}$  is the input signal power and  $E\{n^2(k)\}$  is the noise power. Analysis of the effect of varying the adaptive step size parameter  $\mu$  in [2, 21] showed that  $\mu$

should be chosen to be small in order to ensure numerical stability of the algorithm. Hence, in all the simulation results presented in Section 5 to follow have obtained with an adaptive step size value  $\mu = 1 \times 10^{-3}$ .

It is known that the LMF algorithm has better steady state performance than that of the LMS algorithm for applications in which the noise has a probability density function with short tail. However, its stability properties are worse than those of the LMS algorithm. On the other hand, the LMMN algorithm has better steady state performance than the LMS algorithm and better stability properties than the LMF algorithm [14]. It is for those reasons that we consider the application of LMMN algorithm to adaptive beamforming for robust interference suppression. The steady-state performance of the LMMN algorithm is a function of the norm mixing parameter  $\delta$ . The steady-state performance is quantified in terms of the misadjustment, which is defined as,

$$M = \frac{1}{\sigma_n^2} \lim_{k \rightarrow \infty} E\{\mathbf{w}(k)^H \mathbf{x}(k)\} \quad (27)$$

The effect of varying the norm mixing parameter  $\delta$  on the misadjustment  $M$  is studied in Section 5. The purpose there is to derive an optimal value for  $\delta$  for which the misadjustment  $M$  reaches a minimum value.

The weight initialization is arbitrary in the LMMN algorithm which makes it take longer (i.e., requires more iterations) to converge. To overcome this problem, we use the sample matrix inversion (SMI) algorithm to initialize the weights vector in the LMMN update equation (25). We further improve the performance of the LMMN algorithm by evaluating the initial weights vector in the LMMN weights update equation (25). SMI method is a block-data adaptive algorithm and is known to be the fastest algorithm for estimating the optimum weight vector. Because of its high complexity, SMI algorithm will be used only to estimate the initial weights vector  $\mathbf{w}(0)$  which is obtained as,

$$\mathbf{w}(0) = \mathbf{R}^{-1}(0)\mathbf{r}(0) \quad (28)$$

where the estimates of the covariance matrix  $\mathbf{R}(0)$  and cross-correlation vector  $\mathbf{r}(0)$  are given by,

$$\mathbf{R}(0) = \sum_{k=1}^B \mathbf{x}(k)\mathbf{x}^H(k) \quad (29)$$

$$\mathbf{r}(0) = \sum_{k=1}^B \mathbf{x}(k)d^*(k). \quad (30)$$

In equations (29) and (30),  $B$  represents the block size and is taken to be small just to ensure that the effect due to the change in the signal environment during the block acquisition does not affect the performance of the SMI algorithm. Also, a large block results in more matrix inversions making the algorithm computationally intensive.

The weight initialization as given in equation (28) is not any arbitrary value but an estimate of the optimum



value computed by the SMI algorithm. This means that before the LMMN adaptation begins the antenna beam is already steered to an approximate direction of the desired signal, depending on the initial SMI weight estimate. In this way, the LMMN algorithm takes little time to converge. Also, after an estimate of the initial weights is made using the SMI algorithm, the SMI-initialized LMMN algorithm uses a continuous approach to adapt itself to the changing signal environment by updating the weights for every incoming sample. Since the initial convergence is faster, the SMI-initialized LMMN algorithm takes much less time than the LMMN algorithm to adapt to the signal environment changes. Therefore, the SMI-initialized LMMN algorithm is better suited for continuous transmission systems. Numerical examples in Section 7 illustrate the improved performance of the combined LMMN/SMI algorithm in comparison with the LMMN algorithm.

#### IV. SIMULATION RESULTS

Both the desired and interfering signals take the form of a simple complex sinusoidal-phase modulated signal. By doing so it can be shown in the simulations how interfering signals of the same frequency as the desired signal can be separated to achieve rejection of co-channel interference. For simplicity purpose the reference signal  $d(k)$  is considered to be the same as the desired signal.

##### A. Optimal Value of Norm Mixing Parameter ( $\delta$ )

To find the optimal value of the norm mixing parameter  $\delta$ , Monte Carlo simulations are carried to plot the misadjustment  $M$  as defined in equation (27) versus  $\delta$ . The mixing parameter  $\delta$  is used to calculate the tap weights according to the LMMN weight update equation (25). The LMMN step size is chosen as  $\mu = 1 \times 10^{-3}$ . The mixing parameter  $\delta$  is chosen as 10 equispaced points in  $[0, 1]$ . The input signal,  $\mathbf{x}(k)$  is zero-mean and uniformly distributed with unity power. The noise signal is also zero-mean and it is obtained by adding a Gaussian distributed noise of power  $\sigma_{n_1}^2 = 0.1$  and a uniformly distributed noise of power  $\sigma_{n_2}^2 = 1.0$ . The signal-to-noise ratio ( $SNR$ ) is 10 dB. The values of misadjustment  $M$  are computed at the steady state, after  $10^5$  iterations, from equation (27) by averaging over 50 Monte Carlo trials. The variation of the misadjustment  $M$  with respect to  $\delta$  is shown in Fig. 3, where it is clear that the misadjustment curve has a well defined minimum at  $\delta = 0.40$ . Therefore by choosing  $\delta = 0.40$ , it is expected that the LMMN algorithm performs better than both LMS ( $\delta = 0$ ) and LMF ( $\delta = 1$ ) algorithms. Hence, all simulation results for the LMMN algorithm presented in Section 5 to follow are carried out with a mixing parameter value  $\delta = 0.40$  in order to ensure optimum steady state performance.

##### B. Beam pattern

Consider an array of four elements ( $N = 4$ ) and half-wavelength spacing ( $d = 0.5\lambda$ ). The array is to

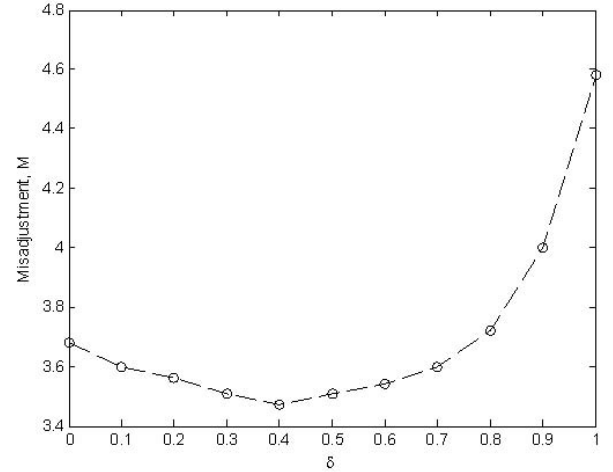


Fig. 3. Misadjustment  $M$  vs. LMMN norm mixing parameter  $\delta$  ( $SNR = 10$  dB,  $\mu = 1 \times 10^{-3}$ ).

maximize output radiation towards a source signal arriving at an angle  $\theta_S = 0^\circ$  having a signal-to-noise ratio  $SNR = 10$ dB. The array is also designed to mitigate an interference signal arriving at an angle  $\theta_I = -60^\circ$  having a signal-to-interference ratio  $SIR = -10$ dB. The number of iterations is 6000 for both algorithms. The step size for both LMS and LMMN algorithms is taken as  $\mu = 1 \times 10^{-3}$  whereas the norm mixing parameter  $\delta$  is fixed at its optimum value obtained in Section 4 as  $\delta = 0.40$ . Results are presented in Fig. 4 where the solid line represents the beam pattern obtained using the LMS algorithm and the dashed line represents the beam pattern obtained from the proposed LMMN algorithm with SMI initialization. It is evident that the pattern nulls in the case of the LMS algorithm (solid line in Fig. 4) are not deep enough to cancel the effect of interfering signals. This means that the LMS algorithm did not converge to the optimum weights solution within the given number of iterations. On the other hand, the SMI-initialized LMMN algorithm (dashed line in Fig. 4) is capable of generating deep pattern nulls (90dB below the maximum) which are strong enough to cancel the effect of the interfering signals. This means that the SMI-initialized LMMN algorithm converges faster as it reached to the optimum weights solution within the given number of iterations. This is due to the fact that the initialization of the weights vector in equation (25) was obtained from the SMI algorithm as described in equations (28) to (30).

##### C. Convergence

Here, simulations are carried out for an array with four elements ( $N = 4$ ) and half-wavelength spacing ( $d = 0.5\lambda$ ). The array is to maximize output radiation towards a source signal arriving at an angle  $\theta_S = 0^\circ$  having a signal-to-noise ratio  $SNR = 10$ dB. The array is also designed to mitigate an interference signal arriving at an angle  $\theta_I = -60^\circ$  having a signal-to-interference ratio

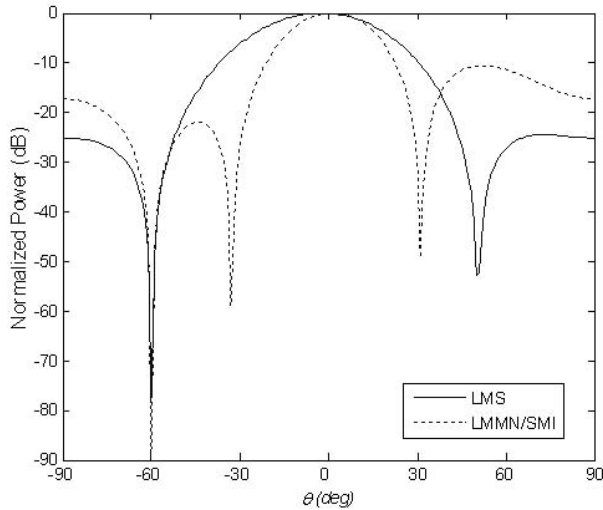


Fig. 4. Output beampattern using LMS and proposed LMMN/SMI algorithm ( $N = 4$ ,  $d = 0.5\lambda$ ,  $SNR = 10$  dB,  $SIR = -10$ dB,  $\mu = 1 \times 10^{-3}$ ,  $\delta = 0.40$ ,  $\theta_S = 0^\circ$ ,  $\theta_I = -60^\circ$ ).

$SIR = -10$ dB. The step size for both LMS and LMMN algorithms is taken as  $\mu = 1 \times 10^{-3}$  whereas the norm mixing parameter  $\delta$  is fixed at its optimum value obtained in Section 4 as  $\delta = 0.40$ . The convergence of the array weights is studied by plotting in Fig. 5 the magnitude of the array weights vector  $\mathbf{w}(1)$  versus number of iterations for both LMS algorithm and proposed LMMN algorithm with SMI initialization, respectively. It is evident from that the array weights obtained using the LMS algorithm (solid line in Fig. 5) did not converge to the steady state optimum solution within the given number of iterations. On the other hand when the proposed SMI-initialized LMMN algorithm (dashed line in Fig. 5) is used, the array weights converge to the stable value within the given number of iterations since the initialization of the weights vector in the LMMN update equation (25) was done using the SMI algorithm as described in equations (28) to (30). This verifies the improved convergence rate that is achieved when the proposed SMI-initialized LMMN algorithm is used.

#### D. Mean Square Error (MSE)

The convergence of the beamforming algorithm is examined by plotting in Figs. 6 and 7 the Mean Square Error (MSE) versus number of iterations for both the LMS algorithm and proposed SMI-initialized LMMN algorithm, respectively. Results of Fig. 7 show a significant improvement in terms of a reduced MSE when the SMI-initialized LMMN algorithm is used indicating that it has a faster convergence rate when compared to the LMS algorithm of Fig. 6. This verifies the improved accuracy that is obtained when the proposed SMI-initialized algorithm is used.

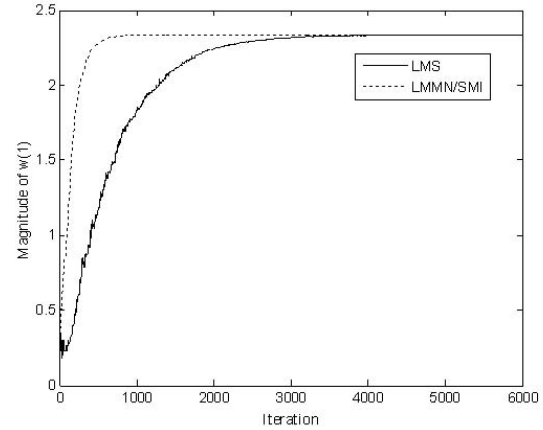


Fig. 5. Convergence of  $\mathbf{w}(1)$  using both LMS and SMI-initialized LMMN algorithm ( $N = 4$ ,  $d = 0.5\lambda$ ,  $SNR = 10$  dB,  $SIR = -10$ dB,  $\mu = 1 \times 10^{-3}$ ,  $\delta = 0.40$ ,  $\theta_S = 0^\circ$ ,  $\theta_I = -60^\circ$ ).

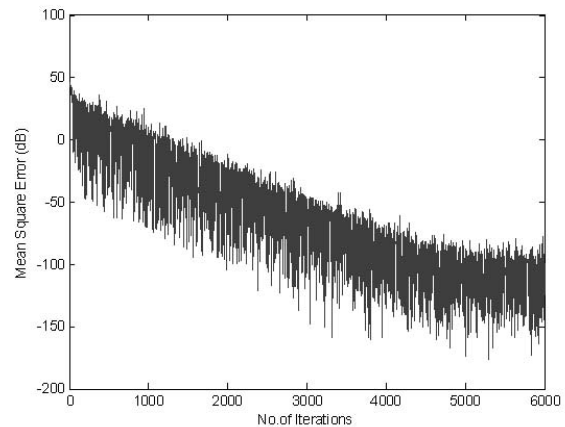


Fig. 6. Mean square error vs. number of iterations for LMS algorithm ( $N = 4$ ,  $d = 0.5\lambda$ ,  $SNR = 10$  dB,  $SIR = -10$ dB,  $\mu = 1 \times 10^{-3}$ ,  $\theta_S = 20^\circ$ ,  $\theta_I = -40^\circ$ ).

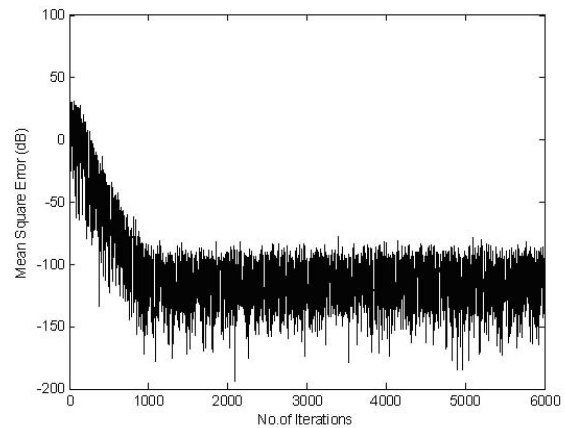


Fig. 7. Mean square error vs. number of iterations for SMI-initialized LMMN algorithm ( $N = 4$ ,  $d = 0.5\lambda$ ,  $SNR = 10$  dB,  $SIR = -10$  dB,  $\mu = 1 \times 10^{-3}$ ,  $\delta = 0.40$ ,  $\theta_S = 20^\circ$ ,  $\theta_I = -40^\circ$ ).

## V. CONCLUSIONS

An accurate and computationally-efficient adaptive beamforming algorithm based on LMMN with SMI initialization was presented. The algorithm uses a mixing parameter  $\delta$  which controls the proportions of the error norms and offers an extra degree of freedom within the adaptation. Monte Carlo simulations show that the misadjustment curve has a minimum at  $\delta = 0.40$  which means that the LMMN algorithm has an optimum steady-state performance at this mixing parameter value. The convergence of the algorithm is further improved by employing SMI to initialize the weights vector in the LMMN update equation. Hence, the SMI-initialized LMMN algorithm provides better steady state performance when compared to the least mean squares (LMS) algorithm and better stability properties when compared to the least mean fourth (LMF) algorithm. Simulation results obtained show that the proposed SMI-initialized LMMN algorithm performs better when compared to the other algorithms. The improved performance of the proposed SMI-initialized LMMN algorithm takes the form of faster convergence rate, less mean square error, as well as deeper nulls placed accurately in the directions of interference signals. These features make the proposed algorithm suitable for the design and implementation of practical smart antenna systems.

## REFERENCES

- [1] L. C. Godara, "Applications of antenna arrays to mobile communications I: performance improvement, feasibility, and system considerations," *Proceedings of the IEEE*, Vol. 85, No. 7, 1031-1060, 1997.
- [2] L. C. Godara, "Application of antenna arrays to mobile communications, Part II: Beamforming and direction-of-arrival considerations," *Proceedings of IEEE*, Vol. 85, No. 8, 1195-1245, 1997.
- [3] M. Mouhamadou, P. Vaudon, and M. Rammal, "Smart antenna array patterns synthesis: null steering and multi-user beamforming by phase control," *Progress In Electromagnetics Research*, PIER 60, 95-106, 2006.
- [4] F. Gozasht, G. Dadashzadeh, and S. Nikmhr, "A comprehensive performance study of circular and hexagonal array geometries in the LMS algorithm for smart antenna applications," *Progress In Electromagnetics Research*, PIER 68, 281-296, 2007.
- [5] F. E. Fakoukakis, S. G. Diamantis, A. P. Orfanides, and G. A. Kyriacou, "Development of an adaptive and a switched beam smart antenna system for wireless communications," *Journal of Electromagnetic Waves and Applications*, Vol. 20, No. 3, 399-408, 2006.
- [6] M. I. Dessouky, H. Sharshar, and Y. A. Albagory, "A novel tapered beamforming window for uniform concentric circular arrays," *Journal of Electromagnetic Waves and Applications*, Vol. 20, No. 14, 2077-2089, 2006.
- [7] R. Vescovo, "Beam scanning with null and excitation constraints for linear arrays of antennas," *Journal of Electromagnetic Waves and Applications*, Vol. 21, No. 2, 267-277, 2007.
- [8] R. M. Shubair and R. S. Nuaimi, "Displaced sensor array for improved signal detection under grazing incidence conditions," *Progress In Electromagnetics Research*, PIER 79, 427-441, 2008.
- [9] Y.-J. Gu, Z.-G. Shi, K. S. Chen, and Y. Li, "Robust adaptive beamforming for steering vector uncertainties based on equivalent DOAs method," *Progress In Electromagnetics Research*, PIER 79, 277-290, 2008.
- [10] Gu Y.-J., Shi Z.-G., Chen K. S., and Li Y., "Robust adaptive beamforming for a class of Gaussian steering vector mismatch," *Progress In Electromagnetics Research*, PIER 81, 315-328, 2008.
- [11] R. M. Shubair and A. Merri, "Robust algorithms for direction finding and adaptive beamforming: performance and optimization," *Proceedings of IEEE International Midwest Symposium of Circuits & Systems (MWSCAS 2004)*, Hiroshima, Japan, 589-592, July 25-28, 2004.
- [12] L. C. Godara, "Improved LMS algorithm for adaptive beamforming," *IEEE Transactions on Antennas and Propagation*, Vol. 38, No. 10, 1631-1635.
- [13] J. A. Chambers, O. Tanrikulu, and A. G. Constantinides, "Least mean mixed-norm adaptive filtering," *IEE Electronic Letters*, Vol. 30, No. 9, 1994.
- [14] O. Tanrikulu O. and J. A. Chambers, "Convergence and steady-state properties of the least-mean mixed norm (LMMN) adaptive algorithm," *IEE Proceedings on Vision, Image and Signal Processing*, Vol. 143, No. 3, 137-142, 1996.
- [15] S. A. Jimaa, M. E. Jadah, and B. S. Sharif, "Least mean mixed-norm adaptive filtering for impulsive DS-CDMA channels," *Proceedings of IEEE International Symposium on Signal Processing and Information Technology (ISSPIT 2004)*, 9-12, 2004.
- [16] E. M. Al Ardi, R. M. Shubair, and M. E. Al Mualla, "Investigation of high-resolution DOA estimation algorithms for optimal performance of smart antenna systems," *Proceedings of IEE International Conference on Third Generation Mobile Communications (3G'03)*, 460-464, 2003.
- [17] E. M. Al Ardi, R. M. Shubair, and M. E. Al Mualla, "Performance evaluation of direction finding algorithms for adaptive antenna arrays," *Proceedings of IEEE International Conference on Electronics, Circuits, and Systems (ICECS'03)*, Vol. 2, 735-738, 2003.
- [18] M. Samahi M. and R. M. Shubair, "Performance of smart antenna systems for signal detection and estimation in multipath fading environment," *Pro-*

- ceedings of IEEE International Conference on Innovations in Information Technology (IIT'06)*, 2006.
- [19] E. M. Al Ardi, R. M. Shubair, and M. E. Al Mualla, "Direction of arrival estimation in a multipath environment: an overview and a new contribution," *Applied Computational Electromagnetics Society Journal: Special Issue on Phased and Adaptive Array Antennas*, Vol. 21, No. 3, 226-239, 2006.
- [20] J. M. Samhan, R. M. Shubair, M. A. Al Qutayri, "Design and implementation of an adaptive smart antenna system," *Proceedings of IEEE International Conference on Innovations in Information Technology (IIT'06)*, 2006.
- [21] R. M. Shubair, M. A. Al Qutayri, and J. M. Samhan, "A setup for the evaluation of the MUSIC and LMS algorithms for a smart antenna system," *Journal of Communications (JCM)*, Academy Publisher, Vol. 2, No. 4, 71-77, 2007.
- [22] H. L. Van Trees, *Detection, Estimation, and Modulation Theory, Part IV: Optimum Array Processing*. John Wiley & Sons, 2002.
- [23] S. Haykin, *Adaptive Filter Theory*. Prentice-Hall, 4th Ed., 2002.

**Raed M. Shubair** received his B.Sc. degree with Distinction and Class Honors from Kuwait University, Kuwait, in June 1989 and his Ph.D. degree with Distinction from the University of Waterloo, Canada, in February 1993, both in Electrical Engineering. From March 1993 to August 1993 he was a Postdoctoral Fellow at the Department of Electrical and Computer Engineering, University of Waterloo, Canada. In August 1993 he joined Khalifa University of Science, Technology and Research (KUSTAR), formerly Etisalat University College, UAE, where he is currently an Associate Professor at the Communication Engineering Department and Leader of the Communication & Information Systems Research Group. His current research interests include adaptive array and multi-channel processing, smart antennas and MIMO systems, as well as applied and computational electromagnetic modeling of RF and microwave circuits for wireless communications. He has published over 80 papers in refereed technical journals and international conferences. He has been a member of the technical program, organizing, and steering committees of numerous international conferences and workshops. He organized and chaired a number of technical sessions in international conferences including IEEE Symposium on Antenna and Propagation (AP-S), IEEE Symposium on Electronics, Circuits and Systems (ICECS), Progress in Electromagnetics Research Symposium (PIERS), and Applied Computational Electromagnetics Symposium (ACES). Dr. Shubair is a Senior Member of IEEE, Member of ACES. He was elected to become Fellow of MIT Electromagnetics Academy in 2007. Dr. Shubair is a founding member of the IEEE UAE Signal Processing and Communications Joint So-

cieties Chapter. He acted as Symposium Vice-Chair of the 2008 IEEE Canadian Conference on Electrical and Computer Engineering: Communications and Networking Symposium. Dr. Shubair is Editor for the Journal of Communications, Academy Publisher. He is listed in Who's Who in Electromagnetics and in several editions of Who's Who in Science and Engineering.

**Shihab A. Jimaa** received the M.Sc. degree in Digital Communications Systems from Loughborough University, UK, in 1986 and the Ph.D., Data transmissions over ADPCM links, from the same university in 1989. In the UK, he has worked at Loughborough University, Warwick University, and University of Hertfordshire. He has also worked at Amman University in Jordan and Near East University in Cyprus. In 2002 he joined Khalifa University of Science, Technology and Research, UAE where he is currently an Associate Professor at the Communication Engineering Department. His research interests include adaptive algorithms for channel equalization and beamforming, system identification, signal processing for communications, and adaptive receiver structures for DS-CDMA mobile communication systems. He has authored over 50 technical papers in various journals and referred international conferences. Dr. Jimaa is a Chartered Electrical Engineer, Member of the IET, Senior Member of the IEEE, and Member of IASTED International Program Committee on Telecommunications.

**Amjad A. Omar** received the B.SC. and M.SC. in Electrical Engineering from Kuwait University in 1985 and 1988, respectively. He received the Ph.D. in Electrical Engineering from the University of Waterloo, Canada, in 1993. He worked at the Communications Research Center in Ottawa for two years (1993-1995). Since 1995 he has been working at several universities in the UAE and Jordan and currently he is an Associate Professor at the Telecommunication Engineering Department of Hijawi College for Engineering Technology at Yarmouk University in Jordan. His research interests include numerical electromagnetics, design of microwave and millimeter wave circuit components, biomedical applications of electromagnetics, and smart antennas. He has over 60 publications in various journals and referred international conferences.

# An Accurate Reduced-Order Polynomial Solution for Root-MUSIC Source Localization Using Displaced Sensor Arrays

R. M. Shubair

Communication Engineering Department  
Etisalat University College, P.O.Box 573, Sharjah, UAE  
Tel: + 971 6 5611333, Fax: + 971 6 5611789  
E-mail: rshubair@ece.ac.ae

**Abstract** – This paper proposes a modified Root-MUSIC direction finding algorithm for source localization using a displaced sensor array (DSA) configuration which utilizes two parallel-displaced arrays in the vertical plane. It is shown that the proposed configuration utilizes the spatial displacement of array sensors in both the horizontal and azimuth directions together with the symmetry of the two parallel arrays in order to reduce the rank of the spatial covariance matrix. This results in a reduced-order Root-MUSIC polynomial for which the complex roots correspond to the desired directions of the radiating sources to be localized. Simulation results show that the developed algorithm outperforms the standard Root-MUSIC algorithm for conventional uniform linear arrays (ULAs) in terms of computational efficiency, numerical accuracy, and angular resolution.

**Keywords:** Smart antennas, source localization, direction finding, and Root-MUSIC.

## I. INTRODUCTION

The development of personal communication devices is a challenging topic in modern electromagnetic research. The number of users that can actually interact at the same time with the base stations is very high. It is therefore necessary to develop efficient methods, which are able to track the desired users and mitigate the effects of interference signals. The use of multiple antennas seems to be very helpful in enhancing the performance of transmit/receive systems in the communication networks. In particular, multiple antennas offer several advantages: They allow increasing the channel capacity; reducing channel fading by using the spatial diversity of the antenna array; and, finally, mitigating co-channel and inter-symbol interferences.

Spatial filtering methods using advanced antenna techniques, smart or adaptive antennas, have received much attention over the last few years. Filtering in the spatial domain can separate spectrally and temporally overlapping signals from multiple mobile users, and hence the performance of a system can be significantly improved. Particular interest in such adaptive antennas

has been shown with regard to code-division multiple-access (CDMA) systems. In CDMA systems, all users communicate simultaneously in the same frequency band, and hence multiple-access interference (MAI) is one of the major causes of transmission impairment [1].

Furthermore, since antenna arrays generate beams with a maximum toward the desired users and nulls in the directions of interferences, they play an important role in improving the performance of both the base stations and the mobiles. To this end, an essential step is the source localization or estimation of the directions of arrival (DOAs) of the waves that impinge on the antenna array. Several methods have been proposed in the literature for source localization. Among them, subspace eigenanalysis-based methods such as MUSIC, and its derivative version Root-MUSIC [2], seem to be popular due to their high resolution capability and low computational complexity.

A smart antenna system at the base station of a cellular mobile system is depicted in Fig. 1. It consists of a uniform linear antenna array for which the currents are adjusted by a set of complex weights using an adaptive beamforming algorithm. The adaptive beamforming algorithm optimizes the array output beam pattern such that maximum radiated power is produced in the directions of desired mobile users and deep nulls are generated in the directions of undesired signals representing co-channel interference from mobile users in adjacent cells. Prior to adaptive beamforming, the directions of users and interferences must be obtained using a source localization or direction-of-arrival estimation algorithm [3].

The concept of displaced sensor array (DSA) was introduced by the author in [4-6] where it was shown that such a configuration, which is composed of two parallel-displaced sensor arrays, can improve the performance of the smart antenna system. The proposed DSA configuration has several other advantages. First, it maintains almost the same radiation aperture as the conventional uniform linear array yet it can handle more signals from users and interferers because it has more array sensors when compared to the conventional uniform linear array. Second, the horizontal displacement between the two parallel arrays allows for resolving correlated signals encountered in multipath propagation environment

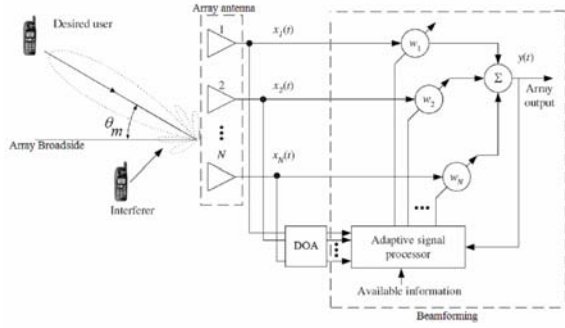


Fig. 1. A functional block diagram of a smart antenna system.

without having to apply spatial smoothing techniques. Moreover, the vertical separation between the two parallel arrays allows for resolving signals arriving in the vertical plane at the endfire direction.

In this paper we show that using the proposed DSA configuration results also in a reduced-order Root-MUSIC polynomial. This in turn leads to a more efficient source localization when compared to conventional Root-MUSIC for uniform linear arrays.

The paper is organized as follows: Section II presents the theory of Root-MUSIC source localization for uniform linear array (ULA). Section III derives a reduced-order modified Root-MUSIC polynomial using the proposed displaced sensor array (DSA) configuration. Section IV presents simulation results showing the performance improvement obtained when the modified Root-MUSIC algorithm proposed in this paper is used for cases involving signals incident close to array broadside as well as endfire direction. Finally, some conclusions are given in Section V.

## II. ROOT-MUSIC FOR UNIFORM LINEAR ARRAY (ULA)

Let a uniform linear array (ULA) be composed of  $N$  sensors, and let it receive  $M$  ( $M < N$ ) narrowband sources  $s_m(t)$  impinging from directions  $\theta_1, \theta_2, \dots, \theta_M$ , as shown in Fig. 2. Assume that there are  $K$  snapshots  $\mathbf{x}(1), \mathbf{x}(2), \dots, \mathbf{x}(K)$  available. The  $N \times 1$  array observation vector is modeled as,

$$\mathbf{x}(t) = \sum_{m=1}^M \mathbf{a}_{ULA}(\theta_m) s_m(t) + \mathbf{n}(t) \quad (1a)$$

$$= \mathbf{A}_{ULAS}(t) + \mathbf{n}(t), \quad (1b)$$

where  $\mathbf{A}_{ULA}$  is the  $N \times M$  matrix of the signal direction vectors and is given by,

$$\mathbf{A}_{ULA} = [\mathbf{a}_{ULA}(\theta_1), \mathbf{a}_{ULA}(\theta_2), \dots, \mathbf{a}_{ULA}(\theta_M)] \quad (2)$$

$\mathbf{a}_{ULA}(\theta_m)$  is the  $N \times 1$  steering vector of a ULA given by,

$$\mathbf{a}_{ULA}(\theta_m) = [1, e^{j(2\pi/\lambda)d \sin \theta_m}, \dots, e^{j(2\pi/\lambda)(N-1)d \sin \theta_m}]^T \quad (3)$$

where  $1 \leq m \leq M$ . In addition  $\mathbf{s}(t)$  is an  $M \times 1$  vector of source waveforms;  $\mathbf{n}(t)$  is an  $N \times 1$  vector of white sensor noise;  $\lambda$  is the wavelength;  $d$  is the inter-element spacing; and  $(\cdot)^T$  is the transpose. Equation (3) can be rewritten as,

$$\mathbf{a}_{ULA}(z_m) = [1 \quad z_m \quad \dots \quad z_m^{N-1}]^T \quad (4a)$$

$$= [1 \quad e^{j\psi_m} \quad \dots \quad e^{j(N-1)\psi_m}]^T \quad (4b)$$

$$= [\exp\{(2\pi/\lambda)(n-1)d \sin \theta_m\}]^T, \quad (4c)$$

where  $z_m = e^{j\psi_m}$ ,  $\psi_m = (2\pi/\lambda)d \sin \theta_m$ , and  $1 \leq n \leq N$ .

The conventional (forward only) estimate of the covariance matrix is defined as,

$$\mathbf{R} = E\{\mathbf{x}(t)\mathbf{x}^H(t)\} \quad (5)$$

where  $E\{\cdot\}$  represents the ensemble average; and  $(\cdot)^H$  is the Hermitian transposition operator. Equation (5) can be approximated by applying temporal averaging over  $K$  snapshots (or samples) taken from the signals incident on the sensor array. This averaging process leads to forming a spatial correlation (or covariance) matrix  $\mathbf{R}$  given by, [7]

$$\mathbf{R} = \frac{1}{K} \sum_{k=1}^K \mathbf{x}(k)\mathbf{x}^H(k). \quad (6)$$

Substituting for  $\mathbf{x}(t)$  from equation (1) in equation (6) yields,

$$\mathbf{R} = \mathbf{A}_{ULA}\mathbf{R}_{ss}\mathbf{A}_{ULA}^H + \sigma_n^2\mathbf{I} \quad (7)$$

where  $\mathbf{R}_{ss} = E\{\mathbf{s}(t)\mathbf{s}^H(t)\}$  is an  $M \times M$  source waveform covariance matrix,  $\sigma_n^2$  is the noise variance, and  $\mathbf{I}$  is an identity matrix. The matrix  $\mathbf{R}$  is centro-Hermitian if, [7]

$$\mathbf{R} = \mathbf{J}\mathbf{R}^*\mathbf{J} \quad (8)$$

where  $\mathbf{J}$  is the exchange matrix with ones on its anti-diagonal and zeros elsewhere, and  $(\cdot)^*$  stands for complex conjugate. The covariance matrix  $\mathbf{R}$  in equation (8) is known to be centro-Hermitian if and only if  $\mathbf{S}$  is a diagonal matrix, i.e., when the signal sources are uncorrelated.

A common subspace based DOA estimation algorithm is MUSIC (Multiple Signal Classification) [3]. This method is based on the eigen-decomposition of the covariance matrix  $\mathbf{R}$  into a signal subspace having  $M$  eigenvalues with corresponding eigenvectors  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_M$ , and noise subspace having  $(N - M)$  eigenvalues with corresponding eigenvectors  $\mathbf{v}_{M+1}, \mathbf{v}_{M+2}, \dots, \mathbf{v}_N$ . Let  $\mathbf{V}_s$  be the matrix whose columns are the source subspace eigenvectors, and  $\mathbf{V}_n$

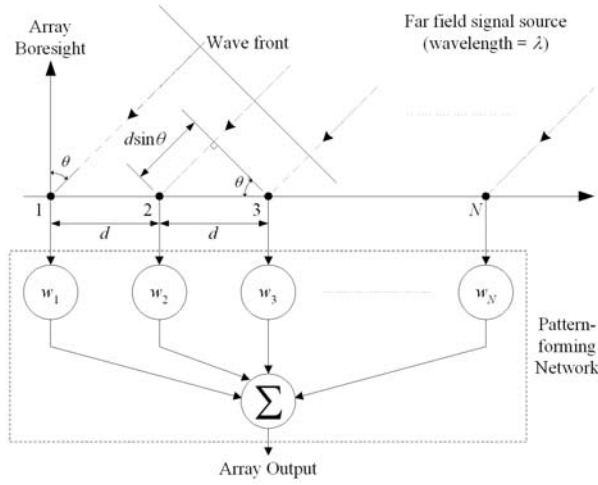


Fig. 2. Geometry of a uniform linear array (ULA).

be the matrix whose columns are the noise subspace eigenvectors, i.e.,

$$\mathbf{V}_s = [\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_M] \quad (9)$$

$$\mathbf{V}_n = [\mathbf{v}_{M+1}, \mathbf{v}_{M+2}, \dots, \mathbf{v}_N]. \quad (10)$$

The covariance matrix  $\mathbf{R}$  can then be expressed as, [7]

$$\mathbf{R} = \mathbf{V}\mathbf{I}\mathbf{V}^H \quad (11a)$$

$$= \mathbf{V}_s \mathbf{\Pi}_s \mathbf{V}_s^H + \sigma^2 \mathbf{V}_n \mathbf{V}_n^H \quad (11b)$$

where the subscripts  $s$  and  $n$  stand for signal and noise subspace, respectively. In equation (11a)  $\mathbf{\Pi}_s$  is a diagonal matrix given by  $\mathbf{\Pi}_s = \text{diag}\{\pi_1, \pi_2, \dots, \pi_M\}$ . Hence, the normalized MUSIC angular spectrum is defined as, [7, 8]

$$P_{ULA}(\theta) = \frac{\mathbf{A}_{ULA}^H \mathbf{A}_{ULA}}{\mathbf{A}_{ULA}^H \mathbf{V}_n \mathbf{V}_n^H \mathbf{A}_{ULA}}. \quad (12)$$

By examining the denominator in equation (12) it is evident that peaks in the MUSIC angular spectrum occur at angles  $\theta$  for which the array manifold matrix  $\mathbf{A}_{ULA}$  is orthogonal to the noise subspace matrix  $\mathbf{E}_n$ . Those angles  $\theta$  define the desired angles of arrival of the of the narrowband source signals impinging on the ULA. A comprehensive performance evaluation of the MUSIC algorithm for DOA estimation can be found in [9, 12].

The denominator in equation (12) can be rearranged to form the conventional Root-MUSIC polynomial as, [8]

$$G_{ULA}(z) = \mathbf{a}_{ULA}^T(1/z) \mathbf{V}_n \mathbf{V}_n^H \mathbf{a}_{ULA}(z). \quad (13)$$

The roots of the polynomial  $G_{ULA}(z)$  defined in equation (13) occur in pairs ( $z$  and  $1/z$ ) with the same argument in the  $z$ -plane. The roots  $z_m$  which lie on the unit circle (or close to it) will correspond to the directions of arrival  $\theta_m$  of the incident signals while the other roots

are spurious roots. The directions of arrival  $\theta_m$  are thus given by,

$$\theta_m = \sin^{-1} \left[ \left( \frac{\lambda}{2\pi d} \right) \arg(z_m) \right]. \quad (14)$$

### III. MODIFIED ROOT-MUSIC FOR DISPLACED SENSOR ARRAY (DSA)

The DSA configuration consists of two parallel ULAs displaced by a horizontal distance  $d = \lambda/4$  and vertical separation  $s = \lambda/2$ , as shown in Fig. 3. Each of the two parallel ULAs consists of  $N$  linear equispaced omnidirectional sensors with inter-element spacing  $d = \lambda/2$ . The two parallel ULAs are positioned along the  $x$  axis with an azimuth angle  $\theta_m$  measured with respect to the  $z$  axis. It is assumed that the DSA configuration receives  $M$  narrowband source signals  $s_m(t)$  from incidence directions  $\theta_1, \theta_2, \dots, \theta_M$ . At a particular instant of time  $t = 1, 2, \dots, K$ , where  $K$  is the total number of snapshots taken, the desired users signal vector  $\mathbf{x}(t)$  is given by,

$$\mathbf{x}(t) = \sum_{m=1}^M [\mathbf{a}_1(\theta_m) + \mathbf{a}_2(\theta_m)] s_m(t) + \mathbf{n}(t) \quad (15)$$

where  $\mathbf{n}(t)$  is the sensor noise vector modeled as temporally white and zero mean complex Gaussian process,  $\mathbf{a}_1(\theta_m)$  and  $\mathbf{a}_2(\theta_m)$  are the steering (or response) vectors for the two parallel arrays with respect to  $\theta_m$ , which represents the angle of arrival of the  $m$ th signal. The first steering vector  $\mathbf{a}_1(\theta_m)$  has dimensions  $N \times 1$  and represents the space factor of the first array with respect to direction  $\theta_m$ . It is given by,

$$\mathbf{a}_1(\theta_m) = \{\exp[j(n-1)\psi_m]\}^T, \quad 1 \leq n \leq N \quad (16)$$

where  $\{\}^T$  is the transposition operator, and  $\psi_m$  represents the electrical phaseshift from element to element along the array defined as  $\psi_m = 2\pi(d/\lambda) \sin \theta_m$ , where  $d$  is the inter-element spacing and  $\lambda$  is the wavelength of the received signal. The second steering vector  $\mathbf{a}_2(\theta_m)$  has dimensions  $N \times 1$  also and represents the space factor of the second array with respect to direction  $\theta_m$ . It is given by,

$$\mathbf{a}_2(\theta_m) = \mathbf{a}_1(\theta_m) \cdot F_s(\theta_m) \cdot F_\Delta(\theta_m) \quad (17)$$

where  $F_s(\theta_m)$  and  $F_\Delta(\theta_m)$  represent the space factors due to the vertical separation  $s$  and horizontal displacement  $\Delta$  of the two parallel arrays, respectively. These are given by,

$$F_s(\theta_m) = \exp \left[ -j2\pi \left( \frac{s}{\lambda} \right) \cos \theta_m \right] \quad (18)$$

$$F_\Delta(\theta_m) = \exp \left[ -j2\pi \left( \frac{\Delta}{\lambda} \right) \sin \theta_m \right]. \quad (19)$$

Equation (15) can be written as,

$$\mathbf{x}(t) = \sum_{m=1}^M \mathbf{a}_{DSA}(\theta_m) s_m(t) + \mathbf{n}(t) \quad (20)$$

where  $\mathbf{a}_{DSA}(\theta_m)$  is the DSA steering vector given as the superposition of the steering vectors of the two parallelly-displaced ULAs, i.e.,

$$\mathbf{a}_{DSA}(\theta_m) = \mathbf{a}_1(\theta_m) + \mathbf{a}_2(\theta_m). \quad (21)$$

The combination of all possible steering vectors forms the array manifold (or steering vector) matrices  $\mathbf{A}_1$  and  $\mathbf{A}_2$  of size  $N \times M$  each, i.e.,

$$\mathbf{A}_1 = [\mathbf{a}_1(\theta_1), \mathbf{a}_1(\theta_2), \dots, \mathbf{a}_1(\theta_M)] \quad (22)$$

$$\mathbf{A}_2 = [\mathbf{a}_2(\theta_1), \mathbf{a}_2(\theta_2), \dots, \mathbf{a}_2(\theta_M)]. \quad (23)$$

The received signal vector  $\mathbf{x}(t)$  of equation (15) can then be written as,

$$\mathbf{x}(t) = \mathbf{A}_{DSA} \mathbf{s}(t) + \mathbf{n}(t) \quad (24)$$

where  $\mathbf{A}_{DSA}$  is the overall array manifold matrix and is given by,

$$\mathbf{A}_{DSA} = \mathbf{A}_1 + \mathbf{A}_2. \quad (25)$$

The covariance matrix for the DSA configuration can then be expressed as,

$$\mathbf{R} = \mathbf{A}_{DSA} \mathbf{R}_{ss} \mathbf{A}_{DSA}^H + \sigma_n^2 \mathbf{I}. \quad (26)$$

Using the expression for  $\mathbf{A}_{DSA}$  as given in (24), the denominator in equation (12) results in the following modified Root-MUSIC polynomial,

$$G_{DSA}(z) = \mathbf{a}_{DSA}^T(1/z) \mathbf{V}_n \mathbf{V}_n^H \mathbf{a}_{DSA}(z). \quad (27)$$

$G_{DSA}(z)$  given in equation (27) for the proposed DSA has a reduced order compared to the order of the conventional Root-MUSIC polynomial  $G_{ULA}(z)$  derived in equation (13) for ULA.

## IV. SIMULATION RESULTS

### A. Detection of Signals Incident Closer to Array Broad-side Direction

The results for Root MUSIC were obtained using a ULA configuration with  $N = 12$  elements in the array, and compared them to those obtained using the modified Root-MUSIC using the DSA configuration with  $N/2 = 6$  elements in each array. Inter-element spacing of  $d = \lambda/2$  is maintained in both configurations. This is essential to

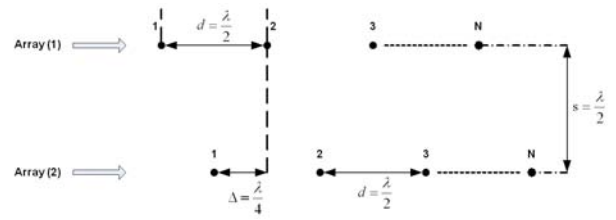


Fig. 3. Proposed displaced sensor array (DSA) configuration consists of two parallelly-displaced uniform linear arrays (ULAs) in the vertical plane displaced horizontally by a distance  $\Delta = \lambda/4$  and separated vertically by a distance  $s = \lambda/2$ .

reduce the effects of inter-element mutual coupling. We have assumed a signal-to-noise ratio  $SNR = 10\text{dB}$  and the number of snapshots  $K = 100$ . There are  $M = 5$  incoming signals to be detected arriving at directions  $\theta_m$  close to the array broadside direction ( $\theta = 0^\circ$ ). These directions are:  $\theta_1 = 30^\circ$ ,  $\theta_2 = 20^\circ$ ,  $\theta_3 = 10^\circ$ ,  $\theta_4 = -20^\circ$ , and  $\theta_5 = -40^\circ$ .

Results are presented in Table 1 showing that more accurate results with less percentage error are obtained for the DOA estimates when the modified Root-MUSIC algorithm is used. These results are also plotted in Fig. 4 which clearly demonstrates that the modified Root-MUSIC for DSA outperforms conventional Root-MUSIC for ULA in terms of numerical accuracy coupled with computational efficiency. The latter feature is evident from the reduced-order complex-root polynomial obtained as a result of the modified Root-MUSIC algorithm proposed for DSA of this paper.

### B. Detection of Signals Incident Closer to Array Endfire Direction

The results for Root MUSIC were obtained using a ULA configuration with  $N = 14$  elements in the array, and compared them to those obtained using the modified Root-MUSIC using the DSA configuration with  $N/2 = 7$  elements in each array. Inter-element spacing of  $d = \lambda/2$  is maintained in both configurations. We have assumed a signal-to-noise ratio  $SNR = 10\text{dB}$  and the number of snapshots  $K = 100$ . There are  $M = 6$  incoming signals to be detected arriving at directions  $\theta_m$  close to the array endfire direction ( $\theta = \pm 90^\circ$ ). These directions are:  $\theta_1 = -85^\circ$ ,  $\theta_2 = -80^\circ$ ,  $\theta_3 = -75^\circ$ ,  $\theta_4 = 75^\circ$ ,  $\theta_5 = 80^\circ$ , and  $\theta_6 = 85^\circ$ .

Results are presented in Table 2 showing that more accurate results with less percentage error are obtained for the DOA estimates when the modified Root-MUSIC algorithm is used. These results are also plotted in Fig. 5 which clearly demonstrates that the modified Root-MUSIC for DSA outperforms conventional Root-MUSIC for ULA in terms of numerical accuracy coupled with computational efficiency. The latter feature is evident from the reduced-order complex-root polynomial obtained as a result of the modified Root-MUSIC algorithm proposed for DSA of this paper.



Table 1. Comparison of Root-MUSIC and Modified Root-MUSIC (incidence close to array broadside direction) ( $M = 5$ ,  $d = 0.5\lambda$ ,  $SNR = 10\text{dB}$ , and  $K = 100$ ).

$\theta_m^{exact}$	Root-MUSIC		Modified Root-MUSIC	
	$\theta_m$	% error	$\theta_m$	% error
$30^\circ$	$30.78^\circ$	2.60%	$30.11^\circ$	0.37%
$20^\circ$	$20.53^\circ$	2.65%	$20.21^\circ$	1.05%
$10^\circ$	$10.16^\circ$	1.60%	$10.02^\circ$	0.20%
$-20^\circ$	$-20.33^\circ$	1.65%	$-20.07^\circ$	0.35%
$-40^\circ$	$-40.6^\circ$	1.50%	$-40.12^\circ$	0.3%

Table 2. Comparison of Root-MUSIC and Modified Root-MUSIC (incidence close to array endfire direction) ( $M = 6$ ,  $d = 0.5\lambda$ ,  $SNR = 10\text{dB}$ , and  $K = 100$ ).

$\theta_m^{exact}$	Root-MUSIC		Modified Root-MUSIC	
	$\theta_m$	% error	$\theta_m$	% error
$-85^\circ$	$-76.2^\circ$	10.35%	$-80.9^\circ$	4.82%
$-80^\circ$	$-73.1^\circ$	8.63%	$-76.8^\circ$	4.00%
$-75^\circ$	$-70.9^\circ$	5.47%	$-73.5^\circ$	2.00%
$75^\circ$	$71.9^\circ$	4.13%	$73.8^\circ$	1.60%
$80^\circ$	$75.3^\circ$	5.88%	$77.2^\circ$	2.25%
$85^\circ$	$77.5^\circ$	8.82%	$82.9^\circ$	2.47%

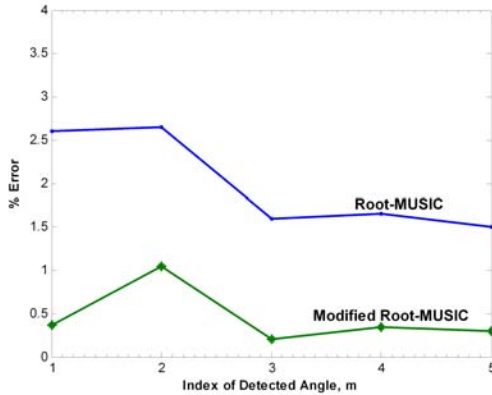


Fig. 4. Comparison of Root-MUSIC and Modified Root-MUSIC for incidence close to the array broadside direction.

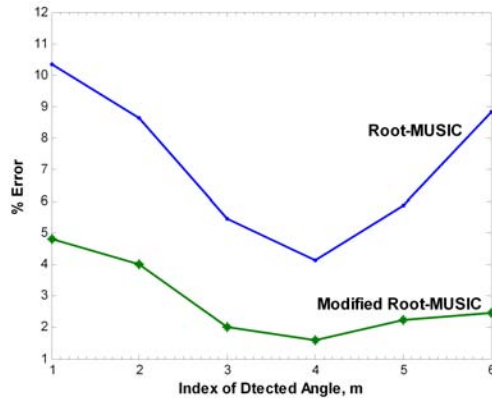


Fig. 5. Comparison of Root-MUSIC and Modified Root-MUSIC for incidence close to the array endfire direction.

## V. CONCLUSIONS

We have used a parallelly-displaced sensor array configuration to derive a reduced-order polynomial expression for the Root-MUSIC source localization method. The derived reduced-order complex-root polynomial makes the proposed method for displaced sensor arrays much more efficient in terms of computational complexity when compared to the conventional Root-MUSIC method used in conjunction with uniform linear arrays. Moreover, the horizontal displacement in the displaced sensor array configuration between the two parallel arrays allows for resolving correlated signals encountered in multipath propagation environment without having to apply spatial smoothing techniques [11]. Moreover, the vertical separation between the two parallel arrays allows for resolving signals arriving in the vertical plane at the endfire direction. Results were presented to show the improved performance of the proposed method in terms of computational efficiency and numerical accuracy.

## REFERENCES

- [1] L. C. Godara, "Applications of antenna arrays to mobile communications I: Performance improvement, feasibility, and system considerations," *Proceedings of IEEE*, vol. 85, no. 7, pp. 1031-1060, July 1997.
- [2] M. Pesavento, A. B. Gershman, and M. Haardt, "Unitary Root-MUSIC with a real-valued eigendecomposition: A theoretical and experimental performance study," *IEEE Transactions on Signal Processing*, vol. 48, no. 5, pp. 1306-1314, May 2000.
- [3] L. C. Godara, "Applications of antenna arrays to mobile communications II: Beamforming and direction-of-arrival considerations," *Proceedings of IEEE*, vol. 85, no. 8, pp. 1195-1245, August 1997.
- [4] R. M. Shubair, "Improved smart antenna design using displaced sensor array configuration," *Proceedings of 2006 International Conference of Applied Computational Electromagnetics Society (ACES 2006)*, Miami, Florida, USA, March 12-16, 2006.
- [5] R. M. Shubair, "Improved smart antenna design using displaced sensor array configuration," *Applied Computational Electromagnetics Society (ACES) Journal*, vol. 22, no. 2, pp. 83-87, March 2007.
- [6] R. M. Shubair, "Reduced-order Root-MUSIC source localization using displaced sensor arrays," *Proceedings of 2007 International Conference of Applied Computational Electromagnetics Society (ACES 2007)*, Verona, Italy, March 19-23, 2007.
- [7] S. Haykin, *Adaptive Filter Theory*. Prentice-Hall, 4th Ed., 2002.
- [8] H. L. Van Trees, *Detection, Estimation, and Modulation Theory, Part IV: Optimum Array Processing*. John Wiley & Sons, 2002.

- [9] E. M. Al Ardi, R. M. Shubair, and M. E. Al Mualla, "Investigation of high-resolution DOA estimation algorithms for optimal performance of smart antenna systems," *Proceedings of IEE International Conference on Third Generation Mobile Communications (3G'03)*, London, UK, pp. 460-464, Jun 25-27, 2003.
- [10] E. M. Al Ardi, R. M. Shubair, and M. E. Al Mualla, "Performance evaluation of direction finding algorithms for adaptive antenna arrays," *Proceedings of IEEE International Conference on Electronics, Circuits, and Systems (ICECS'03)*, Sharjah, UAE, pp. 735-738, December 14-17, 2003.
- [11] E. M. Al Ardi, R. M. Shubair, and M. E. Al Mualla, "Direction of arrival estimation in multipath environment: an overview and a new contribution," *Applied Computational Electromagnetics Society (ACES) Journal*, vol. 21, no. 3, pp. 226-239, November 2006.
- [12] R. M. Shubair, M. A. Al Qutayri, and J. M. Samhan, "A setup for the evaluation of MUSIC and LMS algorithms for a smart antenna system," *Journal of Communications, Academy Publisher*, vol. 3, no. 4, July 2007.

**Raed M. Shubair** received his B.Sc. degree with Distinction and Class Honors from Kuwait University, Kuwait, in June 1989 and his Ph.D. degree with Distinction from the University of Waterloo, Canada, in February 1993, both in Electrical Engineering. From March 1993 to August 1993 he was a Postdoctoral Fellow at the Department of Electrical and Computer Engineering, University of Waterloo, Canada. In September 1993 he joined Etisalat University College, UAE, where he is currently an Associate Professor at the Communication Engineering Department and Leader of the Communication & Information Systems Research Group. His current research interests include adaptive array processing, smart antennas and MIMO systems, as well as applied and computational electromagnetic modeling of RF and microwave circuits for wireless communications. He has published widely in refereed technical journals and international conferences. He has been a member of the technical program, organizing, and steering committees of numerous international conferences and workshops. He organized and chaired a number of technical sessions in international conferences including IEEE Symposium on Antenna and Propagation (AP-S), IEEE Symposium on Electronics, Circuits and Systems (ICECS), Progress in Electromagnetics Research Symposium (PIERS), and Applied Computational Electromagnetics Symposium (ACES). Dr. Shubair is a Senior Member of the IEEE and a Member of the Applied Computational Electromagnetics Society. He was also elected to the MIT Electromagnetics Academy in 2001. While at the University of Waterloo he received several research grants and scholarships from Canada National Science and Engineering Research Council (NSERC), Canada National Research Council (NRC), and Ontario Graduate Scholarship (OGS). Dr. Shubair supervised a large number of research projects including Adaptive Beamforming for Next-Generation Wireless Communications, recipient of the 2004 IEE Award, as well as Design of Optimum SMI Beamformers for Spatial Interference Rejection, recipient of the 2005 IEE Award. His papers were selected for three successive years (2004-2006) amongst the "Selected Best Papers" of the IEEE International Conference on Innovations in Information Technology held annually in Dubai, UAE. Dr. Shubair is a founding member of the IEEE UAE Signal Processing and Communications Joint Societies Chapter. He currently acts as Symposium Vice-Chair of the 2008 IEEE Canadian Conference on Electrical and Computer Engineering: Communications and Networking Symposium. Dr. Shubair is Editor for the *Journal of Communications, Academy Publisher*. He is listed in *Who's Who in Electromagnetics* and in several editions of *Who's Who in Science and Engineering*.

# The Design of a TRL Calibration Kit for Microstrip and its use for Measurement and Modeling of Active and Passive RF Components

D. Elsherbeni, L. Jordan, E. Hutchcraft, D. Kajfez, and R. K. Gordon

Department of Electrical Engineering  
The University of Mississippi, University, MS 38677-1848, USA  
daelsher@olemiss.edu, ltjordan@olemiss.edu, eeweh@olemiss.edu, eedarko@olemiss.edu,  
eegordon@olemiss.edu

**Abstract** – Microstrip TRL standards for the Thru, Reflect, and Line are designed, fabricated and tested. A split fixture design is built which contains microstrip to coaxial transitions to aid in measuring non-coaxial devices. With this microstrip TRL calibration kit, accurate and repeatable calibrations over a broad frequency range will be verified. This kit is used to measure and model several components such as a capacitor, an inductor, a transistor, and a radio frequency identification tag. The TRL kit allows the reference planes to be set at the device under test and allows novel modeling of the device's intrinsic and extrinsic parameters through optimization of a network to minimize the difference between S-Parameter measurements taken with the TRL kit and device under test and the proposed model.

## I. REQUIREMENTS FOR TRL STANDARDS

The Short-Open-Load-Thru (SOLT) technique is the traditional calibration method used primarily for coaxial applications. In non-coaxial measurements it is more difficult to build impedance standards that are easily characterized. In microstrip, short circuits are inductive, open circuits radiate energy, and it is difficult to build a high quality purely resistive load. Because of these limitations, studies are done on an alternative method for calibration in non-coaxial environments called Thru-Reflect-Line (TRL) that uses simple, realizable standards. This form of calibration can provide more accurate results than the SOLT method even for coaxial applications. The standard SOLT calibration depends on a set of four well-defined impedance standards (open, short, load, thru), but TRL only relies on lines with a consistent characteristic impedance and a reflect that doesn't have to be well defined like a short or an open circuit. Because of this, TRL calibration standards are easier to manufacture than SOLT standards, especially for in-fixture environments. In this paper, the TRL calibration technique is studied for use in testing packaged transistors and passive surface mount components that are typically used on microstrip. The

device under test must be physically connected to the network analyzer by some kind of transition. As a solution, we propose the development of a TRL calibration kit in which only microstrip TRL standards are used for S-parameter characterization of non-coaxial devices. The hookup of the three different standards along with the associated error boxes is shown in Appendix I [1]. The thru is obtained by placing the two fixture halves together in between the set of coaxial connectors. The thru standard is of zero length. Its characteristics are perfect transmission ( $S_{21}=1$  with zero degree phase shift) and reflection coefficients equal to zero. At all frequencies, an ideal short is defined to have perfect reflection in which its reflection coefficient,  $S_{11}$ , is equal to negative one. The line length is designed to have an electrical length equivalent to  $90^\circ$  phase shift at the center frequency of the entire range. It is characterized by perfect transmission when  $|S_{21}| = 1$ . The major limitation of the TRL technique is the limited bandwidth of the Line standards. A single line is only usable over a maximum of an 8:1 frequency range, so multiple lines are required for broad frequency coverage. At low frequencies, Line standards can become too long for practical use [2]. In the work reported here, three different line lengths, one to cover the low frequency range of 200 MHz to 1.8 GHz, a second to cover 700 MHz to 6.3 GHz, and a third to cover 1.4 to 12.6 GHz. These three lines would allow measurement up through the X-band. Since the design was to be used in a 50 Ohm system, widths of the 50 ohm microstrip lines were found using a linecalc program with the knowledge of the dielectric constant and the thickness of the board. The board substrate used in the design of the microstrip TRL calibration kit is the RT Duroid 6002 High Frequency Laminate from Rogers Corporation with dielectric constant of 2.94 and thickness of 30 mils [3]. An example of using a linecalc program available in Agilent's Advanced Design System (ADS) is shown in Fig. 1 below [4].

The width and relative dielectric constant are then used in a set of formulas shown below to obtain different frequency Line lengths. A block for each Line length was first designed then built. These blocks represent the Line

standard in the TRL calibration technique. Twenty percent of the center frequency defines the lower part of the operating range. One-hundred and eighty percent of the center frequency defines the upper part of the range. For example, consider the low frequency range:

$$\begin{aligned} &\text{center frequency, } f_c = 1 \text{ GHz} \\ &20\% \text{ of } 1 < f_c < 180\% \text{ of } 1 \\ &200 \text{ MHz} < f_c < 1.8 \text{ GHz} . \end{aligned}$$

From the above calculations, a low frequency Line insert piece was designed to cover 200 MHz to 1.8 GHz. From LineCalc, width,  $w = 1.935 \text{ mm}$  and relative effective dielectric constant,  $\epsilon_{\text{reff}} = 2.383$  were obtained when the center frequency was set at 1 GHz. From these values and the following set of equations we obtain the phase velocity,  $v_p$  and wavelength,  $\lambda_g$ . The line length is designed to be one quarter the wavelength at the center frequency as shown below,

$$\begin{aligned} v_p &= \frac{c}{\sqrt{\epsilon_{\text{reff}}}} = \frac{3 \times 10^8}{\sqrt{2.383}} \\ &= 1.94 \times 10^8 \text{ m/s} \end{aligned} \tag{1}$$

$$\begin{aligned} \lambda_g &= \frac{v_p}{f} = \frac{1.94 \times 10^8}{1 \times 10^9} \\ &= 0.194 \text{ m} \end{aligned} \tag{2}$$

lower frequency Line length,

$$\begin{aligned} L &= \left( \frac{90}{360} \right) \lambda_g = \frac{\lambda_g}{4} = \frac{0.194}{4} \\ &= 0.0485 \text{ m} = 4.85 \text{ cm} . \end{aligned} \tag{3}$$

Following the same calculating procedure as above, with center frequency at 3.5 GHz, a resulting span from 700 MHz to 6.3 GHz characterized a second Line length insert piece for the middle frequency range of the designed microstrip TRL calibration kit. Finally, a third line was designed with center frequency at 7 GHz and a range covering 1.4 GHz to 12.6 GHz. Typically, the VNA has calibration kits for most coaxial devices and the different standards for these kits have been defined within the network analyzer. The VNA also typically has the TRL calibration algorithm embedded into the system software. However, for the TRL algorithm to work, it is necessary to create and define each standard that will be used with the designed test fixture. The user-created standard kit includes the electrical characteristics of each calibration standard. Figure 2 shows the Standard Definition window for a VNA Calibration Kit Program used for out Agilent 8510C VNA [5].

Entered are the Short, Line, and Thru delay values in picoseconds. These values were obtained through a one-port SOLT calibration using the Agilent coaxial calibration kit. Each standard was placed on port 1 of the network analyzer, and with the port extension feature, the length of each standard was recorded and entered into the Offset Delay column in the Standards Definition window.

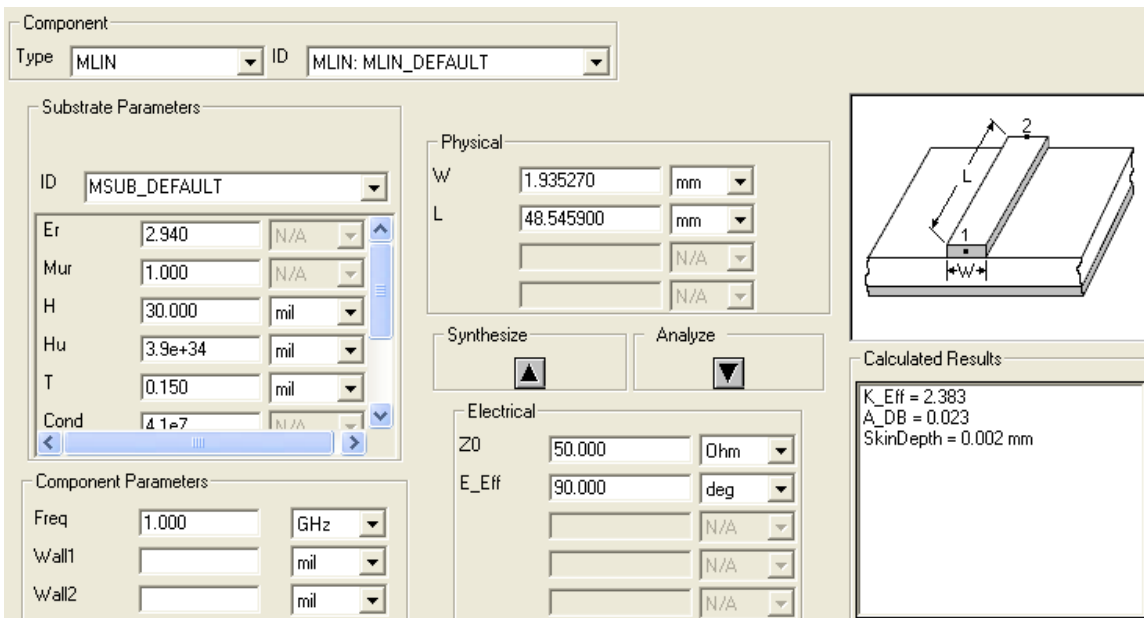


Fig. 1. A typical LineCalc window.

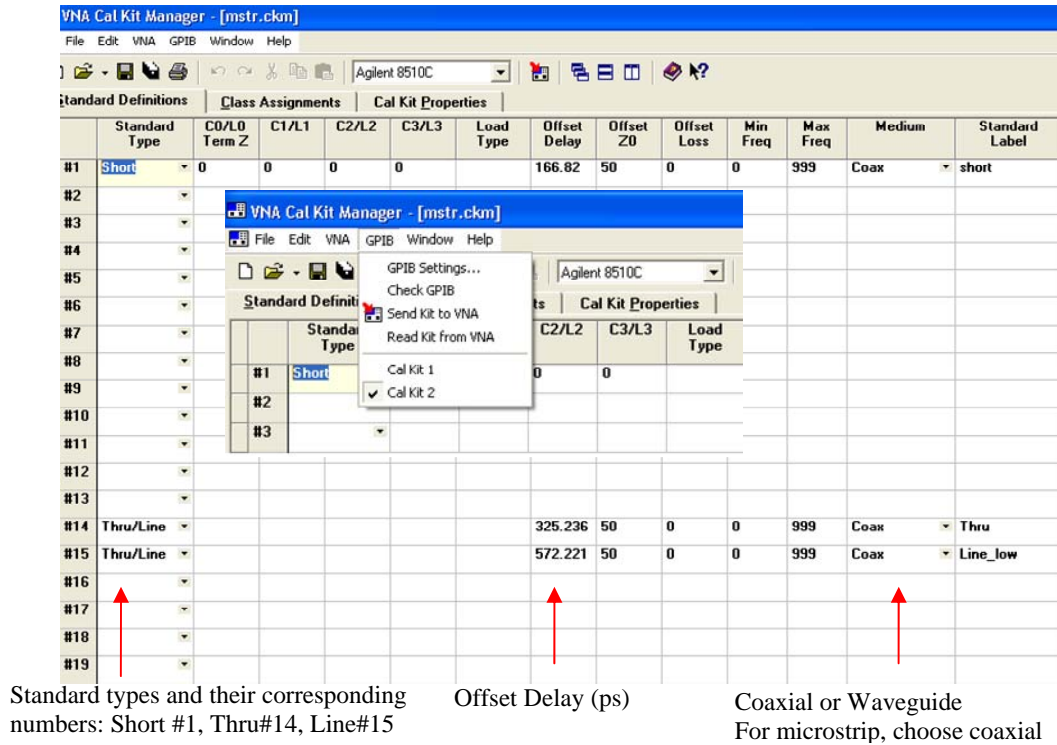


Fig. 2. User-created Cal Kit standards definitions.

## II. SPLIT FIXTURE DESIGN

For non-coaxial device measurement, a test fixture is needed to mount the device under test. The split design gives the ability to mount calibration standards of different lengths and non-coaxial devices to the center block which is inserted between the fixture halves. The test fixtures that represent the TRL calibration standards were made out of brass blocks and built using a CNC machine. Each block part was carefully modeled in the mechanical design software Pro-Engineer [6]. Figure 3 is a layout in Pro-Engineer showing dimensions in cm for the Reflect block, one of the required calibration standards of the TRL technique. Figure 3 simply demonstrates how each part was carefully designed and laid out for machining.

Figure 4 shows the assembled split fixture (zero length Thru standard) with the two block halves and a Teflon bridge on top. The purpose of the bridge is to insure connection between the device under test and the connecting blocks. The connection could be done with solder, but the Teflon bridge will allow connection without having to solder the leads of the device under test to the microstrip fixture leaving the calibration kit general to the various sizes and shapes of RF non-coaxial components. In the middle of the Teflon piece are two fingers. These fingers press down on the leads of the device under test so that there is good electrical

connection between the device under test and microstrip line. The Teflon piece is connected to the block with screws, and these screws provide the pressure to the fingers which are lined with copper strips for conductivity. It should be pointed out that the Teflon could, in fact, introduce discontinuities, but since its dielectric constant is relatively low, and only a small portion touched the microstrip lines, it did not affect the measurement results in the frequency range of interest.

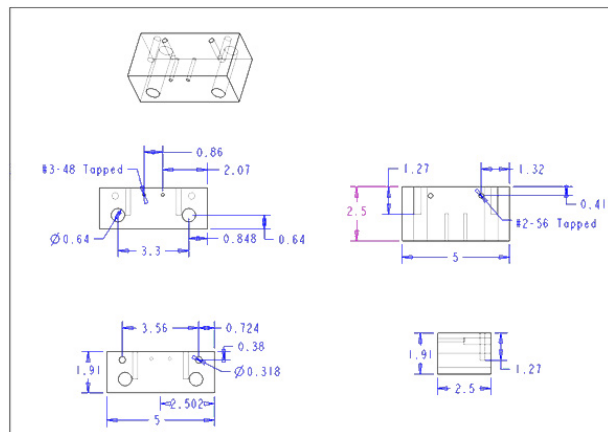


Fig. 3. Pro-Engineer layout for the Reflect standard (2.5 x 5.0 cm block half).

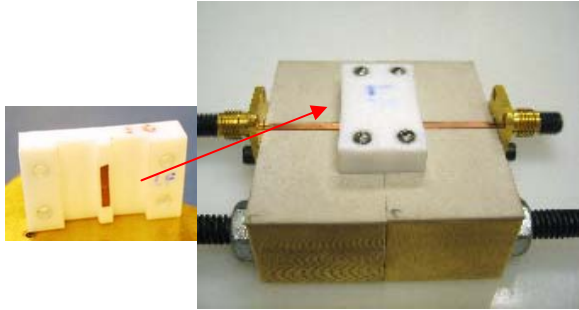


Fig. 4. Thru standard with Teflon bridge connection.

As can be seen from the Pro Engineer design, the individual blocks have assembling bolts and precision guiding pins, for accurate alignment of microstrip sections. Figure 5 points out the low frequency Line standard inserted in between the two halves with guiding pins and bolt holes to ensure a tight fit. Any high Reflect standard can be used and must be the same on port 1 and port 2 of the network analyzer.

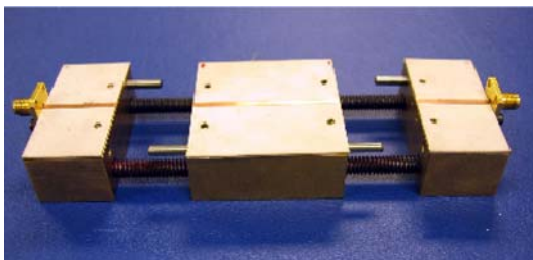


Fig. 5. Low frequency Line insert with guiding pins and bolts.

In this work, the Reflect standard is a 5 x 2.5 cm block with a wire soldered at the end of the microstrip line resembling a short circuit shown in Fig. 6.

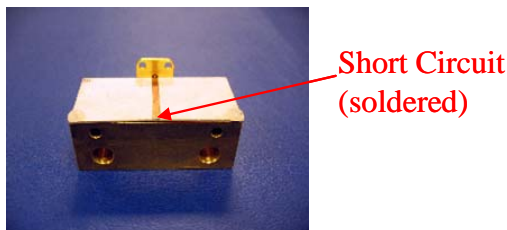


Fig. 6. Reflect standard (short circuit).

### III. CHECKING THE CALIBRATION PERFORMANCE

This section below will check the performance of each TRL standard for the low frequency range. The reflection coefficient of the Reflect standard is supposed to be negative one at all frequencies. The objective is to

obtain total reflection,  $20 \log(1)$  equals 0 dB. Figure 7 is a plot of the reflection coefficient in log magnitude verifying the Reflect standard requirement. Figure 8 verifies both the Line and Thru standards. The low frequency Line length is designed to have an electrical length of  $90^\circ$  at center frequency of 1 GHz.

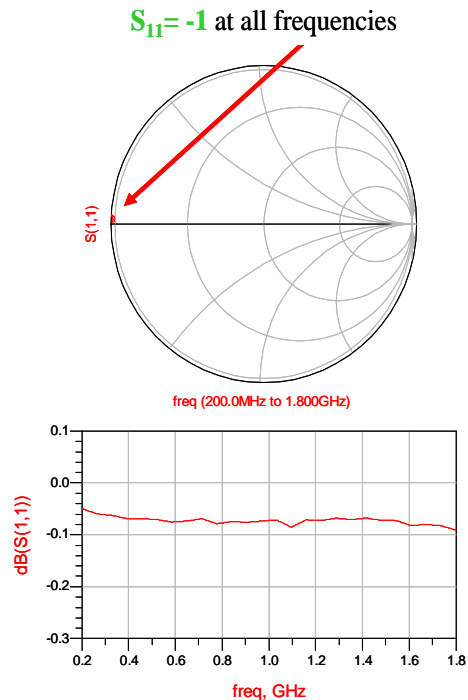


Fig. 7. Short standard verification.

The marker “m1” denotes the center frequency point. Note that the red marker on this transmission plot, is, as designed, nearly one quarter wavelength away from the Thru. The red curve is the measured range between 200 MHz and 1.8 GHz. Notice that the Line length varies as a function of frequency. The phase due to the electrical length of the line is different at each frequency and therefore the Line length is different at each frequency. The thru will remain constant at zero degrees.

The TRL calibration kit has been tested for the lower frequency band, and the results are excellent. The results are similar for the middle and upper bands. However, some difficulties arose with the performance of the Reflect standard when calibrating with the high frequency set of blocks which results in the limitation of its use to only 9 GHz. Further studies will be done regarding the Reflect block to determine the reason for the incorrect reflection coefficient at a few specific frequencies to increase the overall working frequency range. A new Reflect board with grounding vias is being designed, and it is expected that this should cause the short circuit to appear less inductive. It is hoped that this will remedy the problem and allow the use of the high frequency block to 12.6 GHz.

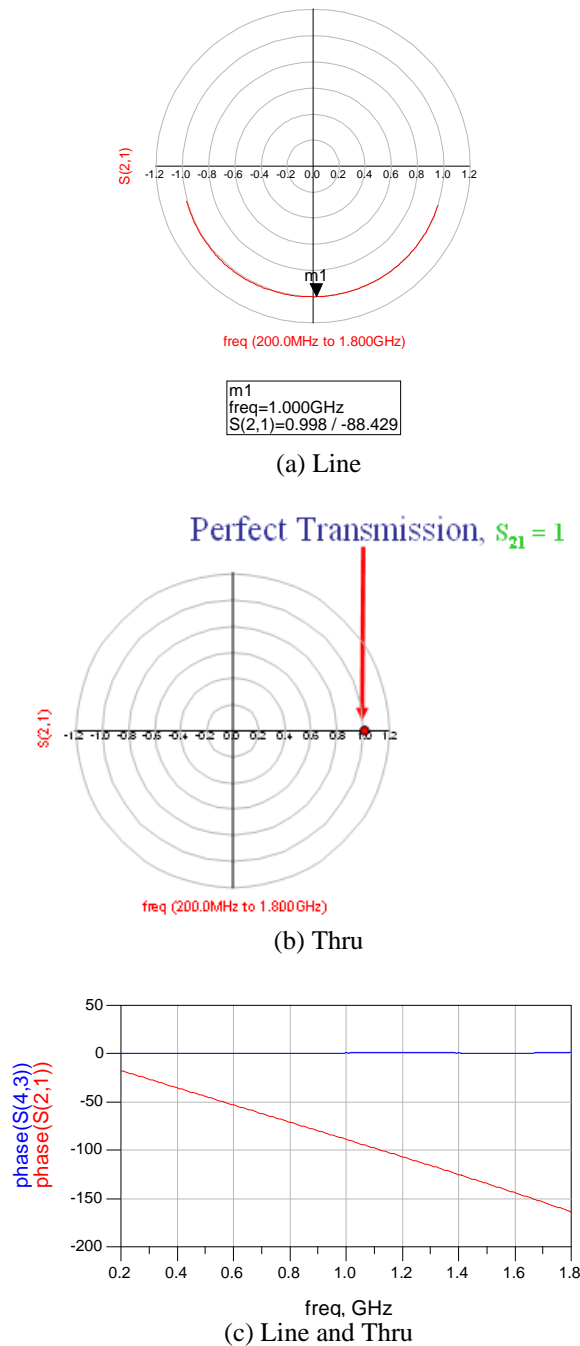


Fig. 8. (a)Line standard verification, (b)Thru standard verification, and (c) Line and Thru- phase in degrees.

**IV. MEASUREMENT AND CHARACTERIZATION OF ACTIVE AND PASSIVE RF COMPONENTS**

Calibration is necessary before taking measurements. In order to remove systematic errors the TRL calibration technique is used to setup the reference plane at the device under test. Because this TRL kit allows the reference plane to easily be set at the test device's

terminals, models can easily be obtained through optimization of a model and the measured S-parameters. This section will show several measurements of non-coaxial devices such as a capacitor, an inductor, and a transistor all setup and tested using the designed microstrip TRL calibration kit. All of these were measured using the low or middle frequency ranges because they were being used in an amplifier that was being designed for the WLAN range (2.4 GHz). In addition, a radio frequency identification (RFID) tag was also measured. For all of these components, the S-parameters were measured, and these S-parameters could be used to obtain a model for the device, or the S-parameters could be used directly in a simulation.

**Capacitor:**

The designed microstrip TRL calibration kit can be used to assess the performance of capacitors. The capacitor used in measurement was the C06CF5R1B9U high Q Multilayer Capacitor from Dielectric laboratories, Inc. [7]. This capacitor was measured using the designed low frequency microstrip TRL calibration kit over the range of 200 MHz to 1.8 GHz. To determine the measured capacitor's value, the capacitor was modeled using an equivalent circuit and optimized in ADS. The equivalent circuit used was a series RLC circuit with another resistor in parallel with the capacitance. A simulation was run to best fit the measured data to an the capacitor circuit model. Results conclude that the optimization yields that the value of the capacitor is 5.41489 pF. It would be hard to accurately measure this capacitance this accurately with an LCR meter, but with the optimization in ADS, the result is to this number of significant digits over the frequency range measured. The data sheet provided by Dielectric Laboratories, Inc lists nominal value for the capacitor is 5.1 pF. Measured results show a good match between the data provided. Figure 9 compares the transmission results of the measured capacitor to the data provided by the manufacturing company in log magnitude format. The measured S-parameters match the model well; the two curves shown illustrated the difference between the company's nominal capacitor value and the particular one that was measured. At the higher frequencies, both are acting as short circuits, as a capacitor should.

**Inductor:**

A hand wound inductor coil that was to be used as an RF choke was made from 28 gauge enamel coated copper wire and soldered to the microstrip boards for measurement. A 3-turn inductor was measured using the low frequency calibration kit over the frequency range of 200 MHz to 1.6 GHz. The picture in Fig. 10 below shows the 3-turn inductor soldered to the 50 Ohm microstrip transmission lines placed on the designed test fixture

which is connected to ports one and two of the network analyzer.

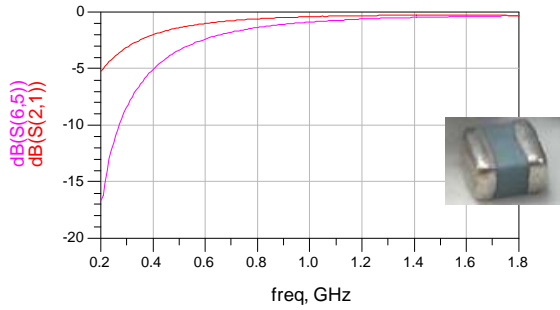


Fig. 9.  $S_{21}$  capacitor results. Red = Measured capacitance (optimized to  $C=5.41489$  pF using an equivalent circuit with parasitics using ADS), Magenta = Typical Manufacturer's Characteristics ( $C=5.1$  pF).

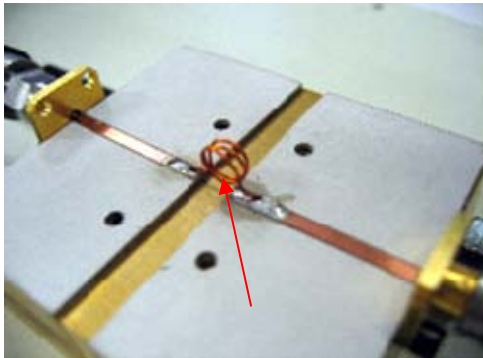


Fig. 10. 3-Turn inductor.

Results for the 3-turn inductor are shown in Fig. 11 in Smith chart form. This inductor was merely fabricated to have impedance much higher than 50 Ohms so that the RF could be blocked in the desired frequency range since it was to be used as an RF choke in an amplifier design. So, to insure the device was acting accordingly, the device was modeled in ADS as shown in Fig. 12 below. This model includes parasitic elements as well as the inductor itself. Optimization enables us to determine the value of the inductor itself. As can be seen from Fig. 12, the inductance value obtained from the equivalent circuit was in good agreement with the design as the equivalent circuit's inductance was approximately 48 nH, which by itself (not including parasitic) has an impedance that is at least ten times larger than 50 Ohms in the amplifier's frequency range of 1.9-2.6 GHz. The resulting equivalent circuit's S-parameter results are shown in blue on Fig. 11. Thus, the equivalent circuit provides an excellent model for the inductor over the desired frequency range.

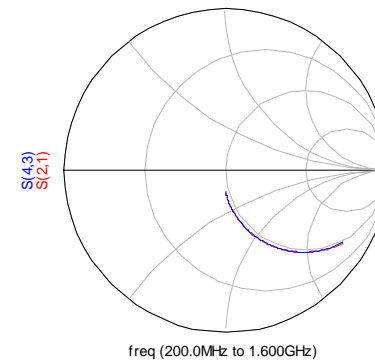
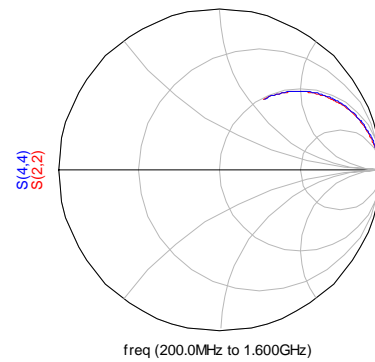
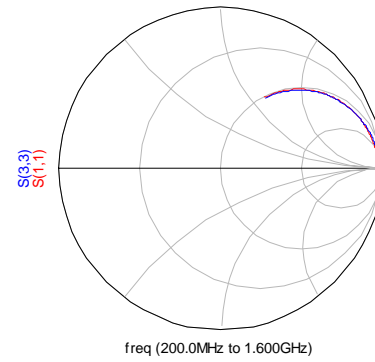
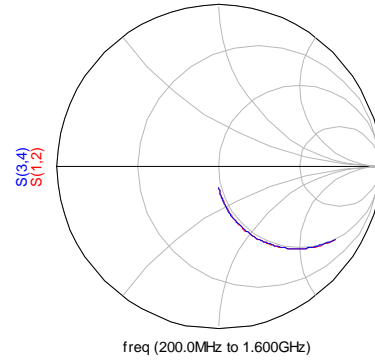


Fig. 11. 3-Turn inductor results. Red = measured, Blue = simulated model including parasitics.



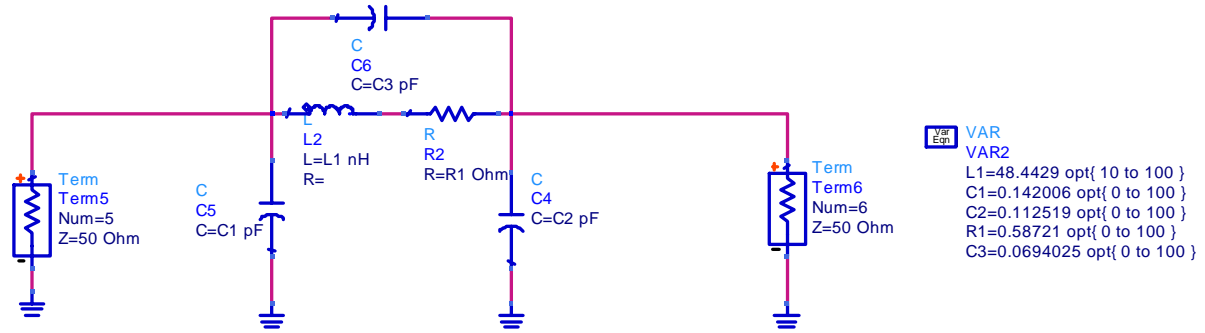


Fig. 12. ADS schematic of the equivalent circuit of an inductor.

**Radio Frequency Identification (RFID) Chip:**

An RFID tag is an object that can be attached to another object for the purpose of identification using radio waves. Chip-based RFID tags contain silicon chips and antennas [8]. A sample commercial RFID tag with an integrated chip was set up for impedance measurement. The setup is depicted in Fig. 13. The terminals shown in the Figure were connected to port 1 and port 2 of the network analyzer. The fingers on the Teflon bridge pressed the edges of the chip down for contact during measurement. Characteristics were not provided by the manufacturer, so the TRL calibration kit enables us to determine the input impedance of the tag.

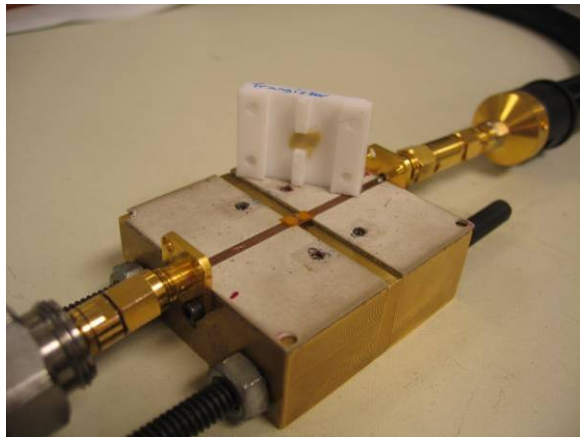


Fig. 13. RFID measurement setup.

The S-parameters were measured. A picture of the RFID chip is shown in Fig. 14 along with  $S_{11}$  data plotted on a Smith chart.

From the measured S-parameters, the RFID tag’s input impedance could be obtained.

**Transistor:**

A silicon carbide (SiC) transistor from the manufacturing company CREE [9] is an active component used in an RF amplifier design. This small

transistor has a large power output capability. So before actual measurements of the amplifier, the designed microstrip TRL calibration kit was used for testing the D.C. biasing and general behavior of the CREE transistor. For the CREE transistor a special block was made so that the bottom of the transistor can sit down into the block allowing the leads to reach the microstrip board. This block is shown in Fig. 15. The transistor leads were soldered to the board. A Teflon piece was pressed over the transistor for electrical conductivity during measurements.

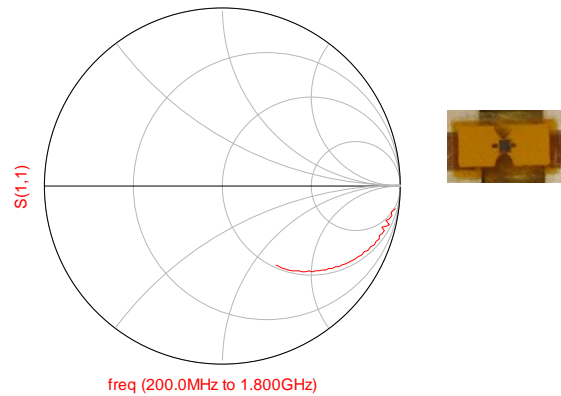


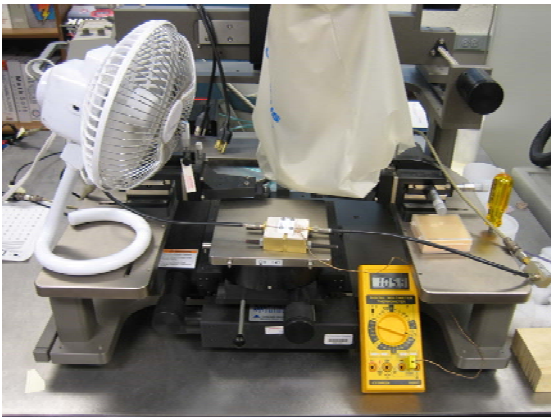
Fig. 14. Reflection,  $S_{11}$ , results and a picture of the measured RFID chip.



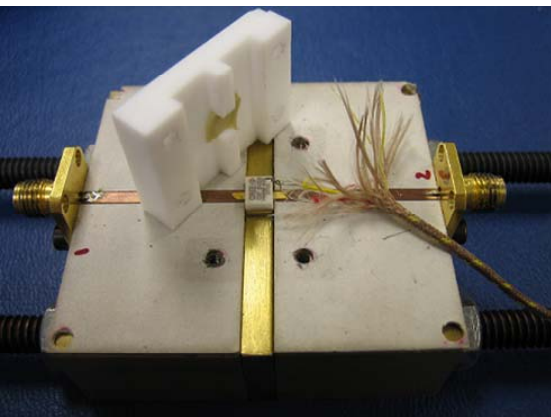
Fig. 15. The CREE transistor sits on this block.

The transistor was measured using the middle frequency range designed microstrip TRL calibration kit but only from 100 MHz to 4 GHz. A thermocouple wire

was used to monitor the temperature as different Q-points were tested. The tip of the thermocouple wire was pressing against the top of the transistor. From above, pressure was applied by the Teflon bridge and a small piece of silicon. Figure 16 contains two pictures. The first picture, (a), is the measurement setup of the CREE SiC transistor as its temperature readings are being taken while the fan cools it off. The second picture labeled (b) is a close up of the thermocouple wire and Teflon bridge that will press over the mounted transistor during measurements.



(a)



(b)

Fig. 16. (a) Measurement of CREE SiC transistor (b) Teflon piece with silicon padding and thermocouple wire for monitoring the temperature.

The fan served as a cooling mechanism as the temperatures started to reach the maximum operating temperature of 125 °C. As the temperature increased, a clamp was added to create better contact with the copper block and dissipate more heat. Before applying any current, the temperature read 22.6 °C from the thermocouple meter. Full two-port measurements were performed at the biasing point of  $V_{GS} = -6.87$  V,  $V_{DS} = 48.0$  V, and  $I_{DS} = 249.4$  mA at 139 degrees Celsius. This

was near the manufacturer's provided data which was taken at  $V_{GS} = -6.0$  V,  $V_{DS} = 48.0$  V, and  $I_{DS} = 250.0$  mA. Figure 17 illustrates a comparison between the  $S_{21}$  (transistor gain) of the closest fit curve that was obtained during measurement from the set of different Q-points listed in Table 1 to CREE's manufacturer's data. As seen from the  $S_{21}$  data, the results are close, and differences are likely due to device variability. The full two-port data was also used in building a circuit model for the SiC transistor.

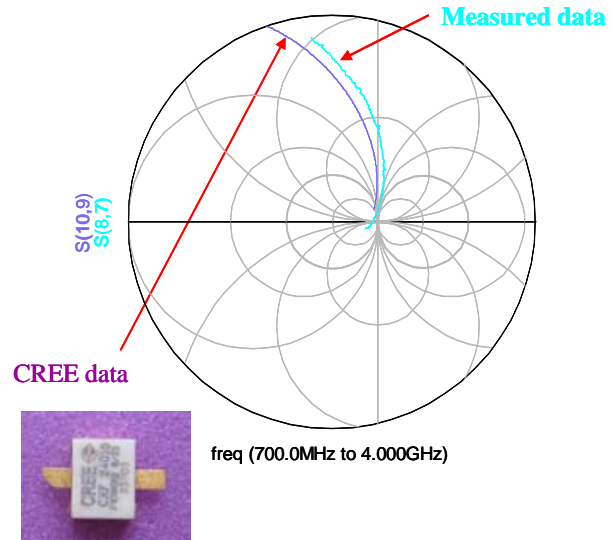


Fig. 17. Gain of measured S (8, 7) versus CREE's manufacturer's curve S (10, 9).

## V. CONCLUSION

In this work, the design and fabrication of a microstrip TRL calibration kit has been developed, discussed, and proven to work for RF device measurements. The Thru, Reflect, and Line standards were verified through the TRL calibration process. The microstrip split fixture design gives ability to mount calibration devices of different lengths and packaged transistors to the center block. Several non-coaxial components such as a capacitor, an inductor, a transistor, and an RFID tag were mounted onto and measured with this split fixture design calibration kit. One key importance of this kit is that it allows extraction of both intrinsic and parasitic device parameters through the novel optimization and circuit simulation shown throughout this paper.

## ACKNOWLEDGEMENTS

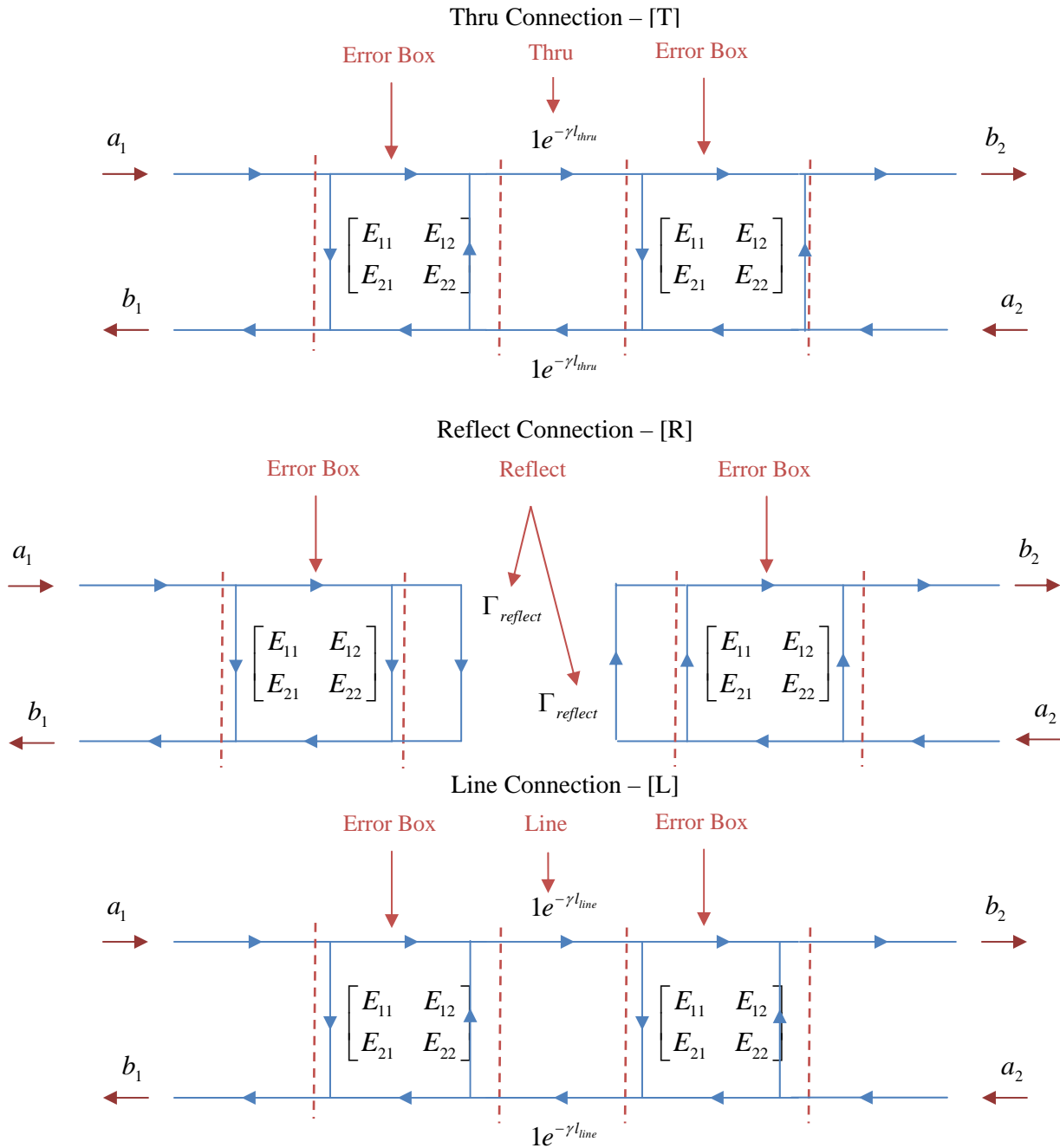
The authors would like to thank the U.S. Army Space and Missile Defense Command for sponsoring this research as part of the Radar Power Technology program.

APPENDIX I

**TRL Technique**

The signal flow graphs of the TRL calibration technique are shown in the figures below. The Error Boxes are the errors in the measurement that are calibrated out after the TRL calibration is performed. These errors include errors associated with directivity, source and load matching, etc. By connecting the Thru, the Reflect, and the Line standards, enough measurements are made to find the unknowns that are

associated with the error boxes, the impedance of the transmission lines, and the reflection coefficient of the Reflect standard. Finally, the reference planes can be set at the DUT (where the Thru, Reflect, and Line are placed in the pictures). Notice in the Thru connection, that  $l_{thru}$  is typically zero, so the resulting terms are just 1. If a nonzero length thru is used, the technique is often called LRL due to the fact that two lines are being used in the calibration. Details of the resulting equations can be found in [1].



## REFERENCES

- [1] David Pozar, *Microwave Engineering*, 3<sup>rd</sup> Ed., John Wiley & Sons, 2005, pp 193-196.
- [2] Agilent Technologies PN 8720-2, In-fixture Microstrip Device Measurements Using TRL\* Calibration.
- [3] Rogers Corporation, Web: [www.rogerscorporation.com](http://www.rogerscorporation.com).
- [4] Agilent Technologies Advanced Design System (ADS) (2004A, Oct.). Web: <http://www.agilent.com>
- [5] VNA Cal Kit Manager, Version 2.10, 1997-2002, Barry A. Brown.
- [6] Pro-Engineer Wildfire 2.0 Design and Modeling Software, [www.PTC.com](http://www.PTC.com).
- [7] Dielectric Laboratories, Web: [www.dilabs.com/pdfs/pages\\_10-11\\_C06.pdf](http://www.dilabs.com/pdfs/pages_10-11_C06.pdf).
- [8] Online: <http://en.wikipedia.org/wiki/RFID>.
- [9] CREE Part Number CRF24010, Web: [http://www.cree.com/products/pdf/CRF24010-Rev1\\_4.pdf](http://www.cree.com/products/pdf/CRF24010-Rev1_4.pdf)



**Dalia Elsherbeni** received an undergraduate degree in electrical engineering from The University of Mississippi in 2004. In December 2006 she received a Master's degree in electromagnetics at the University of Mississippi. Her thesis entitled, "Measurement of Non-Coaxial RF

Components Using the TRL Calibration Technique" was written under her advisor Dr. Elliott Hutchcraft. Ms. Elsherbeni is a member of IEEE. She was the recipient of the Mississippi Academy of Sciences (MAS) Third Place Oral Presentation Award in Vicksburg, MS in 2006. Since 2004 she has been employed full time by Radiance Technologies working in the technology and development department.



**Lisa Jordan Ewell** graduated from the University of Mississippi in May of 2004 with a Bachelor of Science in Electrical Engineering. She continued her education at Ole Miss and pursued a master's degree in electrical engineering with an emphasis in electromagnetics. She received her Master of Science in December, 2006.

While studying for her master's degree, she first worked as a research assistant for Dr. Elliott Hutchcraft and then began working for Radiance Technologies in Oxford, MS. She currently lives in Bangkok, Thailand, teaching English and Java programming at Mahidol Wittayanusorn School, a Thai math and science high school.



**W. Elliott Hutchcraft** was born in Lexington, Kentucky on April 29, 1973. He earned his B.S. in electrical engineering at the University of Mississippi, Oxford, MS in 1996, his M.S. in electrical engineering at the University of Mississippi, Oxford, MS in 1998 and his Ph. D. in electrical engineering at the University of Mississippi, Oxford, MS in 2003. He is an Assistant Professor in the Department of Electrical Engineering at the University of Mississippi in Oxford, Mississippi. Dr. Hutchcraft is a member of Eta Kappa Nu, Sigma Xi, IEEE, Tau Beta Pi, Phi Kappa Phi, and ARFTG.



**Richard K. Gordon** was born in Birmingham, Alabama on November 26, 1959. He earned his B.S. in physics at Birmingham Southern College, Birmingham, AL in 1983, his M.S. in mathematics at the University of Illinois, Urbana, IL in 1986 and his Ph. D. in electrical engineering at the University of Illinois, Urbana, IL in 1990. He is an Associate Professor in the Department of Electrical Engineering at the University of Mississippi in Oxford, Mississippi. Dr. Gordon is a member of Eta Kappa Nu, Phi Beta Kappa, and Tau Beta Pi.



**Darko Kajfez** is Emeritus Professor of Electrical Engineering at the University of Mississippi. He obtained the electrical engineer's degree (Dipl. Ing.) from University of Ljubljana, Slovenia, in 1953, and the PhD degree from U.C. Berkeley in 1967. He co-edited the book *Dielectric Resonators*, and authored the books *Notes On Microwave Circuits*, and *Q Factor*. His research interests include rf and microwave measurement and analysis. He can be reached at [eedarko@olemiss.edu](mailto:eedarko@olemiss.edu).

# A Discrete Random Medium Model for Electromagnetic Wave Interactions with Sea Spray

O. Kilic

Department of Electrical Engineering and Computer Science  
The Catholic University of America, Washington, DC, USA  
Email: kilic@cua.edu

**Abstract** – Understanding the electromagnetic interactions with sea spray is of interest for many applications such as satellite or terrestrial communications, remote sensing systems for surveillance and meteorology. The water droplets in spray are a key factor in the energy transfer between the atmosphere and the ocean, which can help the wind speed retrieval algorithms for the global climate models. Furthermore, the presence of these droplets along the propagation path interferes with both land-based and satellite based remote sensing of the ocean surface. This paper investigates the backscatter response from sea spray, which is modeled as a layer of water droplets over a rough ocean surface. The distorted Born approximation technique is used in conjunction with the analytical wave theory to compute backscattering from the medium.

**Keywords:** remote sensing, scattering, random medium, maritime, ocean, and sea spray.

## I. INTRODUCTION

The presence of water droplets along the propagation path interferes with both land-based and satellite-based remote sensing systems of ocean surface. This can cause adverse effects for weather forecasting, wireless communication and military surveillance systems. Understanding the electromagnetic interactions with sea spray is also important as it plays a critical role in the energy exchange between the surface of the ocean and the atmosphere. This energy transfer helps define the boundary condition for atmospheric and oceanic models and is a key factor for the global and regional climate models. The simulation of the complex interaction of the weather system with the sea is often a major challenge in meteorology.

This paper investigates the backscatter response from a layer of water droplets over a rough ocean surface as shown in Fig. 1. The results of this model have been presented before, [1]. This paper extends the earlier findings, and provides an investigation of cluster effects when the particles get in close proximity to each other.

The water droplets in sea spray are modeled as perfectly conducting spheres of various sizes. The

droplets are randomly distributed inside the layer, that extends from  $z = -d$  to  $z = 0$ . Different probability distribution functions (pdf) are used to generate different scenarios for the size distribution of the droplets inside the medium. It is assumed that the layer extends indefinitely in the  $x$ - $y$  plane. A Lambertian rough surface, with a dielectric constant equal to that of water is assumed in the background. Sea spray is more complicated than the assumptions made in this analysis. Factors such as ligaments from which the droplets are torn, air bubbles and foam at the water-air interface are currently neglected in the backscatter calculations.

## II. THE BACKSCATTER CROSS SECTION

The backscattered fields from the medium are calculated by using the analytical wave theory in conjunction with the distorted Born Approximation. Therefore, the approach is field based and provides more accuracy than power based approaches, such as the radiative transfer. The calculations are valid for a sparse medium; i.e. the fractional volume of the particles is less than the total volume of the medium.

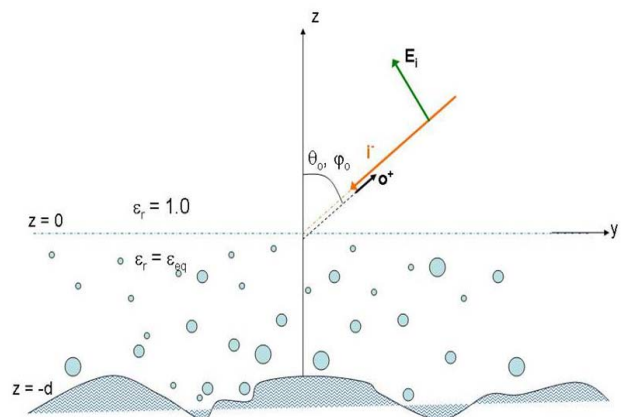


Fig. 1. Sea spray as a layer of discrete random medium.

The Distorted Born Approximation (DBA) is used to calculate the backscattering from the layer of sea spray for both horizontal and vertical polarizations. The approach is similar to the standard Born approximation,

which is a single scatter approximation with the exception of two fundamental assumptions: i) mean field propagating inside the medium is incident on each particle. ii) the scattered fields due to the mean wave are computed using a Green's function for the mean wave.

The scattering from a discrete random medium when illuminated by an electromagnetic pulse has been studied before and applied to vegetation or a layer of rain, [2-4]. The assumptions require the solution for the mean field before the scattered fields can be calculated. The Foldy-Lax technique [5] is used to determine an approximate expression for the mean field inside the medium. This approach assumes a sparse medium; i.e. the fractional volume of the scatterers is very small. The solution for the mean wave suggests that the medium can be characterized by an equivalent dielectric constant [6]. The DBA method allows the mean field to attenuate as it propagates inside the medium, as opposed to the Born Approximation, where the mean field is assumed to be identical to the incident field. As a first step, an equivalent propagation constant in the medium is determined based on the statistical properties and scatterer types in the medium. This equivalent propagation constant defines an equivalent medium that accounts for attenuation as the fields propagate inside. The particles are then embedded inside this equivalent medium for scattering calculations. An application of this approach to vegetation can be found in [7], where all particles inside the medium are assumed identical, and a 2-dimensional problem is studied.

The backscattering coefficient from the medium is given by,

$$\sigma_{pq} = \frac{4\pi r^2 \langle P_{r_p} \rangle}{AP_{i_q}}, \quad p \in \{h_s, v_s\}, \quad q \in \{h_i, v_i\} \quad (1)$$

where  $P_{i_q}$  is the power of the incident plane wave with polarization,  $\hat{q}$  and  $\langle P_{r_p} \rangle$  is the average received power of the  $\hat{p}$  polarized scattered wave, A is the illuminated area on the surface, r is the distance of the radar to this area. The far field is assumed in the calculations; i.e.  $k_0 r \gg 1$ .

The backscattered term for a first order approximation is the sum of three distinct contributions referred to as: direct, direct-reflected, and reflected terms.

$$\sigma_{pq}^o = \sigma_{pqd}^o + \sigma_{pqdr}^o + \sigma_{pqr}^o. \quad (2)$$

The direct term involves volume scattering; i.e., that portion of the backscattered power directly from the particles illuminated.

The direct-reflected term involves a particle and a single bounce from the background before the incident wave is scattered back to the radar. It consists of four components due to the coupling of two possible fields that can arrive back through a single reflection from the ground. The first field directly illuminates the particle and is scattered toward the ground and reaches back through a reflection from the ground. The second field illuminates the particle after being reflected from the ground and is scattered directly back. Both of these fields are in phase and can couple with each other in four different ways in the power calculations. This is an advantage of using a field based approach used in this model, as power based approaches would neglect this coupling effect.

The reflected term arrives back after a double bounce from the background. Figure 2 depicts the different characteristics of these components. The expressions for these terms are given as follows,

$$\text{Direct Term: } \sigma_{pqd}^o = \left[ \sum_{j=1}^{N_{type}} \rho^{(j)} 4\pi \left| f_{pq}^{(j)}(-\hat{i}^-, \hat{i}^-) \right|^2 \right]^* \left[ \frac{1 - e^{-2\text{Im}(\kappa_q + \kappa_p)d}}{2\text{Im}(\kappa_q + \kappa_p)} \right]. \quad (3)$$

$$\text{Direct-Reflected Term: } \sigma_{pqdr}^o = \sum_{j=1}^{N_{type}} \rho^{(j)} \sigma_{qqdr}^{(j)} d \left| \Gamma_{sq} \right|^2, \quad (4a)$$

$$\sigma_{qqdr}^{(j)} = \sigma_{pqdr(1)}^{(j)} + \sigma_{pqdr(2)}^{(j)} + \sigma_{pqdr(3)}^{(j)}, \quad (4b)$$

$$\sigma_{pqdr(1)}^{(j)} = 4\pi \left| f_{qq}^{(j)}(-\hat{i}^+, \hat{i}^-) \right|^2, \quad (4c)$$

$$\sigma_{pqdr(2)}^{(j)} = 4\pi \left| f_{qq}^{(j)}(-\hat{i}^-, \hat{i}^+) \right|^2, \quad (4d)$$

$$\sigma_{pqdr(3)}^{(j)} = 8\pi \text{Re} \left\{ f_{qq}^{(j)}(-\hat{i}^+, \hat{i}^-) f_{qq}^{(j)*}(-\hat{i}^-, \hat{i}^+) \right\}. \quad (4e)$$

$$\text{Reflected Term: } \sigma_{pqr}^o = \left[ \sum_{j=1}^{N_{type}} \rho^{(j)} 4\pi \left| f_{pq}^{(j)}(-\hat{i}^+, \hat{i}^+) \right|^2 \right]^* \left[ \left| \Gamma_{gp} \right|^2 \left| \Gamma_{gq} \right|^2 e^{-2\text{Im}(\kappa_q + \kappa_p)d} * \frac{1 - e^{-2\text{Im}(\kappa_q + \kappa_p)d}}{2\text{Im}(\kappa_q + \kappa_p)} \right]. \quad (5)$$

In the expressions above, the total number of different spheres inside the medium is denoted by  $N_{type}$ , and  $j$  is an index over the different types. The density of type  $j$  in the medium is denoted by  $\rho^{(j)}$ . The scattering amplitudes for a q-polarized incident wave that is scattered back with p-polarization are denoted by  $f_{pq}(\hat{o}, \hat{i})$ , where the first unit vector in the argument ( $\hat{o}$ ) refers to the outgoing (i.e. scattered) field direction and the second unit vector ( $\hat{i}$ ) corresponds to the incoming (or incident) field direction for a particle. The superscripts + and - over the unit vectors describe the direction of the vector along the z-axis, i.e. the axis normal to the medium boundary. Thus,  $i^-$  indicates incidence along a downward direction (i.e., with a negative component along the z-axis), and corresponds to the direction of the incident mean wave. The mean wave can also be incident on a particle after a reflection from the background, and  $i^+$  refers to that direction, which is upward (i.e. with a positive component along the z-axis). Similarly  $-i^-$  and  $-i^+$  correspond to direction of the scattered fields from the particle for the backscatter case. Figure 3 depicts these vectors.

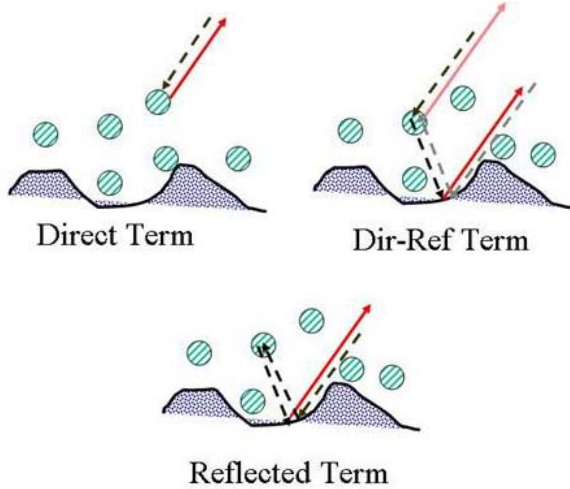


Fig. 2. Backscattering components.

The effective propagation constant inside the medium for different polarization states is represented by  $\kappa_p$  and  $\kappa_q$ , where p and q denote polarization states for the scattered and incident fields, respectively. Both  $\kappa_p$  and  $\kappa_q$  are functions of the equivalent medium characteristics. For an azimuthally symmetric medium, given  $k_z = k_o \cos(\theta_o)$ , the expressions can be written as,

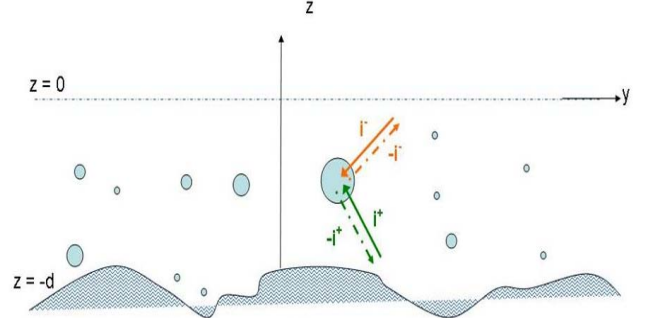


Fig. 3. Unit vector definitions.

$$\kappa_q = k_z + \sum_{j=1}^{N_{type}} \frac{2\pi\rho^{(j)}}{k_z} f_{qq}(\hat{i}^-, \hat{i}^-), \quad (6a)$$

$$q \in \{\hat{h}_i, \hat{v}_i\}$$

$$\kappa_p = k_z + \sum_{j=1}^{N_{type}} \frac{2\pi\rho^{(j)}}{k_z} f_{pp}(-\hat{i}^-, -\hat{i}^-), \quad (6b)$$

$$p \in \{\hat{h}_s, \hat{v}_s\}$$

### III. MODEL RESULTS

Modeling the electromagnetic interactions with sea spray requires knowledge on the characteristics of the droplets. Many approaches for generating a spray generation function exist in literature, [8-10]. The sea spray generation function, commonly denoted as  $dF/dr_0$ , [9] where  $r_0$  is the radius of a droplet at its formation, has units of number of droplets produced per square meter of surface per second per micrometer increment in droplet radius.

There is not a real consensus on how to model this function, which can be affected by various external parameters such as the wind speed, liquid surface tension, etc. However, a noticeable feature observed of natural sprays is that the droplet size distribution is very broad, and demonstrates a skewed distribution, [11] [12]. A comprehensive discussion on the formation of the droplets is given in [13], and demonstrations of the breakup regimes of liquid jets from ligaments when subjected to flowing gas are provided. The droplet distribution studied in this paper is based on [13] where the pdf for droplet sizes is calculated based on the pdf of ligaments of different lengths. This results in an exponential function as follows,

$$p(d) = e^{-nd/(d_0)} \quad (7)$$

where  $d$  is the droplet diameter in millimeters,  $\langle d_0 \rangle$  is the average ligament size, and the parameter  $n$  is approximately given as  $n \approx 3.5$ . The distribution function is normalized in the calculations such that the integral over the possible sizes is one. Therefore equation (7) is modified as follows,

$$f(d) = \frac{n}{\langle d_0 \rangle} e^{-nd/\langle d_0 \rangle}. \quad (8)$$

#### IV. SIMULATION PARAMETERS AND TEST CASES

Simulations for backscattering are based on a  $1 \text{ m}^3$  water amount distributed in a medium with 2.5 m height. The possible droplet sizes in the medium are discretized to four values; i.e.,  $d = [0.05, 0.25, 0.50, 1.00]$  mm, as shown in Fig. 4 where the solid line describes the pdf in equation (8), and the square markers denote the sampled sizes for the simulations.

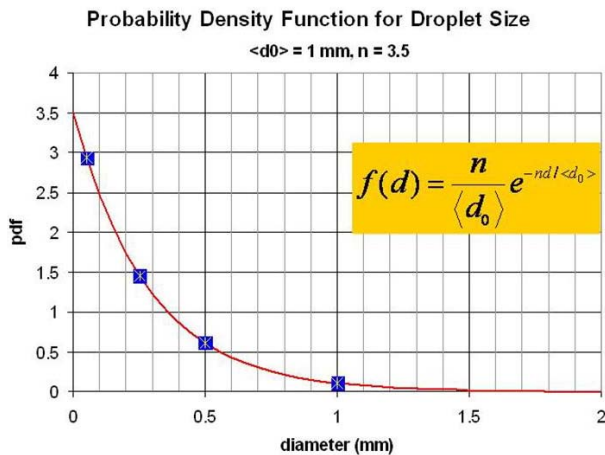


Fig. 4. Droplet size distribution in the simulation based on spray probability density function.

Three different test cases were run using the spray model. Each test case has different scatterer characteristics. The first case represents the  $1 \text{ m}^3$  water using the four types of spheres as identified in Fig. 4. The second case generates the same amount of water by using only the smallest size droplets; i.e., diameter is 0.025 mm. The third case generates the same amount of water by using only the largest size droplets, i.e. diameter is 1.00 mm. The total number for each droplet size for the three cases is then computed such that they add up to the total volume of water; i.e.,  $1 \text{ m}^3$ , in the medium. A Lambertian rough surface is assumed in the background.

The scattering calculations from the spheres are based on expressions in [14], which provide approximations to Mie Theory for different values of  $k_s a$ .

A plane wave at 30 GHz is assumed to be incident on the medium, along a 60 degree incidence angle with respect to the  $z$ -axis. Only hh-polarization is presented in the table, as the results for vv-polarization is identical due to the assumption of perfect spheres for the droplets. There is no cross-polarized component for the same reason. Only the direct and direct-reflected components are calculated for the total return as the reflected term is expected to be smaller than these terms due to the double bounce from the background and longer path traveled inside the medium. The total backscattering cross-section and its different components from the medium are shown in the table in Fig. 5 for the three test cases.

CASE 1	Diam (mm)	# spheres	Sdd (dB)	Sdr (dB)
Type 1	0.05	$2.7389 \times 10^{10}$	-75.3	-172.4
Type 2	0.25	$1.3601 \times 10^{10}$	-36.4	-133.5
Type 3	0.50	$0.5670 \times 10^{10}$	-22.2	-119.3
Type 4	1.00	$9.8526 \times 10^8$	-11.7	-108.8
Sdd = -11.3 dB    Sdr = -108.4 dB    Stot = -11.3 dB    Atten = -108.6 dB				
CASE 2	Diam (mm)	# spheres	Sdd (dB)	Sdr (dB)
Type 1	0.05	$1.5279 \times 10^{13}$	-33.9	-36.3
Sdd = -33.9 dB    Sdr = -36.3 dB    Stot = -31.9 dB    Atten = 0. dB				
CASE 3	Diam (mm)	# spheres	Sdd (dB)	Sdr (dB)
Type 1	1.0	$1.9099 \times 10^9$	-11.7	-207.9
Sdd = -11.7 dB    Sdr = -207.9 dB    Stot = -11.7 dB    Atten = -210.6 dB				

Fig. 5. Backscattering cross-section from the medium for different cases - contribution from different terms. The total water content is  $1 \text{ m}^3$  for all cases.

It is observed from the simulation results that:

- (1) In case 1, strongest contribution to backscatter is from the larger size droplets for the direct term. The dir-ref term is much weaker, as the attenuation is very high (108 db) for this case. Therefore, the total backscattering for this case is due to the direct backscatter from largest particles.
- (2) In case 2, the total backscattering is significantly lower than case 1, although the attenuation is negligible for this case. The smaller particles do not have a high return, although the amount of water is exactly the same as in case 1. The direct-reflected term is comparable to the direct term, but the sum of both direct and dir-ref is still lower than case 1. This is because the particle size is much smaller than the wavelength ( $\lambda = 10 \text{ mm}$  and  $d = 0.05 \text{ mm}$ ), and there is not strong backscattering as the particles behave like point sources.
- (3) Case 3 results are almost identical to case 1, although case 3 has almost twice the number of spheres in case 1. The reason for almost identical backscattering in



both cases although the particle density is doubled in Case 3 can be explained by the increased attenuation due to the doubling of the particle density inside the medium. The incident field penetrates less into the medium for the denser case, resulting in an effectively equal number of particles contributing to backscattering.

**V. CLUSTER EFFECTS**

When the number of particles inside a medium increases, the particles are located much closer to each other. The simulation results for the test cases above do not consider such effects as single scattering is assumed in the calculations for scattered fields. As a matter of fact, increasing the number of spheres in the model further will not increase the backscatter response with this approach. However, multiple scattering or clustering effects becomes important under these circumstances, and can yield higher returns than the predictions based on single scattering.

In order to investigate clustering effects, two cases with identical number and size of water droplets were run. The droplets were paired in one simulation and treated as a scatterer type. The other simulation treated the spheres separately using single scattering as before. For the test case with pairs, the scattering from the pair is calculated using HFSS, which is a full wave analysis tool that uses the finite element method. The full wave approach incorporates multiple scattering effects between the spheres in a pair. The bistatic scattering characteristics of the droplet pair as calculated by HFSS is shown in Fig. 6.

The bistatic scattering result generated by HFSS is incorporated in the model, so that the droplet pair is treated as a scatterer in the medium. The equivalent propagation constant of the medium, the attenuation coefficient and the backscattering terms are computed using the HFSS generated results. All spheres were assumed to have diameters of 1 mm, and the pair test case assumed a center to center separation of 1.2 mm between the droplets. The results of the two cases are shown in Fig. 7. Case 3 corresponds to the single scattering case as before, and case 4 shows the effects of pairing the spheres in the medium. Although we have identical particles and particle densities in both cases, accounting for the close proximity of particles, i.e., cluster effects, enhances the backscatter by almost 4 dB. Two factors contribute to this enhancement effect: (i) the backscattering is more than double for pairs versus single sphere. (ii) The attenuation in the medium is a function of the imaginary part of forward scattering amplitude. HFSS results suggest that there is only 25% increase in the imaginary part of the forward scattering amplitude for a pair of spheres versus a single sphere. Therefore, the medium with droplet pairs have less attenuation compares to the medium with single spheres.

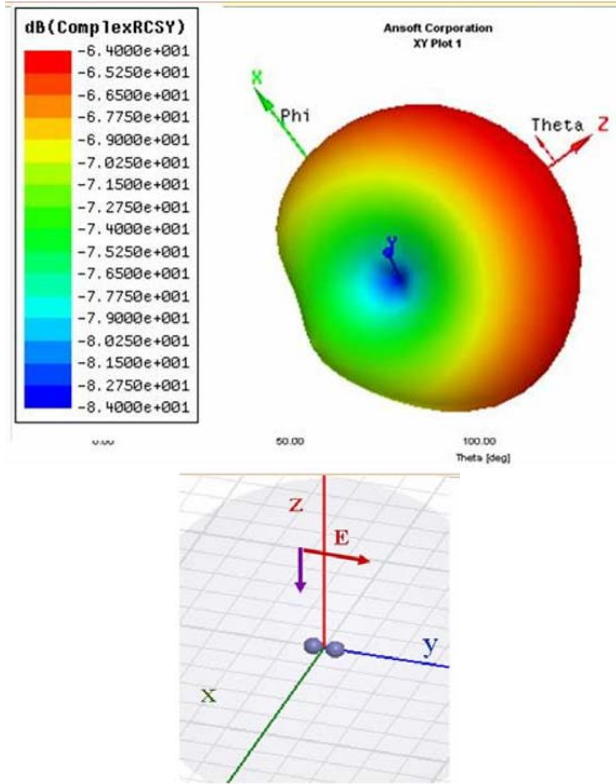


Fig. 6. Bistatic scattering characteristics of the droplet pair - HFSS simulation, 1.2 mm separation between centers of droplets, diameter = 1mm.

CASE 3 (single)	Diam (mm)	# spheres	Sdd (dB)	Sdr (dB)
Type 1	1.0	1.9099x10 <sup>9</sup>	-11.7	-207.9
Sdd = -11.7 dB   Sdr = -207.9 dB   Stot = -11.7 dB   Atten = -210.6 dB				
CASE 4 (pair)	Diam (mm)	# sphere pairs	Sdd (dB)	Sdr (dB)
Type 1	1.0	0.95496x10 <sup>9</sup>	-7.2	-123.6
Sdd = -7.2 dB   Sdr = -128.7 dB   Stot = -7.2 dB   Atten = -128.7 dB				

Fig. 7. Comparison of clustering effects for identical medium, pair versus single droplet.

**VI. CONCLUSIONS**

The interaction of electromagnetic waves with sea spray is a complex phenomenon. Particle size, shape, density and distribution in the medium all play a significant role in modeling these interactions. Furthermore, numerous environmental factors, such as the sea state, wind, and man made influences affect the characteristics of such medium. A statistical electromagnetic model is developed to model this phenomenon as a discrete random medium using single

scattering approach and distorted Born approximation. The water droplets are modeled as perfectly conducting spheres. This assumption simplifies the calculations and provides a first order estimate in understanding the characteristics of backscattering from sea spray. Effects of ligaments which are elongated water particles from which droplets are torn are neglected in the current model. The presence of foam and air bubbles at the interface between air and sea are not included either.

Calculations based on this model demonstrate that the particle size and density determine the attenuation levels in the medium. For a low albedo; i.e., lossy, medium the most significant contribution to backscattering is due to volume scattering; i.e., direct term. The strength of the direct-reflected contribution depends on the attenuation levels in the medium, and can be comparable to volume scattering for lossless media. It is also shown that the close proximity of particles can have a significant effect in terms of backscattering and attenuation. This effect was demonstrated by comparing two cases with identical particle size and density. This implies that cluster effects (i.e., groups of two or more droplets) should be investigated to avoid underestimation of backscattering.

## REFERENCES

- [1] O. Kilic, "Modeling electromagnetic interactions with Sea Spray," *Proc. 2007 ACES Conference*, March 2007, Verona, Italy.
- [2] O. Kilic and R. H. Lang, "Scattering of a pulsed beam by a random medium over ground," *J. of Electromagnetic Waves and Appl.*, vol. 15, no. 4, pp. 481-516, 2001.
- [3] O. Kilic and R. H. Lang, "Pulsed beam scattering from a layer of discrete random medium: A three dimensional model," *Proc. of IEEE AP-S and N. American Radio Sc. Meeting*, 1997.
- [4] O. Kilic and R. H. Lang, "Pulsed beam scattering from a layer of discrete random medium: With applications to spaceborne weather radars," *Proc. of IGARSS*, 1996.
- [5] L. Foldy "The multiple scattering of waves," *Phys. Rev.* vol. 67, no. 3, pp. 107-119, 1945.
- [6] R. H. Lang, "Electromagnetic scattering from sparse distribution of lossy dielectric scatterers," *Radio Sci.*, vol. 16, pp. 15-30, 1981.
- [7] R. H. Lang and J. S. Sidhu, "Electromagnetic backscattering from a layer of vegetation: a discrete approach," *IEEE-GRS*, vol. 21, no. 1, Jan. 1983.
- [8] M. H. Smith, P. M. Park, and I. E. Consterdine, "Marine aerosol concentrations and estimated fluxes over the sea," *Quart. J. Roy. Meteor. Soc.*, vol. 119, pp. 809-824, 1993.
- [9] E. C. Monahan, "The ocean as a source for atmospheric particles," *The Role of Air-Sea*

*Exchange in Geochemical Cycling*, P. Buat-Me'nard, Ed., D. Reidel, pp. 129-163, 1986.

- [10] J. Wu, "Production of spume drops by the wind tearing of wave crests: The search for quantification," *J. Geophys. Res.*, vol. 98, pp. 18 221-18 227, 1993.
- [11] E. L. Andreas, M. Pattison, and S.E. Belcher, "Production rates of sea-spray droplets: clarification and elaboration," *J. Geophys. Res.* vol. 106 (C4), pp. 7157-7161, 2001.
- [12] M. Anguelova and R. P. Barber, "Spume drops produced by the wind tearing of wave crests," *J. Phys. Oceanogr.* vol. 29, pp. 1156-1165, 1999.
- [13] P. Marmottant, and E. Villermaux, "On spray formation," *J. Fluid Mech.* vol. 498, pp. 73-111, 2004.
- [14] Ruck, et. al. *Radar Cross-Section Handbook*, Volume I, Plenum Press.



**OZLEM KILIC** (S '91, M '95, SM '99) received her D.Sc. and M.S. degrees from the George Washington University, Washington, DC in 1996 and 1991, respectively, and B.S. degree from the Bogazici University, Istanbul, Turkey in 1989, all in electrical engineering. She joined the Catholic University of America as an Assistant Professor in the Department of Electrical Engineering and Computer Science in 2005. During her first year at CUA she has lead the effort of establishing the IRIS (Interdisciplinary Remote Imaging and Sensing) Research Center and has been serving as its director. Before joining CUA, she was at the Army Research Laboratory, Adelphi, MD, as an Electronics Engineer (2002-2005) and was the Technical Manager for multiple Small Business Innovative Research (SBIR) programs. Dr. Kilic has over five years of industry experience at COMSAT Laboratories, Clarksburg, MD as a Senior Member of the Technical Staff and a Program Manager with specialization in satellite communications, link modeling and analysis, and modeling, design and test of phased arrays and reflector antennas for satellite communications system.

# Abstraction of Graphics Hardware Through The Use of Modern Interfaces to Increase Performance of Linear Algebra Routines

M. Woolsey, W. E. Hutchcraft, and R. K. Gordon

Department of Electrical Engineering  
University of Mississippi, University, MS 38677, USA  
muwoolse@olemiss.edu, eeweh@olemiss.edu, eegordon@olemiss.edu

**Abstract** – General purpose computation on graphics processing units (GPGPU) is introduced through the application of modern interfaces that abstract graphics hardware. In order to provide an example of these techniques, implementation of an iterative matrix solving algorithm is detailed using two interfaces – Stanford's BrookGPU and Accelerator from Microsoft Research. Performance of the Accelerator implementation is then analyzed.

**Keywords:** Graphics processing unit, GPGPU, and parallel computing.

## I. BACKGROUND

Graphics processing units (GPUs) utilize a parallel pipeline architecture to render graphics onto a 2D screen. A traditional GPU consists of vertex processors, a rasterizer, and pixel processors, as shown in Fig. 1. The vertex processor handles operations such as geometric transformations and per-vertex lighting. The rasterizer converts the vertex data to a 2D array of pixels, and the pixel processors perform texturing and per-pixel lighting operations. Early GPUs employed fixed function pipelines, in which the vertex and pixel processors performed a set of predefined operations. Later GPUs, however, incorporated programmable pipelines, in which programs called shaders can be passed to the vertex and pixel processors. For general purpose applications, most computations are performed within pixel shaders.

The programmable pipelines incorporate many parallel processors that function according to a single instruction, multiple data programming scheme. A pixel operation, for example, can be applied independently to every pixel in a scene. This data-parallel processing capability is the primary draw to general purpose utilization. In addition to parallel processing, graphics hardware also provides a number of useful built-in data types and operations. Because the hardware is optimized to operate on 3D graphics and lighting, data types include multicomponent floating point vectors, and instruction sets contain useful operations such as dot products.

In the latest generation of graphics hardware, both vertex and pixel operations are carried out on processors called unified shaders or simply referred to by their function – stream processors. This architecture results in a more traditional parallel computing environment in which data can be spread over a number of identical processors. The current generation NVidia 8800GT used in this research contains 112 stream processors, with a peak performance of over 500 GFLOPS, and a cost of less than \$250.

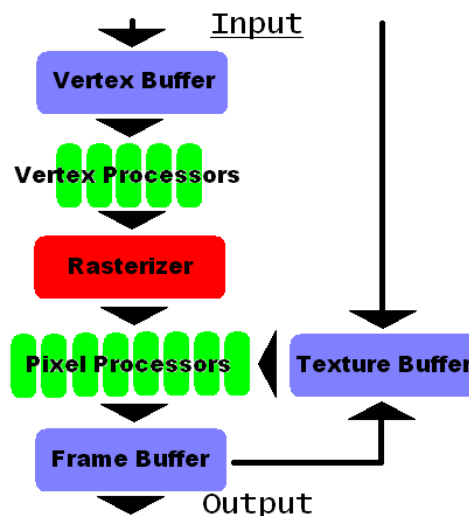


Fig. 1. Traditional GPU architecture.

Processing general purpose data on a GPU, however, has required a creative approach since the hardware is specialized for graphics use. In order for data arrays to be processed within a GPU, they must be stored as textures in graphics memory. In addition, the shader must be explicitly loaded into the GPU memory. By rendering a quadrilateral, data values and pixels can be mapped to one another, and the shader operates on the texture values. The rendering phase may be iterated as necessary, redirecting the shader output as an input texture. The final output may then be read from graphics memory.

## II. GPU PROGRAMMING

Until recently, programming graphics hardware required the knowledge of a shader language. These are languages that can be compiled to run on pixel and vertex processors, such as the OpenGL Shading Language (GLSL), Microsoft's High Level Shader Language (HLSL), and NVidia's C for graphics (Cg). GLSL is a cross platform shading language, due to its OpenGL heritage. HLSL is used within the Windows operating system, as it links to DirectX. Cg, however, is unique in that it can target either OpenGL or DirectX as a compiler option. These languages are common in graphics programming, and they require extensive knowledge about the graphics processes as described in the previous section as well as the selected graphics interface – OpenGL or DirectX. While these languages are a great improvement over hardware-specific assembly language, a higher level of graphics hardware abstraction is necessary for general purpose computation.

The ability to perform general purpose computations on graphics hardware without extensive background in graphics programming is now possible due to the abstraction provided by modern interfaces including Brook for GPUs (BrookGPU) from Stanford University and Accelerator from Microsoft Research. BrookGPU and Accelerator extend C and Microsoft C Sharp (C#) respectively with new syntax and data types allowing data to be transferred to and from the GPU and shaders to be loaded and configured as necessary without complication. An example program will be shown for both, and a brief performance analysis will be conducted on the Accelerator program.

## III. JACOBI ALGORITHM

The Jacobi iterative matrix solving algorithm has been written for both the CPU and GPU as a demonstration. The algorithm computes  $\bar{u}$  from  $\bar{M}\bar{u} = \bar{f}$ , and can be explored by solving for  $u_1$ .

$$M_{11}u_1 + M_{12}u_2 + \dots + M_{1N}u_N = f_1,$$

$$u_1 = \frac{f_1 - M_{12}u_2 - \dots - M_{1N}u_N}{M_{11}}.$$

Let  $\bar{D}$  be a diagonal matrix that holds only the main diagonal of  $\bar{M}$ . The above expression may be written as,

$$u_1 = \frac{f_1 - (M_{11} - D_{11})u_1 - \dots - (M_{1N} - D_{1N})u_N}{D_{11}}.$$

The RHS of this equation can be used to update the LHS

value, which results in the following iterative matrix equation.

$$\bar{u}^{n+1} = \bar{D}^{-1}\bar{f} - \bar{D}^{-1}(\bar{M} - \bar{D})\bar{u}^n$$

$\bar{M}'$  can be defined as  $\bar{M} - \bar{D}$  for simplicity, and the elements of  $\bar{D}$  can be stored in the vector  $\bar{d}$ . This allows the following expression, in which “ $\div$ ” represents an element-by-element division.

$$\bar{u}^{n+1} = (\bar{f} - \bar{M}'\bar{u}^n) \div \bar{d}$$

The procedure is performed by a matrix-vector multiplication, followed by an element-by-element vector subtraction, and then an element-by-element vector division. The equation is successively evaluated for  $\bar{u}^{n+1}$ , and at the end of each iteration, the components of  $\bar{u}$  are updated simultaneously [1].

The advantage of Jacobi iteration in this demonstration is that the evaluation of  $\bar{u}$  components can be performed independently, in parallel between updates. The algorithm may be summarized in the flowchart shown in Fig. 2.

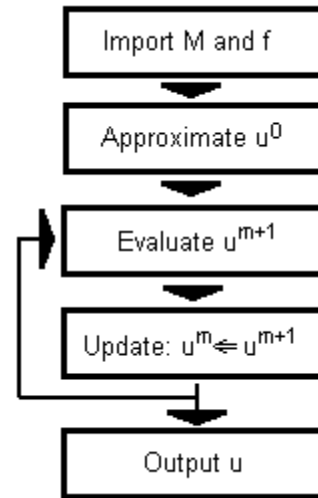


Fig. 2. Iterative algorithm.

In this research, the CPU and GPU Jacobi solvers are applied to a matrix generated by a finite element program. Other numerical methods may utilize a completely different approach. In a finite difference analysis, for example, the only elements contributing to the value of an unknown are its neighbors, so the algorithm only examines a point's neighbors to update its value, rather than an entire matrix row. A performance increase has been demonstrated in finite difference time domain by applying GPU programming in this way [2].

#### IV. BROOKGPU

BrookGPU was developed at Stanford University as an offshoot of the Brook stream processing language, created for the Merrimac streaming supercomputer. Brook extends C with the concepts of *streams* and *kernels*. A stream is similar in concept to an array, and can contain float values as well as the multicomponent vector types – float2, float3, and float4 as shown in Fig. 3. Data may be passed from an array to a stream using the streamRead operator and vice-versa using streamWrite. Other than these operators, streams may only be accessed in functions called kernels. Within a kernel, stream operations are performed in parallel, and the concept of stream shape must be considered to ensure proper operation [3]. As an example, the following program segment contains the Jacobi iteration using BrookGPU.

The first step in writing a BrookGPU program is to define and initialize the stream. Data arrays declared outside the provided program segment are read into the streams using the StreamRead function. For clarity, all streams use unmodified names, while arrays used by the CPU are given a suffix – i.e. *d* is a GPU stream while *d\_array* is stored in system memory. The Jacobi algorithm is divided into four kernels within the main *for* loop. First a multiplication is performed in which each row of matrix *M* is multiplied element-by-element with the row vector *u*. The resulting *temp* matrix is sent to a reduction kernel, in which its rows are summed yielding a column vector. This vector represents the matrix-vector product, which is then transposed to a row vector. The statement *indexof row.yx* returns an index of the current row element with x and y switched, such that row[x,y] = column[y,x]. The final kernel updates by evaluating the statement of row vectors element-by-element. After the iterations are complete, the computed value of *u* is written back to a standard array from the GPU.

A BrookGPU source file containing streams and kernels must be converted to C and shader code using the brcc compiler. After C compilation, program execution calls upon the Brook Runtime, which controls implementation of the kernels on the GPU. Compiler switches applied to brcc allow conversion to Cg or HLSL and optimizations for ATI and nVidia hardware. The last official release of BrookGPU is v0.4, released October 15, 2004, so development using modern hardware and drivers may not be efficient. The Jacobi program produced oddly scaled results until graphics drivers were updated. Even after the driver update, however, the scaling problems would return for all but impractically small matrix sizes. For this reason, a performance evaluation will not be conducted on the BrookGPU Jacobi implementation. At the time of writing, a long awaited update (billed as the v0.5 “test release”) has been made available for download, but has not been tested in this project.

#### Main Program Segments:

```
// DEFINE STREAMS
float d<1,n>; //row
float f<1,n>; //row
float u<1,n>; //row
float columnProduct<n,1>; //column
float rowProduct<1,n>; //row
float M<n,n>; //matrix
float temp<n,n>; //matrix

// INITIALIZE STREAMS
streamRead( d, d_array ); //diagonal
streamRead( f, f_array ); //forcing vector
streamRead( M, M_array ); //matrix M
streamRead( u, u_array ); //initial approx.

// RUN JACOBI ITERATIONS
for( i=0; i<iterations; i++ )
{
    mul( M, u, temp );
    sum( temp, columnSum );
    transpose( columnProduct, rowProduct );
    update( d, f, rowProduct, u );
}

streamWrite( u, u_array ); //output u
```

#### Kernels:

```
kernel void mul( float a<>, float b<>, out float c<> )
    c = a * b;

reduce void sum( float a<>, reduce float r<> )
    r += a;

kernel void transpose(float column[[ ]], out float
row<>)
    row=column[ indexof row.yx ];

kernel void update( float diagonal<>, float forcing<>,
float product<>, out float new<> )
    new = ( forcing - product ) / diagonal;
```

Fig. 3. Sample from BrookGPU Jacobi program.

#### V. ACCELERATOR

Accelerator was produced by Microsoft Research, and is designed to abstract hardware such as GPUs and cell processors. The current implementation provides abstraction for GPGPU programming on C#, and correspondence with a developer on the project suggests that a native C++ version may eventually be released. Accelerator provides a ParallelArray class that contains all necessary functions – I/O, element operations, reductions, transformations, and linear algebra. Linear

algebra routines include vector and matrix multiplications. The `ParallelArray` class also contains several subclasses such as `IntParallelArray` and `Float4ParallelArray` which define data types for single and multicomponent parallel data. Unlike `BrookGPU`, the shaders are not explicitly separated in the form of kernels. Instead, pixel shaders are created from the Accelerator operations automatically. This does not allow as much control over the underlying shaders, but it allows them to be created and optimized by the compiler at runtime [4].

In the Accelerator program of Fig. 4, the GPU must be initialized before additional calls to the `ParallelArray` class, which has been shortened to `PA` for simplicity. Four arrays are read into disposable float parallel arrays, which must be disposed at the end of the program. The Jacobi iteration is straightforward: the matrix-vector product is performed using the `InnerProduct` function of the `ParallelArray` class, and the new value of  $u$  is calculated as in the `BrookGPU` program. An additional step is required for memory management. Every time `PA.Evaluate()` is called, a new array is allocated on the GPU. The old array is explicitly disposed with each iteration in order to free memory. At the completion of the iterations, the `ToArray` function is used to recover the computed value of  $u$ , and all other memory allocated on the GPU is freed.

```
// INIT & UPLOAD TO GPU
PA.InitGPU();
DFPA d = new DFPA( d_array );
DFPA f = new DFPA( f_array );
DFPA M = new DFPA( M_array );
DFPA u = new DFPA( u_array );
DFPA uNew = null;

// BEGIN JACOBI ITERATIONS
for( i=0; i<iterations; i++ )
{
    // MULTIPLY MATRIX BY u
    FPA product = PA.InnerProduct( M, u );
    // UPDATE u
    uNew = PA.Evaluate( ( f - product ) / d );
    u.Dispose();
    u = uNew;
}

// DOWNLOAD RESULT FROM GPU & CLEAN UP
PA.ToArray( u, out u_array );
d.Dispose();
f.Dispose();
M.Dispose();
uNew.Dispose();
PA.UnInit();
```

Fig. 4. Sample from MSR Accelerator Jacobi program.

Programming with accelerator requires Microsoft Visual Studio or Visual C# Express (available for download) and naturally targets the DirectX graphics interface. No additional steps are required in the compilation process other than including the `accelerator.dll` file within the project. A disadvantage, however, is that Accelerator ties to DirectX through the .NET framework requiring a Microsoft language such as C#. Interfacing existing programs to Accelerator involves use of the .NET framework, or by writing to a data file that can be imported by an Accelerator program. The current version of Accelerator – available from [research.microsoft.com](http://research.microsoft.com) – is v1.1, last updated July 9, 2007.

## VI. ELECTROSTATIC EXAMPLE

As a practical example of electromagnetics computations on GPUs, a simple electrostatic problem domain is studied. Consider a rectangular cross-section. The sides and bottom of this domain are maintained at ground potential, while the top is excited with the positive half-cycle of a unit-sinusoidal potential source. The cross-section is discretized by a triangular mesh [5], and a nodal finite element analysis is performed.

Although the Jacobi program is a proof-of-concept rather than an optimized solver, it has been successfully applied to the matrix equation resulting from the finite element analysis. For the finer mesh of Fig. 5, the RMS error present between the computed and theoretical potentials is 0.0075. The error between the computed CPU and GPU results is negligible, suggesting no significant loss of precision between the two architectures in this case.

## VII. PERFORMANCE ANALYSIS

Performance of the GPU (using Accelerator) and CPU Jacobi algorithm implementations was examined. Various mesh densities provided differing numbers of unknowns in order to compute speed factors for varying matrix sizes. The speed factor for this application is defined as the ratio of CPU to GPU processing time. Sufficient parallel computation must be performed in order to overcome the communication and setup penalties of the GPU.

For consistency, the initial approximation is set to 0.5 for each unknown. While no test for convergence is employed in the current version of the program, the number of iterations chosen to be ten times the number of unknowns for each mesh in order to assure convergence without an explicit test. For trials of less than 500 unknowns, the communication and set up time required by the GPU outweighs any performance increase, which can be noticed from the data presented in Figs. 6 and 7. The speed factor increases to 21.1 for the case of 4000 unknowns, beyond which the CPU runs were not feasible

on the test machine.

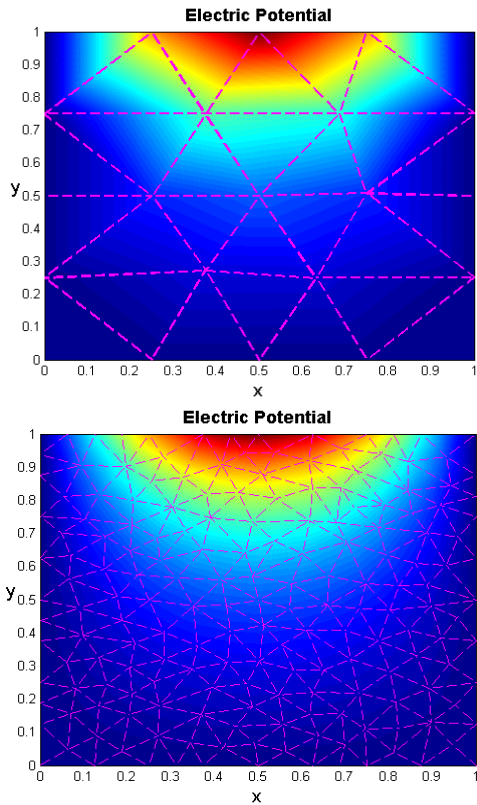


Fig. 5. Electrostatic potential in rectangular tube using different meshes.

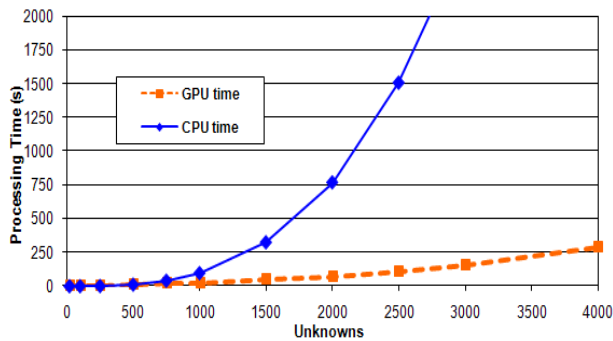


Fig. 6. Processing time for Jacobi iterations. NVidia 8800GT GPU, 2.2GHz Athlon 64 CPU.

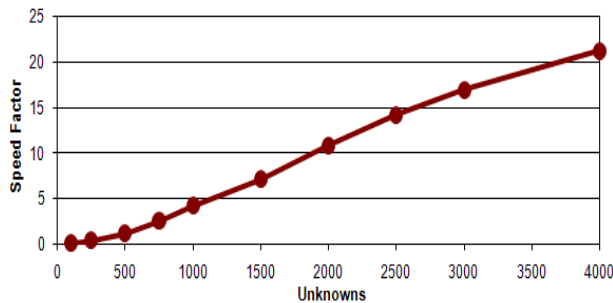


Fig. 7. Speed factor (CPU time / GPU time). NVidia 8800GT GPU, 2.2GHz Athlon 64 CPU.

Figure 8 illustrates the performance increase by generation of graphics technology using a benchmark of 2000 unknowns. The devices used in this graph spanned from a GPU produced in October 2002 (earliest technology compatible with Accelerator) to a current generation one. All other GPU results were produced using an NVidia 8800GT, while CPU results were produced using a single core of a 2.2GHz AMD Athlon 64 processor.

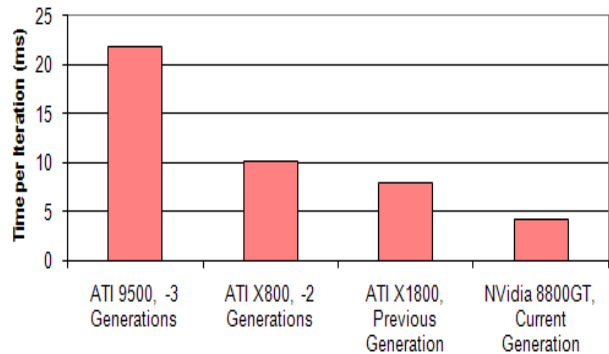


Fig. 8. Accelerator benchmark.

### VIII. CONCLUSIONS

General purpose computation on graphics processing units is now available to scientific and engineering programmers through the rise of high level interfaces such as BrookGPU and Accelerator. Through the implementation of the Jacobi algorithm, both BrookGPU and Accelerator syntax and programming issues have been discussed. Performance analysis of the Accelerator program has provided insight on the current power and continuing performance increases available through the use of GPUs.

### REFERENCES

- [1] E. Kreyszig, *Advanced Engineering Mathematics, 8<sup>th</sup> edition*, John Wiley and Sons, New York, 1999.
- [2] M. J. Inman and A. Z. Elsherbeni, "Programming video cards for computational electromagnetics applications," *IEEE Antennas and Propagation Magazine*, vol. 47, no. 6, December 2005.
- [3] I. Buck, T. Foley, D. Horn, J. Sugarman, K. Fatahalian, M. Houston, and P. Hanrahan, "Brook for GPU's: Stream computing on graphics hardware," *Transactions on Graphics* 23, August 2004.
- [4] D. Tarditi; S. Puri, and J. Oglesby, "Accelerator: using data parallelism to program GPU's for general-purpose uses," *Proceedings of the 12th international conference on Architectural support for programming languages and operating systems 2006*, San Jose, CA, USA, October 2006.
- [5] J. R. Shewchuk, "Triangle: engineering a 2D quality

mesh generator and delaunay triangulator,” *Applied Computational Geometry: Towards Geometric Engineering*, vol. 1148 Lecture Notes in Computer Science, pp. 203-222, Springer-Verlag, Berlin, May 1996.



**Maxwell Woolsey** was born in Oxford, Mississippi on October 25, 1981. He earned his B.S. *summa cum laude* in electrical engineering from the University of Mississippi in 2005 and is currently pursuing his M.S. degree in electromagnetics also at the University of Mississippi. His primary focus is the

application of FEM techniques to high-frequency electromagnetics problems. Other related interests include audio and microwave circuit design as well as parallel computing.



**W. Elliott Hutchcraft** was born in Lexington, Kentucky on April 29, 1973. He earned his B.S. in electrical engineering at the University of Mississippi, Oxford, MS in 1996, his M.S. in electrical engineering at the University of Mississippi, Oxford, MS in 1998 and his Ph. D. in electrical

engineering at the University of Mississippi, Oxford, MS in 2003. He is an Assistant Professor in the Department of Electrical Engineering at the University of Mississippi in Oxford, Mississippi. Dr. Hutchcraft is a member of Eta Kappa Nu, Sigma Xi, IEEE, Tau Beta Pi, Phi Kappa Phi, and ARFTG.



**Richard K. Gordon** was born in Birmingham, Alabama on November 26, 1959. He earned his B.S. in physics at Birmingham Southern College, Birmingham, AL in 1983, his M.S. in mathematics at the University of Illinois, Urbana, IL in 1986 and his Ph. D. in electrical engineering at the

University of Illinois, Urbana, IL in 1990. He is an Associate Professor in the Department of Electrical Engineering at the University of Mississippi in Oxford, Mississippi. Dr. Gordon is a member of Eta Kappa Nu, Phi Beta Kappa, and Tau Beta Pi.



## 2008 INSTITUTIONAL MEMBERS

AUSTRALIAN DEFENCE LIBRARY  
Northcott Drive  
Canberra, A.C.T. 2600 Australia

BAE SYSTEMS  
Technology Center – Library  
W. Hanningfield Road  
Great Baddow, Chelmsford  
United Kingdom CM2 8HN

BEIJING BOOK COMPANY  
701 East Linden Avenue  
Linden, NJ 07036

DARTMOUTH COLLEGE  
6193 Murdough Center  
Hanover, NH 03755-3560

DSTO-DSTORL EDINBURGH  
Jets AU/33851-99, PO Box 562  
Milsons Point, NSW  
Australia 1565

DTIC-OCP/LIBRARY  
8725 John J. Kingman Rd. Ste 0944  
Ft. Belvoir, VA 22060-6218

ELSEVIER  
Bibliographic Databases  
PO Box 2227  
Amsterdam, Netherlands 1000 CE

ENGINEERING INFORMATION, INC  
PO Box 543  
Amsterdam, Netherlands 1000 Am

ETSE TELECOMUNICACION  
Biblioteca, Campus Lagoas  
Vigo, 36200 Spain

FGAN-FHR  
Neuenahrerstrasse 20  
Wachtberg, Germany 53343

FLORIDA INTERNATIONAL UNIV  
10555 W. Flagler Street  
Miami, FL 33174

GEORGIA TECH LIBRARY  
225 North Avenue, NW  
Atlanta, GA 30332-0001

HANYANG UNIVERSITY  
Paiknam Academic Info. Ctr Library  
17 Haengdang-Dong  
Seongdong-Ku  
Seoul, South Korea 133-791

HELSINKI UNIVERSITY  
Otakaari 5-A, PO Box 3000  
Espoo, Finland 02015 TTK

HRL LABS, RESEARCH LIBRARY  
3011 Malibu Canyon  
Malibu, CA 90265

IEE INSPEC/Acquisitions Section  
Michael Faraday House  
6 Hills Way  
Stevenage, Herts UK SG1 2AY

INSTITUTE FOR SCIENTIFIC INFO.  
Publication Processing Dept.  
3501 Market St.  
Philadelphia, PA 19104-3302

IPS RADIO & SPACE SERVICES  
PO Box 1386  
Haymarket NSW Australia 1240

KUWAIT UNIVERSITY  
Postfach/ PO Box 432  
Basel, Switzerland 4900

KYUNGPOOK NATL UNIVERSITY  
CPO Box 5907  
Seoul, South Korea 100-659

LIBRARY – DRDC OTTAWA  
3701 Carling Avenue  
Ottawa, Ontario, Canada K1A 0Z4

LIBRARY of CONGRESS  
Reg. Of Copyrights  
Attn: 40T Deposits  
Washington DC, 20559

LINDA HALL LIBRARY  
5109 Cherry Street  
Kansas City, MO 64110-2498

MISSISSIPPI STATE UNIV LIBRARY  
PO Box 9570  
Mississippi State, MS 39762

MISSOURI S&T  
400 W 14<sup>th</sup> Street  
Rolla, MO 56409

MIT LINCOLN LABORATORY  
Periodicals Library  
244 Wood Street  
Lexington, MA 02420

NAVAL POSTGRADUATE SCHOOL  
Attn: J. Rozdal/411 Dyer Rd./ Rm 111  
Monterey, CA 93943-5101

NAVAL RESEARCH LAB  
Code 3516  
4555 Overlook Avenue SW  
Washington, DC 20375

NDL KAGAKU  
C/O KWE-ACCESS  
PO Box 300613 (JFK A/P)  
Jamaica, NY 11430-0613

OHIO STATE UNIVERSITY  
1320 Kinnear Road  
Columbus, OH 43212

OVIEDO LIBRARY  
PO Box 830679  
Birmingham, AL 35283

PENN STATE UNIVERSITY  
126 Paterno Library  
University Park, PA 16802-1808

CHRIS A. PENWELL  
3500 Long Drive  
Minder, NV 89423

PHILIPS RESEARCH LABORATORY  
Cross Oak Lane, Stella Cox  
Salfords, Redhill  
UK RH1 5HA

DAVID J. PINION  
1122 E Pike Street #1217  
Seattle, WA 98122

SOUTHWEST RESEARCH  
INSTITUTE  
6220 Culebra Road  
San Antonio, TX 78238

SWETS INFORMATION SERVICES  
160 Ninth Avenue, Suite A  
Runnemede, NJ 08078

TECHNISCHE UNIV. DELFT  
Mekelweg 4, Delft, Holland, 2628 CD  
Netherlands

TELSTRA  
TRL/M2/770 Blackburn Road  
Clayton, Victoria, Australia 3168

TIB & UNIV. BIB. HANNOVER  
DE/5100/G1/0001  
Welfengarten 1B  
Hannover, Germany 30167

TU DARMSTADT  
Schlossgartenstrasse 8  
Darmstadt, Hessen  
Germany D-64289

UNIV OF CENTRAL FLORIDA LIB.  
4000 Central Florida Boulevard  
Orlando, FL 32816-8005

UNIV OF COLORADO LIBRARY  
Campus Box 184  
Boulder, CO 80309-0184

UNIV OF MISSOURI-ROLLA LIB.  
1870 Miner Circle  
Rolla, MO 65409-0001

USAE ENG. RES. & DEV. CENTER  
Attn: Library/Journals  
72 Lyme Road  
Hanover, NH 03755-1290

# ACES COPYRIGHT FORM

This form is intended for original, previously unpublished manuscripts submitted to ACES periodicals and conference publications. The signed form, appropriately completed, MUST ACCOMPANY any paper in order to be published by ACES. PLEASE READ REVERSE SIDE OF THIS FORM FOR FURTHER DETAILS.

TITLE OF PAPER:

RETURN FORM TO:

Dr. Atef Z. Elsherbeni  
University of Mississippi  
Dept. of Electrical Engineering  
Anderson Hall Box 13  
University, MS 38677 USA

AUTHORS(S)

PUBLICATION TITLE/DATE:

---

## PART A - COPYRIGHT TRANSFER FORM

(NOTE: Company or other forms may not be substituted for this form. U.S. Government employees whose work is not subject to copyright may so certify by signing Part B below. Authors whose work is subject to Crown Copyright may sign Part C overleaf).

The undersigned, desiring to publish the above paper in a publication of ACES, hereby transfer their copyrights in the above paper to The Applied Computational Electromagnetics Society (ACES). The undersigned hereby represents and warrants that the paper is original and that he/she is the author of the paper or otherwise has the power and authority to make and execute this assignment.

**Returned Rights:** In return for these rights, ACES hereby grants to the above authors, and the employers for whom the work was performed, royalty-free permission to:

1. Retain all proprietary rights other than copyright, such as patent rights.

2. Reuse all or portions of the above paper in other works.

3. Reproduce, or have reproduced, the above paper for the author's personal use or for internal company use provided that (a) the source and ACES copyright are indicated, (b) the copies are not used in a way that implies ACES endorsement of a product or service of an employer, and (c) the copies per se are not offered for sale.

4. Make limited distribution of all or portions of the above paper prior to publication.

5. In the case of work performed under U.S. Government contract, ACES grants the U.S. Government royalty-free permission to reproduce all or portions of the above paper, and to authorize others to do so, for U.S. Government purposes only.

**ACES Obligations:** In exercising its rights under copyright, ACES will make all reasonable efforts to act in the interests of the authors and employers as well as in its own interest. In particular, ACES REQUIRES that:

1. The consent of the first-named author be sought as a condition in granting re-publication permission to others.

2. The consent of the undersigned employer be obtained as a condition in granting permission to others to reuse all or portions of the paper for promotion or marketing purposes.

In the event the above paper is not accepted and published by ACES or is withdrawn by the author(s) before acceptance by ACES, this agreement becomes null and void.

---

AUTHORIZED SIGNATURE

TITLE (IF NOT AUTHOR)

---

EMPLOYER FOR WHOM WORK WAS PERFORMED

DATE FORM SIGNED

## Part B - U.S. GOVERNMENT EMPLOYEE CERTIFICATION

(NOTE: if your work was performed under Government contract but you are not a Government employee, sign transfer form above and see item 5 under Returned Rights).

This certifies that all authors of the above paper are employees of the U.S. Government and performed this work as part of their employment and that the paper is therefor not subject to U.S. copyright protection.

---

AUTHORIZED SIGNATURE

TITLE (IF NOT AUTHOR)

---

NAME OF GOVERNMENT ORGANIZATION

DATE FORM SIGNED

---

## PART C - CROWN COPYRIGHT

(NOTE: ACES recognizes and will honor Crown Copyright as it does U.S. Copyright. It is understood that, in asserting Crown Copyright, ACES in no way diminishes its rights as publisher. Sign only if ALL authors are subject to Crown Copyright).

This certifies that all authors of the above Paper are subject to Crown Copyright. (Appropriate documentation and instructions regarding form of Crown Copyright notice may be attached).

---

AUTHORIZED SIGNATURE

TITLE OF SIGNEE

---

NAME OF GOVERNMENT BRANCH

DATE FORM SIGNED

### Information to Authors

#### ACES POLICY

ACES distributes its technical publications throughout the world, and it may be necessary to translate and abstract its publications, and articles contained therein, for inclusion in various compendiums and similar publications, etc. When an article is submitted for publication by ACES, acceptance of the article implies that ACES has the rights to do all of the things it normally does with such an article.

In connection with its publishing activities, it is the policy of ACES to own the copyrights in its technical publications, and to the contributions contained therein, in order to protect the interests of ACES, its authors and their employers, and at the same time to facilitate the appropriate re-use of this material by others.

The new United States copyright law requires that the transfer of copyrights in each contribution from the author to ACES be confirmed in writing. It is therefore necessary that you execute either Part A-Copyright Transfer Form or Part B-U.S. Government Employee Certification or Part C-Crown Copyright on this sheet and return it to the Managing Editor (or person who supplied this sheet) as promptly as possible.

#### CLEARANCE OF PAPERS

ACES must of necessity assume that materials presented at its meetings or submitted to its publications is properly available for general dissemination to the audiences these activities are organized to serve. It is the responsibility of the authors, not ACES, to determine whether disclosure of their material requires the prior consent of other parties and if so, to obtain it. Furthermore, ACES must assume that, if an author uses within his/her article previously published and/or copyrighted material that permission has been obtained for such use and that any required credit lines, copyright notices, etc. are duly noted.

#### AUTHOR/COMPANY RIGHTS

If you are employed and you prepared your paper as a part of your job, the rights to your paper initially rest with your employer. In that case, when you sign the copyright form, we assume you are authorized to do so by your employer and that your employer has consented to all of the terms and conditions of this form. If not, it should be signed by someone so authorized.

**NOTE RE RETURNED RIGHTS:** Just as ACES now requires a signed copyright transfer form in order to do "business as usual", it is the intent of this form to return rights to the author and employer so that they too may do "business as usual". If further clarification is required, please contact: The Managing Editor, R. W. Adler, 798 Lighthouse Ave. PMB 331, Monterey, CA 93940-1010, USA (408)656-2352.

Please note that, although authors are permitted to re-use all or portions of their ACES copyrighted material in other works, this does not include granting third party requests for reprinting, republishing, or other types of re-use.

#### JOINT AUTHORSHIP

For jointly authored papers, only one signature is required, but we assume all authors have been advised and have consented to the terms of this form.

#### U.S. GOVERNMENT EMPLOYEES

Authors who are U.S. Government employees are not required to sign the Copyright Transfer Form (Part A), but any co-authors outside the Government are.

Part B of the form is to be used instead of Part A only if all authors are U.S. Government employees and prepared the paper as part of their job.

**NOTE RE GOVERNMENT CONTRACT WORK:** Authors whose work was performed under a U.S. Government contract but who are not Government employees are required so sign Part A-Copyright Transfer Form. However, item 5 of the form returns reproduction rights to the U. S. Government when required, even though ACES copyright policy is in effect with respect to the reuse of material by the general public.

January 2002

## INFORMATION FOR AUTHORS

### PUBLICATION CRITERIA

Each paper is required to manifest some relation to applied computational electromagnetics. **Papers may address general issues in applied computational electromagnetics, or they may focus on specific applications, techniques, codes, or computational issues.** While the following list is not exhaustive, each paper will generally relate to at least one of these areas:

- 1. Code validation.** This is done using internal checks or experimental, analytical or other computational data. Measured data of potential utility to code validation efforts will also be considered for publication.
- 2. Code performance analysis.** This usually involves identification of numerical accuracy or other limitations, solution convergence, numerical and physical modeling error, and parameter tradeoffs. However, it is also permissible to address issues such as ease-of-use, set-up time, run time, special outputs, or other special features.
- 3. Computational studies of basic physics.** This involves using a code, algorithm, or computational technique to simulate reality in such a way that better, or new physical insight or understanding, is achieved.
- 4. New computational techniques** or new applications for existing computational techniques or codes.
- 5. “Tricks of the trade”** in selecting and applying codes and techniques.
- 6. New codes, algorithms, code enhancement, and code fixes.** This category is self-explanatory, but includes significant changes to existing codes, such as applicability extensions, algorithm optimization, problem correction, limitation removal, or other performance improvement. **Note: Code (or algorithm) capability descriptions are not acceptable, unless they contain sufficient technical material to justify consideration.**
- 7. Code input/output issues.** This normally involves innovations in input (such as input geometry standardization, automatic mesh generation, or computer-aided design) or in output (whether it be tabular, graphical, statistical, Fourier-transformed, or otherwise signal-processed). Material dealing with input/output database management, output interpretation, or other input/output issues will also be considered for publication.
- 8. Computer hardware issues.** This is the category for analysis of hardware capabilities and limitations of various types of electromagnetics computational requirements. Vector and parallel computational techniques and implementation are of particular interest.

Applications of interest include, but are not limited to, antennas (and their electromagnetic environments), networks, static fields, radar cross section, inverse scattering, shielding, radiation hazards, biological effects, biomedical applications, electromagnetic pulse (EMP), electromagnetic interference (EMI), electromagnetic compatibility (EMC), power transmission, charge transport, dielectric, magnetic and nonlinear materials, microwave components, MEMS, RFID, and MMIC technologies, remote sensing and geometrical and physical optics, radar and communications systems, sensors, fiber optics, plasmas, particle accelerators, generators and motors, electromagnetic wave propagation, non-destructive evaluation, eddy currents, and inverse scattering.

Techniques of interest include but not limited to frequency-domain and time-domain techniques, integral equation and differential equation techniques, diffraction theories, physical and geometrical optics, method of moments, finite differences and finite element techniques, transmission line method, modal expansions, perturbation methods, and hybrid methods.

Where possible and appropriate, authors are required to provide statements of quantitative accuracy for measured and/or computed data. This issue is discussed in “Accuracy & Publication: Requiring quantitative accuracy statements to accompany data,” by E. K. Miller, *ACES Newsletter*, Vol. 9, No. 3, pp. 23-29, 1994, ISBN 1056-9170.

### SUBMITTAL PROCEDURE

All submissions should be uploaded to ACES server through ACES web site (<http://aces.ee.olemiss.edu>) by using the upload button, journal section. Only pdf files are accepted for submission. The file size should not be larger than 5MB, otherwise permission from the Editor-in-Chief should be obtained first. Automated acknowledgment of the electronic submission, after the upload process is successfully completed, will be sent to the corresponding author only. It is the responsibility of the corresponding author to keep the remaining authors, if applicable, informed. Email submission is not accepted and will not be processed.

### PAPER FORMAT (INITIAL SUBMISSION)

The preferred format for initial submission manuscripts is 12 point Times Roman font, single line spacing and single column format, with 1 inch for top, bottom, left, and right margins. Manuscripts should be prepared for standard 8.5x11 inch paper.

### EDITORIAL REVIEW

**In order to ensure an appropriate level of quality control,** papers are peer reviewed. They are reviewed both for

technical correctness and for adherence to the listed guidelines regarding information content and format.

### **PAPER FORMAT (FINAL SUBMISSION)**

Only camera-ready electronic files are accepted for publication. The term “**camera-ready**” means that the material is neat, legible, reproducible, and in accordance with the final version format listed below.

The following requirements are in effect for the final version of an ACES Journal paper:

1. The paper title should not be placed on a separate page. The title, author(s), abstract, and (space permitting) beginning of the paper itself should all be on the first page. The title, author(s), and author affiliations should be centered (center-justified) on the first page. The title should be of font size 16 and bolded, the author names should be of font size 12 and bolded, and the author affiliation should be of font size 12 (regular font, neither italic nor bolded).
2. An abstract is required. The abstract should be a brief summary of the work described in the paper. It should state the computer codes, computational techniques, and applications discussed in the paper (as applicable) and should otherwise be usable by technical abstracting and indexing services. The word “Abstract” has to be placed at the left margin of the paper, and should be bolded and italic. It also should be followed by a hyphen (–) with the main text of the abstract starting on the same line.
3. All section titles have to be centered and all the title letters should be written in caps. The section titles need to be numbered using roman numbering (I. II. ....)
4. Either British English or American English spellings may be used, provided that each word is spelled consistently throughout the paper.
5. Internal consistency of references format should be maintained. As a guideline for authors, we recommend that references be given using numerical numbering in the body of the paper (with numerical listing of all references at the end of the paper). The first letter of the authors’ first name should be listed followed by a period, which in turn, followed by the authors’ complete last name. Use a coma (,) to separate between the authors’ names. Titles of papers or articles should be in quotation marks (“ ”), followed by the title of journal, which should be in italic font. The journal volume (vol.), issue number (no.), page numbering (pp.), month and year of publication should come after the journal title in the sequence listed here.
6. Internal consistency shall also be maintained for other elements of style, such as equation numbering. As a guideline for authors who have no other preference, we suggest that equation numbers be placed in parentheses at the right column margin.

7. The intent and meaning of all text must be clear. For authors who are not masters of the English language, the ACES Editorial Staff will provide assistance with grammar (subject to clarity of intent and meaning). However, this may delay the scheduled publication date.
8. Unused space should be minimized. Sections and subsections should not normally begin on a new page.

ACES reserves the right to edit any uploaded material, however, this is not generally done. It is the author(s) responsibility to provide acceptable camera-ready pdf files. Incompatible or incomplete pdf files will not be processed for publication, and authors will be requested to re-upload a revised acceptable version.

### **COPYRIGHTS AND RELEASES**

Each primary author must sign a copyright form and obtain a release from his/her organization vesting the copyright with ACES. Copyright forms are available at ACES, web site (<http://aces.ee.olemiss.edu>). To shorten the review process time, the executed copyright form should be forwarded to the Editor-in-Chief immediately after the completion of the upload (electronic submission) process. Both the author and his/her organization are allowed to use the copyrighted material freely for their own private purposes.

Permission is granted to quote short passages and reproduce figures and tables from an ACES Journal issue provided the source is cited. Copies of ACES Journal articles may be made in accordance with usage permitted by Sections 107 or 108 of the U.S. Copyright Law. This consent does not extend to other kinds of copying, such as for general distribution, for advertising or promotional purposes, for creating new collective works, or for resale. The reproduction of multiple copies and the use of articles or extracts for commercial purposes require the consent of the author and specific permission from ACES. Institutional members are allowed to copy any ACES Journal issue for their internal distribution only.

### **PUBLICATION CHARGES**

All authors are allowed for 8 printed pages per paper without charge. Mandatory page charges of \$75 a page apply to all pages in excess of 8 printed pages. Authors are entitled to one, free of charge, copy of the journal issue in which their paper was published. Additional reprints are available for a nominal fee by submitting a request to the managing editor or ACES Secretary.

Authors are subject to fill out a one page over-page charge form and submit it online along with the copyright form before publication of their manuscript.

**ACES Journal is abstracted in INSPEC, in Engineering Index, DTIC, Science Citation Index Expanded, the Research Alert, and to Current Contents/Engineering, Computing & Technology.**