APPLIED

COMPUTATIONAL

ELECTROMAGNETICS

SOCIETY

JOURNAL

THE APPLIED COMPUTATIONAL ELECTROMAGNETICS SOCIETY JOURNAL

TABLE OF CONTENTS – EXPRESS PAPERS

Half-Loop Segmented Antenna with Omnidirectional Hemispherical Coverage for Wireless Communications

Payam Nayeri [1], Atef Z. Elsherbeni [1], Roger Hasse [2], and Darko Kajfez [3]

[1] Department of Electrical Engineering and Computer Science
Colorado School of Mines, Golden, CO 80401, USA
pnayeri@mines.edu, aelsherb@mines.edu

[2] Georgia Tech Research Institute
Georgia Institute of Technology, Atlanta, GA 30332, USA
roger.hasse@gtri.gatech.edu

[3] Department of Electrical Engineering
University of Mississippi, University, MS 38677, USA
eedarko@olemiss.edu

Abstract ─ Segmented antennas loaded with reactance elements at the junctions between segments provide additional parameters for shaping antenna characteristics. The design procedure consists of first, circuit analysis of the multiport antenna description to determine the initial values of the lumped element reactances and second, of fine tuning the antenna dimensions full-wave simulation software. Measured results show that a good matching and an omnidirectional radiation pattern in the vertical plane is maintained up to $\theta = 70°$ with a circular ground plane of 1m diameter.

Index Terms ─ Antenna directivity, antenna impedance matrix, design optimization, omnidirectional antenna, partitioned antenna, printed circuit antenna, segmented antenna.

I. INTRODUCTION

Many applications of wireless communication require antennas that provide omnidirectional radiation patterns. For example, an antenna designed for communication with non-stationary low-earth-orbit satellites should display an omnidirectional pattern in the upper half plane of a spherical system of coordinates. One possibility of creating such a radiation pattern is with the use of a half-loop segmented antenna [1] that can easily be fabricated in a printed-circuit form. As this antenna creates a linearly polarized field, a combination of two orthogonal antennas driven in a phase quadrature would be necessary for achieving circularly polarized radiation. This paper presents a new design of segmented half loop antenna that achieves omnidirectional radiation above a ground plane in one azimuthal plane.

The need for segmentation can be best understood by considering the characteristics of a solid (unsegmented) antenna such as shown in Fig. 1. We assume that the antenna circumference 2a is approximately one-half wavelength, and the two generators are in phase opposition. A transmitting loop antenna placed above a perfectly conducting ground plane as shown in Fig. 1, produces the same field in the upper half-space as another double-sized, full-loop antenna, located in free space. To distinguish the antenna in Fig. 1 from other loop antennas, we will call it a *half-loop* antenna.
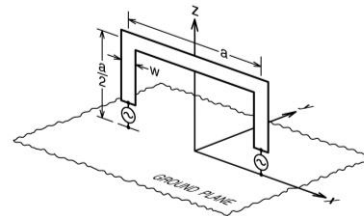
Fig. 1. A solid half-loop antenna above a perfectly conducting ground plane.

The directivity pattern created by such antenna is well known in the literature as is not shown for brevity. This pattern differs significantly from an ideal radiation pattern produced by an elementary magnetic dipole placed in the horizontal direction. In particular the directivity pattern of such an antenna drops by 5 decibels already at an angle of $\theta = 60°$, instead of being constant up to the horizon as in the case of the elementary dipole. Note that in spherical system of coordinates, angle θ is counted from the vertical z axis. On the other hand, when the half-loop antenna of the

same overall dimensions is partitioned into 7 total segments, separated by lumped reactances, the simulated radiation pattern becomes as shown in Fig. 2, which is almost identical with an ideal magnetic dipole pattern. In addition, the segmented antenna is better matched than the solid version. More details about the segmentation procedure will be provided in Section III. This pattern has been obtained by the use of the full wave EM simulation software Ansys HFSS [2], assuming an infinite ground plane.
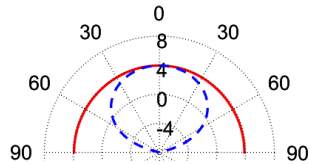
Fig. 2. Simulated directivity patterns of half-loop segmented antennas. The solid and dashed lines correspond to the *xz* and *yz* planes, respectively.

II. EQUIVALENT CIRCUIT OF THE HALF-LOOP SEGMENTED ANTENNA

For a half-loop antenna with electrical length no greater than one-half wavelength, the radiation pattern could be made omnidirectional in the orthogonal plane, if one could maintain a stationary phase of the current distribution along the antenna conductor. The control of the current distribution along the antenna will be accomplished by segmenting the half-loop antenna into 7 segments as shown in Fig. 3. Without the loss of generality, the substrate used for the loop is a 32 mil Rogers RO4003 laminate.
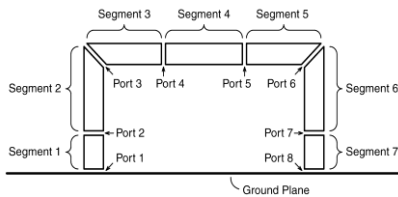
Fig. 3. The conductor of a half-loop antenna in Fig. 1 is divided into 7 segments.

Between each two neighboring segments a lumped reactance will be inserted leading to an equivalent circuit as shown in Fig. 4. From the circuit theory point of view, we see two impedance matrices; the antenna is described by matrix $\mathbf{Z}_a$ while the "external" reactances constitute diagonal impedance matrix $\mathbf{Z}_{ex}$:

$$\mathbf{Z}_{ex} = diag(0, jX_2, jX_3, \dots jX_7, 0). \qquad (1)$$

The capacitive reactances are negative, and inductive reactances are positive. In order to insure that the radiation in the positive *x* direction is of the same intensity as the one in the negative *x* direction, the reactance values should be made symmetric around the *z* axis, i.e., $X_2=X_7$, $X_3=X_6$ and $X_4=X_5$. Furthermore the source voltages should be made equal to each other ($V_{g1} = V_{g8}$). Then, the two input impedances will be equal to each other ($Z_{in1} = Z_{in8}$), and the two powers delivered to the antenna ports will also be equal to each other ($P_{in1} = P_{in2}$). Considering the structure in Fig. 3 to be an array comprised of 7 elements, the impedance matrix $\mathbf{Z}_a$ can be generated using full-wave electromagnetic simulation software [2].
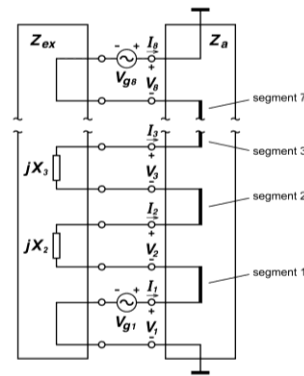
Fig. 4. Equivalent circuit of the half-loop segmented antenna.

When the voltages at the individual ports are grouped in the vector $|V\rangle$ and the port currents in a vector $|I\rangle$ the network in Fig. 4 is described by:

$$|I\rangle = (\mathbf{Z}_a + \mathbf{Z}_{ex})^{-1}|V_g\rangle. \qquad (2)$$

The first component of vector $|V_g\rangle$ is equal to V_{g1} and the 8th component is equal to V_{g8} while the rest of components are being equal to zero. For the operation in the phase opposition, $V_{g1} = V_{g8}$, and it is therefore possible to evaluate the voltages and currents at all the ports. Thus, for any combination of reactances X_2 to X_7 it is possible to determine the *exact* input impedance with the use of circuit-theory Equation (2), without the need to refer to a full wave simulation software. Furthermore, it is also possible to *approximately* compute the radiation pattern based on the results obtained by the circuit theory as follows. Each of the segments is considerably shorter than the free-space wavelength, so one can assume that the current distribution along an individual segment will not display significant fluctuations. Thus, it is reasonable to assume that the current on the segment *i* is equal to the average value of port currents entering and exiting that segment:

$$I_{si} = (I_i + I_{i+1})/2. \qquad (3)$$

The antenna system in Fig. 3 can be viewed to represent an array of known linear current sources I_{si}, so

that it is possible to evaluate the total radiated field using antenna theory [3]. It is convenient to define current moments as follows:

$$I_{mi} = \frac{2\pi d_i}{\lambda} I_{si}, \qquad (4)$$

where d_i is the length of the segment i and λ is the free space wavelength. When the segments i ($i = 1$ to 7) are rotated by angles γ_i with respect to the z-axis, their centers being located in the xz-plane at coordinates x_i and z_i, the general formula for directivity in the xz plane becomes:

$$D(\theta) = \frac{7.5}{P_{in1}} \left| \sum_{i=1}^{14} I_{mi} \sin(\theta - \gamma_i) e^{jk(x_i \sin\theta + z_i \cos\theta)} \right|^2. \qquad (5)$$

Figure 5 displays the individual current moments that are used for computing the directivity. Although there are only 7 physical segments on the antenna, there are also 7 images below the ground plane that must be added in the summation (5). It is important to note that the number of segments needed for any given design is subjective. In the proposed design initially 5 segments were considered, however it was observed that adding segments on the top corners were necessary, hence this number was increased to 7. While adding more segments is also possible, due to added complexity it should generally be avoided.
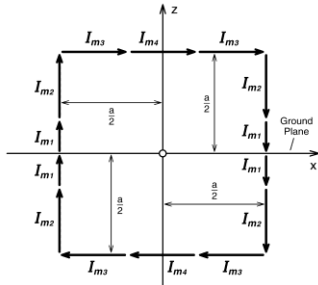
Fig. 5. The positions and orientations of the current moments on the segmented half-loop antenna.

III. OPTIMIZATION PROCEDURE FOR INDIVIDUAL REACTANCES

Expression (5) enables one to estimate the radiation pattern as a function of θ. For optimization purposes, directivity D is evaluated at a number of θ points. Afterwards, the mean value mD and the standard deviation sD of all the points are computed. The objective function for optimization is then selected to be:

$$U = \frac{w_1}{mD} + w_2 \cdot sD + w_3 \cdot \rho. \qquad (6)$$

In this way, the mean directivity is maximized, the standard deviation of directivity (i.e., the departure from the omnidirectional shape) is minimized, and the input reflection coefficient is minimized. The weight w_1 is used to emphasize the mean directivity; w_2 is used to emphasize the uniform directivity; and w_3 is used to emphasize the impedance match.

For demonstration and verification purposes, the side length of the half-loop from Fig. 6 is selected to be $a = 62$ mm, and the strip width is 5 mm. The operating frequency is selected to be $f = 1.2276$ GHz (GPS L2 band). The 8-port antenna impedance matrix is created by HFSS, while the initial optimization is performed with the Matlab® program "fminsearch" [4].

At the start of optimization, all reactances are set to be equal to -259 Ω (i.e., 0.5 pF). The three optimized reactances come out to be X_2= -500 Ω, X_3= 92.7 Ω and X_4= -432 Ω. Thus, the second and the fourth loading reactances are capacitive, while the third one is inductive. The amplitudes and phases of the current moments at the start of optimization are shown in Fig. 6 (a), and after optimization they are shown in Fig. 6 (b). It can be seen that the amplitudes of the current moments after optimization have become several times larger, because of an improved impedance match. The phase distribution has also changed somewhat after optimization, but phases still remain close to each other. The predicted mean directivity is mD= 3.6 dBi, and the predicted input return loss is 20 dB.

(a)

(b)
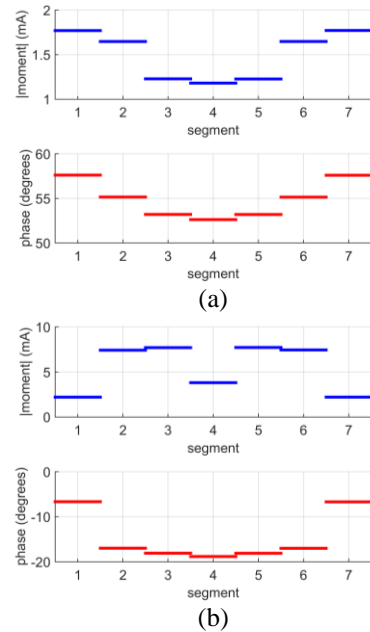
Fig. 6. Amplitudes and phases of current moments on individual segments: (a) before optimization, and (b) after optimization.

When the optimized values of lumped reactances are inserted in the antenna model, the radiation pattern obtained by HFSS is shown in Fig. 2 (b). As expected, the shape of the pattern closely resembles that of an ideal magnetic moment oriented along the y axis. As far as the directivity of this oversized magnetic dipole is concerned, the directivity value computed by HFSS is

slightly larger than the 3.6 dBi predicted by circuit analysis, namely 4.1 dBi in the z direction, and it drops to 3.8 dBi in the x direction. The simulated return loss is approximately 21.6 dB at 1.2276 GHz, which is in good agreement with the value obtained by circuit theory of 20 dB at 1.2276 GHz.

The final step is to model the required reactances of the segmented antenna. The inductance corresponding to X_3 is L_3=12.03 nH. This value is realized by a small circular loop in each of the two corners of the half-loop. The trace width for this inductor was set to 1 mm. The other two lumped elements are capacitances 0.26 and 0.30 pF, which are realized by overlapping the strips located on the opposite sides of the substrate. For these three sets of printed reactive elements, the initial dimensions were determined based on static approximation. The final dimensions, i.e., the loop radius for the inductors and overlap length for the capacitors, were determined with direct optimization of these parameters in the full-wave software. This secondary optimization resulted in an omnidirectional directivity of 3.8 dBi in both x and z directions, with the input reflection coefficient magnitude of -21.6 dB at 1.2276 GHz. The dimensions for the inductive loop radius, and the overlap capacitor lengths, were 2.22 mm, 1.55 mm, and 1.95 mm, respectively. The printed circuit arrangement, and a photo of its prototype, are shown in Fig. 7. A lumped port excitation was used in the simulation model to represent the SMA end-launch connector of the fabricated prototype. The substrate used for the ground plane and feed network is a 62 mil Rogers RT/Duroid 5880. Note that the feed network is a power divider which feeds the two ports of the HLA out of phase [1].

(a)   (b)
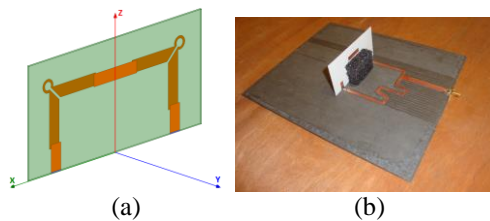
Fig. 7. (a) Printed circuit version of the antenna. (b) Half-loop segmented antenna prototype.

IV. PROTOTYPE MEASUREMENTS

The radiation pattern measured with a ground plane of 1 meter diameter is compared with the simulated pattern in Fig. 8. The antenna radiation pattern and efficiency were measured using the SG64 Starlab measurement system at Kennesaw, GA. Because of the limited ground plane size, the edge diffraction effect is clearly visible. The measured value drops below the simulated value at $\theta = 70°$. For a larger ground plane size, one would expect wider omnidirectional coverage

angle. The measured radiation efficiency at the center frequency is 93%. The measured reflection coefficient magnitude is given in Fig. 9.

(a)   (b)
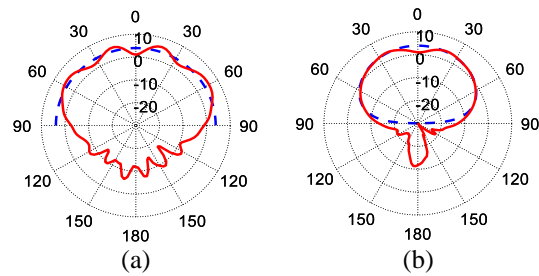
Fig. 8. Measured (solid) and simulated (dashed) radiation patterns of the antenna at 1.2276 GHz in: (a) xz or $\varphi = 0°$ and (b) yz or $\varphi = 90°$ planes.
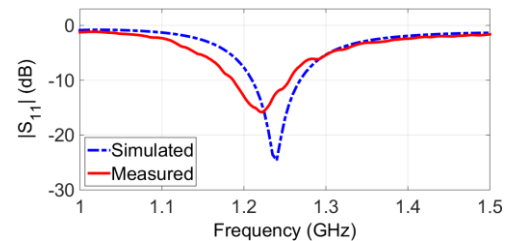
Fig. 9. Reflection coefficient magnitude of the antenna.

V. CONCLUSIONS

By allowing the individual reactances in a segmented antenna to take distinct values, it becomes possible to simultaneously optimize the radiation pattern and the input match of the antenna. A preliminary optimization is considerably speeded up by analyzing the antenna equivalent circuit. The required capacitances and inductances can be easily integrated into the printed circuit of the segmented antenna, such that no external matching circuit is necessary.

ACKNOWLEDGMENT

REFERENCES

[1]   P. Nayeri, D. Kajfez, and A. Z. Elsherbeni, "Design of a segmented half-loop antenna," *IEEE Antennas and Propagation Society International Symposium*, Memphis, TN, July 2014.
[2]   ANSYS HFSS, v. 15.0, ANSYS Inc.
[3]   R. E. Collin, *Radiation from Simple Sources*, Ch. 2, Antenna Theory, Part I, New York: McGraw-Hill, 1969.
[4]   Matlab®, The Mathworks Inc., Natick, MA.

Computational Electromagnetic Modelling of Compact Antenna Test Range Quiet Zone Probing

C. G. Parini [1], R. Dubrovka [1], and S. F. Gregson [2]

[1] School of Electronic Engineering and Computer Science
Queen Mary University of London, London, E1 4FZ, UK
c.g.parini@qmul.ac.uk, r.dubrovka@qmul.ac.uk

[2] Nearfield Systems Inc.
19730 Magellan Drive, Torrance, CA, 90502, USA
sgregson@nearfield.com

Abstract — This paper extends the authors previous simulation study [1, 2] that predicted the quality of the pseudo plane wave of an offset compact antenna test range (CATR). In this paper, the quiet-zone performance predictions are extended to rigorously incorporate the effects of probing the CATR quiet-zone using various field probes. This investigation leads to recommendations as to the optimal field probe choice and measurement uncertainties. The results of these new simulations are presented and discussed.

Index Terms — Compact antenna test range, field-probe, quiet-zone probing, reaction theorem.

I. INTRODUCTION

The single-offset compact antenna test range (CATR) is a widely deployed measurement technique for the broadband characterization of electrically large antennas at reduced range lengths. The CATR collimates the quasi-spherical wave radiated by a low gain feed into a pseudo transverse electric and magnetic (TEM) plane-wave. The coupling of this locally plane-wave into the aperture of an antenna under test (AUT) creates the classical measured "far-field" pattern. The accuracy of an antenna measured using a CATR is therefore primarily determined by the uniformity of the amplitude and phase of this illuminating pseudo plane-wave.

Traditionally, the quality of the pseudo plane wave has been assessed by "probing" the amplitude and phase across a transverse planar surface with the results being tabulation on, typically, a plane-polar grid consisting of a series of linear scans in the horizontal, vertical and perhaps inter-carinal planes. A number of workers have utilized portable planar near-field antenna test systems to acquire two-dimensional plane-rectilinear data sets that can be used to provide far greater insight into the behavior of the field in the quiet-zone (QZ) and additionally for the purposes of chamber imaging to provide angular image maps of reflections [3]. However, when mapping the CATR QZ the finitely large aperture of any realized field probe will inevitably affect the mapped fields by way of the convolution process between the pseudo plane wave of the CATR and the aperture illumination function of the scanning near-field probe, *cf.* [4] Potentially, such a discrepancy can lead to confusion when comparing CATR QZ predictions obtained from standard computational electromagnetic (CEM) models and empirical measurements as this "boxcar" field averaging process is not automatically incorporated within the numerical simulation. Several authors have undertaken CATR performance prediction modeling [7, 8, 9] with increasing levels of complexity. This paper extends our recently published comprehensive CATR QZ performance prediction software tool [1, 2] to incorporate the directive properties of several commonly used field probes so that recommendations can be made as to the most appropriate probe to use as well as providing estimates for the upper bound measurement uncertainty.

II. CATR QZ SIMULATION

The field illuminating the CATR offset parabolic reflector is typically derived from the assumed known far-field pattern of the feed antenna. This pattern could be derived CEM simulation, as is the case here, or from empirical range measurements. Figure 1 contains a mechanical drawing of the WR430 choked cylindrical waveguide feed that was used during these simulations with the realised feed shown in Fig. 2. Here, the feed is assumed nominally vertically polarised within its local coordinate system. When computing CATR QZ simulations for a horizontally polarised feed a vector isometric rotation [4] can be used to rotate the probe by 90° about its local *z*-axis so as to produce equivalent far-field patterns for a horizontally polarised probe.

Figures 3 and 4 respectively illustrate the far-field

amplitude and phase cardinal cuts of the feed antenna when resoled onto a Cartesian polarisation basis. These patterns were obtained from a proprietary three-dimensional full-wave CEM solver that used the finite difference time domain (FDTD) method. Here, the difference in beam-widths is exacerbated by presenting the patterns resolved onto a Cartesian polarisation basis [4].

Fig. 1. Model of WR430 CATR feed.
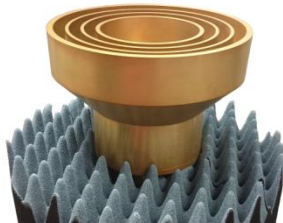
Fig. 2. Realised WR430 CATR feed.
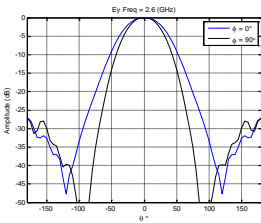
Fig. 3. Amplitude cardinal cuts of feed at 2.6 GHz.
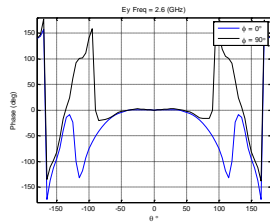
Fig. 4. Phase cardinal cuts of feed at 2.6 GHz.

The location of the phase centre was determined by means of a best fit parabolic function over the $-50° \leq \theta \leq 50°$ angular range [5]. The maximum polar angle of 50° was selected as this is the maximum angle subtended at the feed by the CATR parabolic reflector. For angles larger than this, the feed pattern spills over from the reflector and the feed pattern function for angles larger than this are unimportant. Here, the phase centre of this circular feed was determined as being at $x = y = 0$ m and $z = -0.1377$ m and was found to be extremely stable across the operating bandwidth. The phase patterns were compensated for this parabolic phase function which conceptually corresponds to installing the phase centre of the feed at the focus of the CATR parabolic reflector. The field illuminating the parabolic reflector can then be determined from far-field antenna pattern function by reintroducing the (conventionally suppressed) spherical phase function and the inverse r term. The corresponding magnetic field, as required by the field propagation algorithm, can be computed from the electric field from the TEM far-field condition [4].

As a result of the requirement to minimise feed induced blockage, as described in [1, 2], a single offset reflector CATR design is harnessed. Here, it is assumed that the vertex of the reflector is coincident with the bottom edge of the main reflector. Thus, the feed is required to be tilted up in elevation so that the boresight direction of the feed is orientated towards the centre of the reflector surface. In this case, the CATR main reflector is formed from an offset parabolic reflector with a focal length of 12' = 3.6576 m. The reflector was 4.71 m wide by 3.9 m high with serrations of 0.76 m in length. The following figure shows a false-colour plot of the magnitude of the illuminating electric field as radiated by the WR430 feed. Here, the boresight direction of the feed is pointing through the geometric centre of the reflector which corresponds to an elevation tilt angle of approximately 28°. Although this is a non-optimum illumination angle, in actuality a larger elevation angles is used to improve the CATR QZ amplitude taper by compensating for the spherical loss factor, this value was used for the sake of consistency with prior simulations [1, 2]. Within Fig. 5, the white space corresponds to regions where the reflectivity of the reflector is zero. Figure 6 shows an image of the reflector once installed within the test chamber.
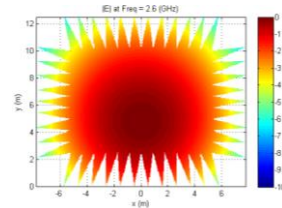
Fig. 5. Magnitude of incident electric field.

Fig. 6. Realised CATR main reflector.

The current element method [1, 2, 6] replaces fields with an equivalent surface current density $\underline{J}_s$ which is used as an equivalent source to the original fields. The surface current density across the surface of the reflector can be obtained from the incident magnetic fields and the surface unit normal using,

$$\underline{J}_s = 2\hat{\underline{n}} \times \underline{H}_i = 2\hat{\underline{n}} \times \underline{H}_r . \qquad (1)$$

The surface current density approximation for $\underline{J}_s$ (as embodied by the above expression) is known as the physical-optics approximation and allows for the computation of valid fields outside of the deep shadow region. The infinitesimal fields radiated by an electric current element can be obtained from the vector potential and the free-space Green's function [1, 6],

$$d\underline{H}(P) = \frac{da}{4\pi} \underline{J}_s \times \nabla \psi . \qquad (2)$$

This is an exact equation. When the field point is more than a few wavelengths from the radiating elemental source, the corresponding elemental electric fields can be obtained conveniently from the elemental magnetic fields using the far-field TEM condition using,

$$d\underline{E} = Z_0 \left(d\underline{H} \times \hat{\underline{u}} \right). \qquad (3)$$

Thus, both the electric and magnetic fields can be obtained from the elemental fields by integrating across

the surface of the parabolic reflector. In practice, for the case of a CATR with a QZ located at a distance z that is larger than the focal length of the reflector, the difference between the electric field as computed using the TEM condition and the exact formula is typically on the order of the limit of double precision arithmetic with this error being negligible. Figures 7 and 8 contain respectively false colour plots of the amplitude and phase patterns of the horizontally polarised electric field components of the pseudo-plane wave over the surface of a transverse plane located down-range at $z = 1.8f$, where f is the focal length of the CATR reflector at a frequency of 2.6 GHz. Figures 7, 8, 9 and 10 contain the E_x and E_y polarised amplitude and phase patterns for the horizontally polarised feed case. Although not shown, the equivalent magnetic fields were also computed. When interpreting these plots it is important to recognise that these are the fields one would measure if an infinitesimal electric (*i.e.*, Hertzian) dipole probe were used to sample the QZ fields [4]. This is in agreement with theory and standard CEM modelling tools. In practice, it is not possible to use an infinitesimal current element as a field probe and the following section examines how these patterns can be modified to include the effects of a finitely large, *i.e.*, directive, field probe.
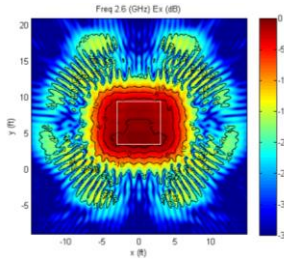
Fig. 7. E_x polarised QZ electric field amplitude.
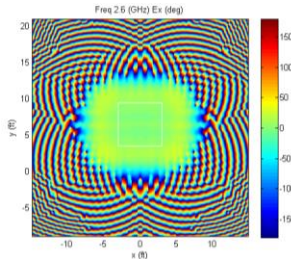
Fig. 8. E_x polarized QZ electric field phase.
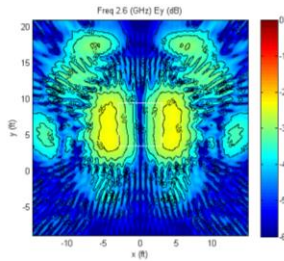
Fig. 9. E_y polarized QZ electric field amplitude.
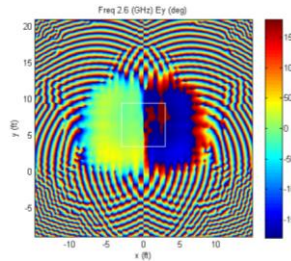
Fig. 10. E_y polarised QZ electric field phase.

III. CATR QZ PROBING SIMULATION

CATR QZ probing is usually accomplished by translating a field probe across a plane that is transverse to the z-axis of the CATR at several positions down-range. An example of a CATR QZ field probe can be seen presented in Fig. 11. Here, the electrically small field probe can be seen positioned at the limit of travel of the 6' linear translation stage. Generally, pyramidal horns, *e.g.*, *circa* 16 dBi standard gain horns (SGH) [6], are used as CATR QZ probes as they have excellent polarisation purity, are easy to align, have some gain and therefore provide some immunity from reflections from the side and back walls of the anechoic chamber. An alternative choice of field probe is a *circa* 6 dBi gain open ended rectangular waveguide probe (OEWG) [6].

Fig. 11. CATR QZ field being probed using a linear translation stage and a plane-polar acquisition scheme.

The clear difference in the electrical size of aperture of these two antennas and their directive properties and spatial filtering can be expected to result in some differences being observed between the probe measured QZ fields with the effects being quantifiable through an application of the reaction theorem which is a well-known method for analyzing general coupling problems [2, 4]. This theorem states that, provided the electric and magnetic field vectors ($\underline{E}_1$, $\underline{H}_1$) and ($\underline{E}_2$, $\underline{H}_2$) are of the same frequency and are monochromatic, then the mutual impedance, Z_{21}, between two radiators, *i.e.*, antennas 1 and 2, in the environment described by ε, μ can be expressed in terms of a surface integration [2, 4],

$$Z_{21} = \frac{V_{21}}{I_{11}} = -\frac{1}{I_{11}I_{22}} \int_{S_2} \left(\underline{E}_2 \times \underline{H}_1 - \underline{E}_1 \times \underline{H}_2 \right) \cdot \hat{\underline{n}} \, ds. \quad (4)$$

Here, $\underline{n}$ is taken to denote the outward pointing surface unit normal. The subscript 1 denotes parameters associated with antenna 1 whilst the subscript 2 denotes quantities associated with antenna 2, where the surface of integration encloses antenna 2, but not antenna 1. Here, I_{11} is the terminal current of antenna 1 when it transmits and similarly, I_{22} is the terminal current of antenna 2 when it transmits. Note that this integral does *not* compute transferred power as there are no conjugates present and as such, crucially, phase information is preserved. Here, the fields $\underline{E}_1$ and $\underline{H}_1$ are used to denote the CATR QZ whilst fields $\underline{E}_2$ and $\underline{H}_2$ denote fields associated with the QZ field probe. From reciprocity, the mutual impedance, $Z_{12} = Z_{21}$, is related to the coupling between the two antennas. Clearly the mutual impedance will also be a function of the displacement between the antennas, their relative orientations, their directivities and their respective polarization properties. Once the

impedance matrix is populated, this can be inverted to obtain the admittance matrix whereupon the required scattering matrix can be computed [4]. The elements $S_{1,2} = S_{2,1}$ of this two port scattering matrix are the complex transmission coefficients for the coupled antenna system which represent a single point in the quiet-zone probing measurement. Although the integration can be performed across any convenient free-space closed surface, in this application integrating across the planar aperture of the OEWG or SGH antenna is perhaps the most computationally efficient strategy. Aperture fields can be obtained from analytical models [4] as in this case, from CEM simulation [4] or from measurement with the choice being determined by the accuracy needed and the available information.

Figure 12 presents a comparison of the CATR QZ amplitude horizontal cut as obtained using an infinitesimal electric dipole (red trace) and an equivalent cut as obtained by using an OEWG probe (blue trace). A measure of the similarity between the respective measurements is provided by the equivalent multipath level (EMPL) [4] (magenta trace). From inspection of Figs. 12 and 13, it is evident that the ideal (dipole) and OEWG measurements are in very good agreement, both in amplitude and phase for the horizontal cuts. This is further confirmed by the EMPL level that is at or below -60 dB right across the pattern peak which corresponds to the useable QZ region.
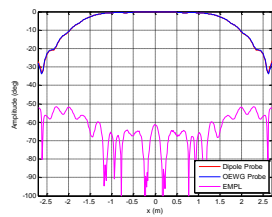
Fig. 12. Horizontal amplitude cut using dipole and OEWG field probe.
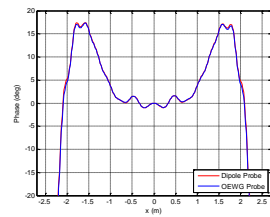
Fig. 13. Horizontal phase cut using dipole and OEWG field probe.

Figures 14 and 15 contain equivalent figures for the case where a SGH has been used as a pyramidal horn probe. Here it is evident from inspection of the amplitude and phase results that the high spatial frequency information within the QZ plots has been attenuated with the larger aperture effectively averaging out the measured response and thereby reducing the observed amplitude and peak-to-peak phase ripple. This is further confirmed by the circa 15 dB increase in the EMPL level between dipole probe and horn probe. Although not shown due to lack of space, equivalent results for the vertical cut exhibited similar phenomena. This probe dependent QZ is a well-known measurement effect but for the first time it has been possible to bound the SGH upper-bound measurement uncertainty and to provide tools necessary for verifying the appropriate choice of field probes.
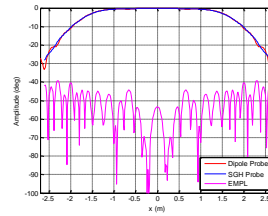
Fig. 14. Horizontal amplitude cut using dipole and SGH probe.
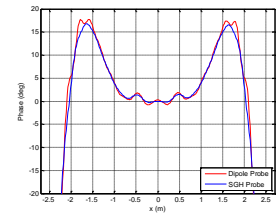
Fig. 15. Horizontal phase cut using dipole and SGH probe.

IV. CONCLUSION

The construction of a complete end-to-end CEM model of a CATR including CATR QZ probing has enabled the validity of standard CATR probing techniques to be objectively and quantitatively examined. Here, it was found that the standard practice of employing a pyramidal horn, e.g., a SGH, as a field probe increases EMPL by *circa* 15 db by reducing peak-to-peak ripple across the probed QZ. This works also confirmed that an electrically small OEWG probe provides highly accurate measure of the QZ fields with an EMPL < -60 dB and in nearly all the range <-70 dB. As this paper details ongoing research, the planned future work is to include obtaining additional verification of the modeling technique using the alternative plane-wave spectrum scattering matrix representation of antenna-to-antenna coupling.

REFERENCES

[1]   C. G. Parini, R. Dubrovka, and S. F. Gregson, "CATR quiet zone modelling and the prediction of "measured" radiation pattern errors: comparison using a variety of electromagnetic simulation methods," *AMTA*, October 2015.

[2]   C. G. Parini, R. Dubrovka, and S. F. Gregson, "Compact range quiet zone modelling: quantitative assessment using a variety of electromagnetic simulation methods," *LAPC*, November 2015.

[3]   G. E. Hindman and D. Slater, "Anechoic chamber diagnostic imaging," *AMTA Symposium*, 1992.

[4]   S. F. Gregson, C. G. Parini, and J. McCormick, *Principles of Planar Near-Field Antenna Measurements*, IET Press, 2007.

[5]   P. N. Betjes, "An algorithm for automated phase center determination and its implementation," *AMTA Symposium*, 2007.

[6]   C. G. Parini, S. F. Gregson, J. McCormick, and D. Janse van Rensburg, *Theory and Practice of Modern Antenna Range Measurements*, IET Press, 2014.

[7]   M. Philippakis and C. G. Parini, "Compact antenna range performance evaluation using simulated pattern measurements," *IEE Proceedings Microwaves, Antennas and Propagation*, vol. 143, iss. 3, pp. 200-206, 1996, DOI: 10.1049/ip-map:19960398

[8] C. G. Parini and M. Philippakis, "The use of quiet zone prediction in the design of compact antenna test ranges," *IEE Proc., Microwave Antennas Propagation*, vol. 143, no. 3, pp. 193-199, 1996.

[9] C. Cappellin, S. Busk Sørensen, M. Paquay, and A. Østergaard, "An accurate and efficient error predictor tool for CATR measurements," 4th EuCAP 2010, Barcelona, 12-16 April 2010.

Enhanced Artificial Immune System Algorithm and Its Comparison to Bio-Inspired Optimization Techniques for Electromagnetics Applications

Ozlem Kilic and Quang M. Nguyen

Department of Electrical Engineering and Computer Science
The Catholic University of America, Washington, DC, 20064, USA
kilic@cua.edu, 93nguyen@cardinalmail.cua.edu

Abstract ─ This paper introduces an enhanced artificial immune system algorithm (EAIS) that benefits from a hybrid approach by integrating concepts from the genetic algorithm (GA) and particle swarm optimization (PSO). The potential of the EAIS algorithm is demonstrated by comparing its performance with other bio-inspired optimization algorithms; namely the particle swarm optimization (PSO) and the conventional artificial immune system (AIS) when applied to two electromagnetics applications, such as the design of antireflective surfaces, and microstrip electromagnetic band gap (EBG) structures.

Index Terms ─ Antireflective surface, artificial immune system (AIS), bio-inspired optimization, enhanced artificial immune system (EAIS), microstrip (EBG), particle swarm optimization (PSO).

I. INTRODUCTION

Bio-inspired optimization techniques typically rely on a number of agents that simultaneously sample the optimization space in a random fashion. The process is iterated until a desired solution or the maximum number of iterations is reached. The initial step is purely random, hence requires no a-priori guess of the solution. More intelligence is added to the heuristic search at each iteration by taking advantage of the accumulated knowledge of the search domain among the agents. This intelligence is based on the computation of a cost function, which is a measure of how well each agent has performed with respect to the desired state; with high costs referring to poor solutions. The basic principles of bio-inspired optimization methods are shown in Fig. 1.

The application of bio-inspired optimization techniques to engineering problems is not a new concept. One of the well-known and original algorithms of this nature is the genetic algorithm (GA), which is based on the genetic recombination and mutation of species, [1]. Some other bio-inspired optimization techniques include the particle swarm optimization (PSO), which is inspired by the intelligent search of bees in a swarm to find the best food source in a field, [2], and the artificial immune system (AIS), which is based on the behavior of our

immune system in defending our body against viruses by adapting to the optimal antibody for a given antigen, [3]. Another recent bio-inspired algorithm is the covariance matrix adaption evolutionary strategy, which was applied to the design of linear arrays in [4]. Among these methods, GA and PSO have been applied to electromagnetics problems before. However, the application of AIS in electromagnetics have been relatively scarce. Most uses for these algorithms have been in the area of networks, resource constructed scheduling, data mining, etc. [5]. Our group has investigated AIS in the context of radar absorbing material design, and multi-beam satellite antenna side lobe control in [6]. In this paper, we further enhance the AIS algorithm (EAIS) and compare its performance to the conventional AIS and PSO for antireflective surfaces, and microstrip electromagnetic band gap (EBG) structures. We show that the consistent success of the enhanced AIS algorithm makes it a promising tool for the electromagnetics community.
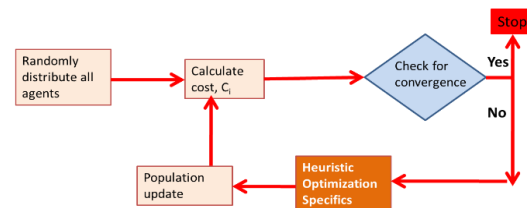
Fig. 1. A general block diagram of bio-inspired optimization methods.

The rest of the paper is organized as follows. Section 2 describes the principles of EAIS. Section 3 investigates the performance of AIS, PSO and EAIS when they are applied to electromagnetics designs. Finally, Section 4 concludes the paper.

II. ENHANCED ARTIFICIAL IMMUNE SYSTEM ALGORITHM (EAIS)

The conventional AIS optimization is based on the clonal selection principles of our immune system

response to potential disease generating metabolisms, and simulates human body's defense system against viruses. Our adaptive immune system produces antibodies whose purpose is to bind to any antigen that it recognizes. For engineering applications, antibodies represent a possible solution to the optimization problem. The optimization space is discretized in order to emulate the binary form of gene behavior. The generic "Heuristic Optimization Specifics" step shown in Fig. 1 is replaced by four steps in AIS: cloning, mutation, combination and sorting as shown in Fig. 2.

Fig. 2. Conventional AIS procedure steps.

In challenging problems, AIS can reach a state of stagnation where all good solutions in the set may differ by only a few bits [6]. In this paper, we incorporated modifications to the conventional AIS algorithm to improve its performance by bringing more intelligence to the mutation stage as well as by introducing concepts from other evolutionary algorithms. The EAIS specific procedures are shown in Fig. 3, where the dark shaded boxes indicate new steps to the process, and the light shade for the mutation step indicates a modification.

Fig. 3. Enhanced AIS procedure steps.

We modify the heuristic optimization specifics in AIS by modifying the mutation concept, and introducing a cross over operation. We also revise the population update process. The mutation stage benefits from the velocity and position update mechanism of PSO. The new cross over stage enhances the randomness using concepts from GA. Finally the population update stage utilizes PSO's global memory principles. The details of each enhancement stage are provided as follows.

(i) Mutation Stage Enhancements
Mutation enables AIS to randomly explore the search space by allowing it to move out of a local optimum and avoid stagnation. The conventional AIS algorithm carries this out by randomly flipping a certain number of bits in the solution set. The only intelligence incorporated in the conventional AIS at this stage is the number of bits to be flipped, with good solutions going through fewer flips than poor solutions. The proposed enhancement introduces a process similar to the velocity update mechanism of the PSO, where the bees get pulled towards the global best value. The value of the new

antibody Ab^* after the mutation process is computed as in Equation (1):

$$Ab^* = Ab + \alpha \times rand \times Ab + \alpha \times rand \times (Ab - Ab_{best}), \quad (1)$$

where Ab_{best} is the value of the best solution in the set, Ab is the value before mutation, and α is the mutation rate which exponentially varies as a function of the rank of the antibody, i.e., $\alpha = e^{-1/k}$ and *rand* is a random number between [0-1]. This process influences the mutation towards the best solution in the set, [7].

(ii) Crossover Stage Enhancements
Crossover is a process where a new binary set is produced from the existing set by randomly combining portions from different solutions. It is used in GA to create "children" from "a parent." This concept is applied to EAIS to create a new set of antibodies by crossing the antibodies created as a result of the cloning and mutation stages. The modified algorithm selects the top N_β antibodies from the output of the conventional AIS, i.e., the combined set of cloning and mutation operations. Each antibody is split into 2 segments. The number of bits in each segments is defined by the user, based on the cross-over split ratio $n_1 : n_2$ where n_1 denotes the number of bits with the most significant bits in the string, and n_2 denotes the number of remaining bits. The two segments are randomly crossed among each other to create a new set of N_x antibodies, where $N_x = \rho_x \times N_a$ is the size of the cross-over set, and ρ_x is the cross over ratio. The concept of crossover is demonstrated in Fig. 4. Finally, the crossover set is combined with the cloned and mutated sets. The antibodies are sorted one more time with respect to their cost values. The top N_a antibodies are selected from this set to start over again.
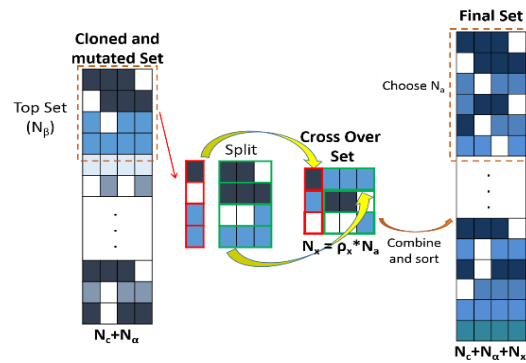
Fig. 4. Crossover concept of EAIS.

(iii) Population Update Stage Enhancements
At this point, the original population has been replaced with the new set of (N_a) antibodies utilizing cloning, mutation and cross over. The AIS algorithm can

continue to iterate these steps to search for the solution. In challenging problems, AIS can reach a state of stagnation where all good solutions in the set may differ by only a few bits. Another enhancement inspired by PSO is implemented at this stage. The solution set of N_a antibodies is split into three sets for separate treatments. The top few solutions from Set 1 are kept in order to remember the best solutions achieved. This is similar to the group memory concept in PSO in the context of the global best term. The rest of the antibodies are divided into two groups, i.e., Sets 2 and 3. Antibodies in Set 2 are replaced by a local random search in the vicinity of the best solution. Set 3 is replaced by a global search carried out randomly in the entire optimization space. The range of the local search can be adjusted dynamically to focus more in the vicinity of the best solutions as the number of iterations increase. The concept of these three sets and how they are treated are shown in Fig. 5. For the local search in Set 2, the approach is to utilize PSO with only a few agents and for a few iterations. The remaining solutions in Set 3 are simply replaced randomly. It should be noted that the cost function computation time is the dominant term for these random search optimization algorithms. Thus, for an equivalent problem; i.e., the same number of cost function calls per iteration, EAIS computation time is similar to AIS regardless of the few extra steps added.
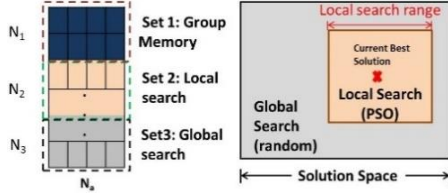
Fig. 5. Population update process of EAIS.

III. APPLICATION TO ELECTROMAGNETIC PROBLEMS

A. Antireflective Surface (AR) design

The first application we consider addresses an anti-reflective surface design, which is useful for both military (e.g., camouflaging) and commercial applications (e.g., efficient collection of energy such as in microwave or millimeter wave lenses). The design is inspired by the moth-eye structure, which consists of multiple layers of periodic gratings that enable the absorption of light at a wide range of incidence angles, [8]. In our application, we use a dielectric slab with an inverse moth eye pattern (i.e., holes rather than protrusions) applied to both top and bottom surfaces of the structure. We use the rigorous coupled wave algorithm (RCW) to simulate the model. Our goal is to achieve a desired total reflection coefficient of $\Gamma \le -30\,dB$ within the Ka-band over the 32-38 GHz bandwidth.

The design has an infinitely thick (i.e., half space) substrate with a dielectric constant of $\varepsilon_r = 2.56$, and an index of refraction of $n_s = 1.6$. The holes are backfilled with air, $n_h = 1$ and the grating period, Λ, is fixed at 3.1 mm, as depicted in Fig. 6. The grating heights, h_1 and h_2, and the hole diameters, d_1 and d_2, are the variables of the design. The incident field is assumed to be a TE polarized plane wave obliquely incident on the surface. The cost function of EAIS is defined as the square of a difference between the desired criteria and the reflection coefficient. The parameters of the optimization algorithms are summarized in Table 1.
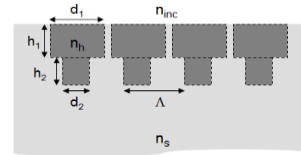
Fig. 6. Inverse half space AR surface design.

Table 1: Algorithm parameters for AR surface design

	PSO	AIS	EAIS
	$c_1 = c_2 = 1$	$\rho_c = 0.6$ $\rho_m = 0.5$	$\rho_c = 0.6,\ \rho_x = 1$ $n_1 / n_2 = 0.2$
N_a	96	28	18
N_b	n/a	12	12

The bandwidth was sampled in steps of 1 GHz. The best solution achieved at the end of 100 iterations is shown in Table 2 along with the best minimum reflection coefficients over the 32-38 GHz bandwidth. The optimized reflection coefficients is plotted in Fig. 7 (a) as a function of the incidence angle at the center frequency, i.e., f = 35 GHz for the three algorithms. We observe that for this particular angle, AIS does not converge at $\theta=40^0$. Also, overall it is evident that PSO has reached a better solution for a wide range of incident angles. However, it should be noted that achieving a value better than the desired criteria, i.e., $\Gamma \le -30\,dB$, is not part of the objectives for any of these algorithms, i.e., each algorithm accepts any value of $\Gamma \le -30\,dB$ as a converged solution. From the perspective of the rate of convergence, EAIS performs the best as shown in Fig. 7 (b), where Γ is plotted against the number of iterations.

Table 2: AR surface design solution

AR Design	PSO	AIS	EAIS
h_1 (mm)	2.58	2.67	2.60
h_2 (mm)	1.82	1.74	1.80
d_1 (mm)	3.02	3.01	3.02
d_2 (mm)	2.15	2.09	2.15
Γ (dB)	-28.62	-28.52	-28.62

(a)   (b)
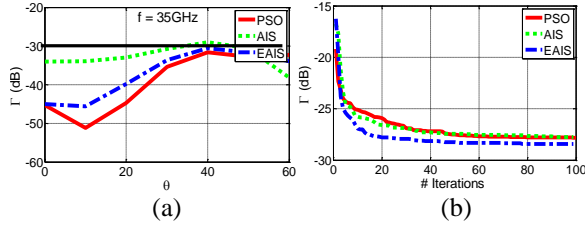
Fig. 7. Optimized AR surface design at f = 35 GHz and performance of three algorithms.

(a)   (b)

Fig. 9. Optimized microstrip EBG structure and performance of three algorithms.

B. Microstrip EBG structures

In this application, we design a low pass filter operating at [4-7.5 GHz] such that $S_{21} \geq$ -5 dB across [4-6 GHz], and $S_{21} \leq$ -25 dB over [6.5-7.5 GHz]. To achieve the performance, we consider a 50 Ohm microstrip line symmetrically residing on a periodically etched ground plane, where 50 mil thick, 2.33x1.53 inches RT/Duroid 6010 ($\varepsilon_r = 10.2$) is used as the substrate. To model the structure, we use EM commercial software, FEKO, which is based on the Method of Moments (MoM).

The simulated model for our design is shown in Fig. 8. The optimization variables; i.e., D_1 and D_2 are set to be between 50-250 mils. The optimization parameters used for each algorithm are summarized in Table 3. The solution achieved by each algorithm is summarized in Table 4, and the achieved performance is plotted in Fig. 9. We observe that, all three algorithms converge to similar solutions. A plot of the best cost versus the number of iterations is shown Fig. 9 (b). Although EAIS reaches the vicinity of the desired solution fastest, PSO achieves the best cost in the end.

IV. CONCLUSION

The robustness of the EAIS algorithm was demonstrated in comparison to PSO and AIS as they were applied to two electromagnetics applications; antireflective surface design using RCW, and microstrip electromagnetic band gap structure using MoM. EAIS consistently performed more robust than the other two algorithms. While it can never be claimed that a particular random search algorithm will always perform better than others, as the nature of a problem might fit the principles of an algorithm better at times, EAIS was shown to be consistently robust presenting itself as a viable tool for challenging electromagnetics problems.

REFERENCES

[1] R. L. Haupt, "An introduction to genetic algorithms for electromagnetics," *IEEE Antennas Propagat. Mag.*, vol. 37, pp.7-15, 1995.

[2] J. Kennedy and R. Eberhart, "Particle swarm optimization," *Proc. IEEE Int. Conf. Neural Networks*, vol. 4, pp. 1942-1948, 1995.

[3] L. N. De Castro and F. J. Von Zuben, "Learning and optimization using the clonal selection principle," *IEEE Trans. Evol. Comput.*, vol. 6, pp. 239-251, 2002.

[4] M. D. Gregory and D. H. Werner, "Design of high performance compact linear ultra-wideband arrays with the CMA evolutionary strategy," *IEEE Int'l Symp. on Ant. and Propag.*, Toronto, ON, 2010.

[5] R. S. Parpinelli, H. S. Lopes, and A. A. Freitas, "Data mining with an ant colony optimization algorithm," *IEEE Trans. Evol. Comp.*, vol. 6, iss. 4, pp. 321-332.

[6] O. Kilic and Q. Nguyen, "Application of artificial immune system algorithm to electromagnetics problems," *PIER B*, vol. 20, pp. 1-17, 2010.

[7] S. A. P. Ramaswamy, G. K. Venayagamoorthy, and S. N. Balakrishnan, "Optimal control of class of non-linear plants using artificial immune systems: application of the clonal selection algorithm," *22nd IEEE Int'l Symp. on Intelligent Control*, pp. 249-254, 1-3 Oct. 2007.

[8] M. S. Mirotznik, B. L. Good, P. Ransom, D. Wikner, and J. N. Mait, "Broadband antireflective properties of inverse Motheye surfaces," *IEEE Trans. Antennas Propagat.*, AP-58, 9, pp. 2969-2980, Sep. 2010.

(a) Top view   (b) Bottom view

Fig. 8. Microstrip EBG line with etched squares in the ground plane.

Table 3: Algorithm parameters for microstrip EBG

	PSO	AIS	EAIS
N_a	10	4	2
N_b	n/a	12	12
# Cost computations/iteration	10	10	10
N_{max}	14	14	14

Table 4: Microstrip EBG design solution

	PSO	AIS	EAIS
D_1 (mil)	50	55	58
D_2 (mil)	222	216	213

Experimental Benchmarking of Unstructured Transmission Line Modelling (UTLM) Method in Modelling Twisted Wires

Xuesong Meng*, Phillip Sewell, Nur H. A. Rahman, Ana Vukovic, and Trevor M. Benson

George Green Institute for Electromagnetics Research
University of Nottingham, Nottingham, NG7 2RD, UK
*xuesong.meng@nottingham.ac.uk

Abstract ─ In this paper the Unstructured Transmission Line Modelling (UTLM) method based on a tetrahedral mesh has been applied to modelling of the coupling between a single wire and a twisted wire pair. The effects of wire twisting on the crosstalk and coupling between wires are modelled by explicitly meshing wire geometries; simulation results are compared with experimental ones. Excellent agreement between simulated and measured results validates the viability and accuracy of the UTLM method and indicates the potential of the UTLM method for modelling complex wire structures.

Index Terms ─ Crosstalk, experimental benchmarking, Transmission Line Modelling (TLM) method, twisted wires, twisting effects, Unstructured Transmission Line Modelling (UTLM) method.

I. INTRODUCTION

Wires and cables play an important role in modern electronic systems, especially in the aerospace and automotive industries. They transmit signals between pieces of equipment and at the same time couple with the ambient electromagnetic fields. The electromagnetic interference (EMI) between wires and cables may affect the normal operation of equipment. It is therefore important to develop efficient and versatile methodologies that can predict the coupling strength between wires and cables.

The Unstructured Transmission Line Modelling (UTLM) method, fully presented and validated in [1-2], is a Transmission Line Modelling (TLM) method based on tetrahedral meshes [3]. Whilst the use of a TLM method based on structured cuboidal meshes has been widely reported, especially for Electromagnetic Compatibility (EMC) studies and microwave modelling, the advantages and characteristics of the UTLM method have not as yet been fully explored for such a variety of applications. For instance, when modelling curved structures, a very fine mesh is needed for the cuboidal mesh based TLM method to approximate the curved boundaries. This not only leads to large computational resources, even when using Octree sub-division technique [4], but also leads to stair-stepping problems [5]. Furthermore, its accuracy in representing the curved boundaries is piece-wise constant. In contrast, the UTLM method can better describe arbitrary shaped geometries, especially those involving curved structures, with no stair-stepping approximations and with a piece-wise-linear accuracy. One key feature of UTLM is the wide dynamic range of cell sizes that can be used. Importantly, a clustering technique whereby clusters of very small cells are grouped into larger cell entities for which the scattering is done implicitly, has been integrated into the UTLM method to allow a practical small time step to be used in a simulation [6]. These characteristics make the UTLM method a very good candidate to model in detail the coupling between wires and cables, a feature which is especially important within an aerospace context.

Some initial experimental benchmarking of a UTLM model explicitly meshing wires has been presented in [7], where the UTLM method was applied to the simulation of a canonical two parallel wire coupling problem. The paper discussed that small wire diameters lead to large computational costs in a discretized numerical simulation. Although an embedded thin wire model [8] has been widely adopted to reduce computational costs, it has difficulties in dealing with twisted wires, especially when other structures are in close proximity to the wire geometries. The purpose of this paper is to demonstrate how the wide dynamic range of cell sizes within ULTM enables the modelling of the coupling between a single wire and a twisted wire pair by explicitly meshing them, and to validate the approach by comparison with experiments. In the scenario considered the single wire is used for excitation and the twisted wire pair is used as victim on which crosstalk is observed. The crosstalk between wires is analyzed for different wire terminations, and both with and without twisting of the wire pair. The simulated results are compared with measurements. The paper is organized as follows: in the next section the experimental set up is described, followed by a description of the simulated model. Section IV outlines the comparison of the crosstalk between simulated and measured results and

Section V summarizes the main conclusions of the paper.

II. PROBLEM DESCRIPTION

In this section, the experimental setup used to study the twisting effects on coupling between wires is described.

In order to consider the twisting effects, a single core copper wire is excited as the source and a twisted wire pair is used as the victim. The single core wire has a radius of $r_w = 0.04$ cm. The separations between ports are $s1 = 2$ cm and $s2 = 1$ cm. The wires are placed near a L shape ground plane as shown in Fig. 1 (a) and mounted on to the two metal bulkheads using SMA bulkhead connectors as shown in Fig. 1 (b). The wires are situated above the bottom ground plane by a height, $h = 8$ cm. The distance between port 1 and the left ground plane is $d = 12.5$ cm as shown in Fig.1 (a). The length of the wires is $L = 100$ cm. The twisted wire pair has 20 twists along its length. The metal bulkheads each have dimensions of 30 cm $\times$ 30 cm.
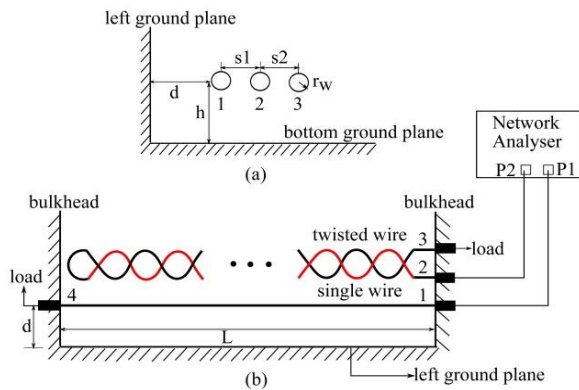
Fig. 1. (a) A single wire and a twisted wire pair are near the L shape ground plane; (b) the top view of the setup.

In the experiment, the twisted wire pair is made by folding a single core copper wire of length *2L* and twisting the two halves to create the configuration shown in Fig. 1 (b). At the right end, one wire of the twisted wire pair is connected to port 2 and the other one is connected to port 3. Port 1 is connected to the right end of the single wire as an excitation. Measurements are made on port 2 using a Network Analyzer (Agilent E5062A). Port 3 and port 4 (the left end of the single wire) are connected to a load, which could be a 50 ohms load, a short circuit or an open circuit.

III. NUMERICAL MODEL

The problem defined in Section II is described for numerical modelling purposes using University of Nottingham (UoN) in-house geometry software that provides a triangulated surface representation of the structure. The single wire is built as a metal cylinder.

Both ends of the single wire are connected to the cores of two coaxial probes, respectively. One of them is modelled as port 1 in Fig.1 (b) to excite the fundamental TEM mode and the other one is modelled as port 4 to terminate the wire.

The twisted wire pair is built by twisting two metal cylinders using a bifilar helix model. The circular cross-section of each wire is discretized by a N_1-sided polygon (where N_1 is an integer number). Each wire is then represented by N_2 piecewise linear segments (N_2 is also an integer number) that follow a helical path rotating around the axis of the twisted wires in such a manner that each wire's cross section remains perpendicular to its own axis. Figure 2 shows one example of a two wire twisted pair built with the UoN in-house geometry software.

Fig. 2. The geometrical model for a two wire twisted pair.

At one end of the twisted wire pair, the wires are connected using a metal cylinder. The wires follow a curved route defined by a Bezier curve. At the other end, the two wires are connected to the cores of two coaxial probes, respectively. One of the coaxial probes is modelled as port 2 in Fig. 1 (b) to observe the coupling to the TEM mode of the twisted wire pair; the other one is modelled as port 3 in Fig. 1 (b) to terminate the wire.

The structure is meshed using our UoN in-house Delaunay Mesher software as a hybrid tetrahedral-cubic mesh. The meshed structure is shown in Fig. 3.
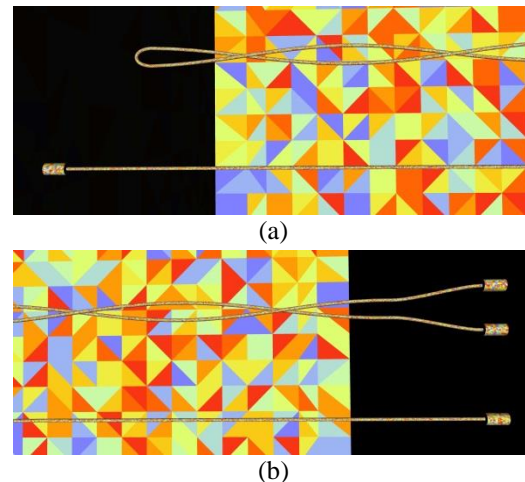
Fig. 3. The meshed geometry showing the triangulated interfaces between different materials: (a) left end of the structure and (b) the right end of the structure with a mesh size of 1 cm. Randomized colouring is used to show the triangle sizes and shapes.

Figures 3 (a) and (b) show the left and right ends of the structure as in Fig. 1 (b), respectively. The wires are meshed using a tetrahedral mesh, which is better able to describe curved structures, and the surrounding environment is meshed using a cuboidal mesh with a mesh size of 1 cm. Although the mesh is very small around the wires, the cell clustering enabled a time step of 0.08 ps to be used in the simulation [6]. The total number of time steps is 2,000,000. The success of meshing twisted wires with such small diameters indicates the great potential of the UTLM method in dealing with complex wire structures.

IV. CROSSTALK BETWEEN WIRES

In this section, the twisting effects on the coupling between wires are discussed for different termination conditions, using the UTLM simulation and experimental measurements.

The crosstalk between wires is described using the S_{21} parameter. In order to account for the influence of the twisting of the wires, the crosstalk between the single wire and a pair of parallel wires is also simulated and measured.

Figure 4 compares the UTLM simulation and experiment results for the crosstalk between the excitation single wire and the pair of twisted wires and the crosstalk between the single wire and the pair of parallel wires for different port 3 and 4 terminations, namely, (a) a short circuit, (b) a 50 ohms load and (c) an open circuit. The UTLM simulation results show very good agreement with those from experiment for all three terminations, so validating the accuracy of the UTLM simulations including the twisted wires.

The twisting effects on the coupling between wires can also be observed from Fig. 4. For short circuit and 50 ohms load terminations, the coupling between wires is reduced greatly in the relatively low frequencies. For example, at 1 MHz, a 20 dB crosstalk reduction for short circuit termination and a 17 dB crosstalk reduction for a 50 ohms load termination are observed in Figs. 4 (a) and (b). As the frequency increases, the crosstalk reduction becomes smaller. For open circuit termination, the twisting does not have any significant effect on the coupling between wires.

The twisting effects on the coupling between wires can be explained as follows. In the relatively low frequencies, the coupling between wires is the combination of inductive coupling and capacitive coupling as shown in [9]. Twisting the wires mainly reduces the inductive coupling, while it has no effect on the capacitive coupling. For low impedance terminations, the inductive coupling dominates the capacitive coupling, so the twisting reduces the total coupling; for high impedance terminations, the capacitive coupling dominates the inductive coupling, so the twisting has no significant effect on the total coupling.

(a)

(b)

(c)
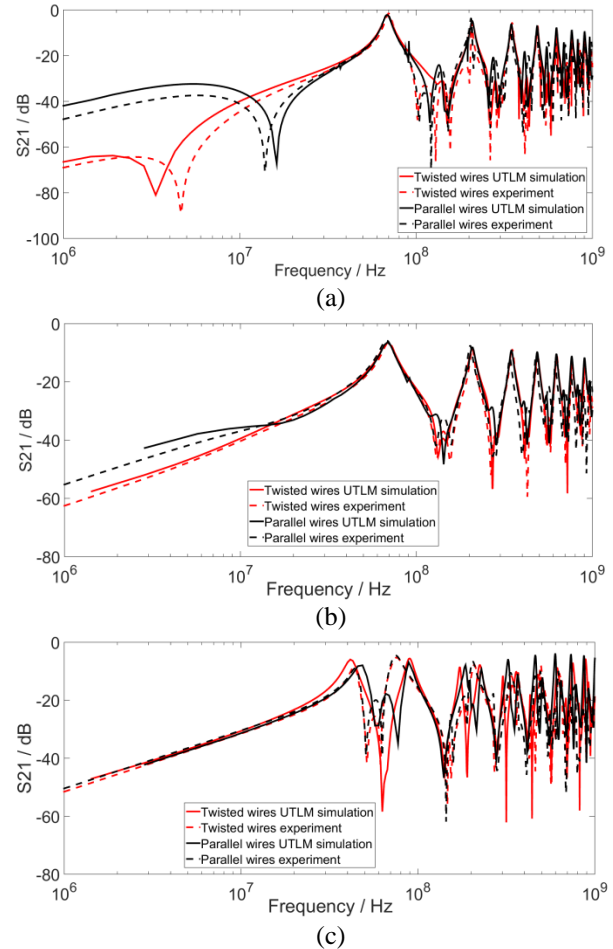
Fig. 4. Comparison of the UTLM simulation and experiment results for the crosstalk between the single wire and the pair of twisted wires and between the single wire and the pair of parallel wires, when ports 3 and 4 are terminated with: (a) a short circuit, (b) a 50 ohms load, and (c) an open circuit.

It is also noted that in Fig. 4 (a), for short circuit termination, there is a big discrepancy in the frequency at which the first dip in the crosstalk between the single wire and the twisted wire pair occurs. This can be explained since, for very low impedance (short circuit), the crosstalk at relatively low frequencies is very sensitive to the twist [9]. Even a very small non-uniformity in the twist could lead to a big change in the crosstalk. In the experiment, although every effort was made to make the twist uniform along the length, it was still unlikely to be a perfectly uniform twist. In the simulation it is quite easy to incorporate a perfect twist. To explore this further, two experiments have been undertaken, in which twisted wire pairs were made from two identical wires with a twist rate of 20 twists/m. The S_{21} parameters measured for each case are shown in Fig. 5. It is seen that the two results agree well with each other

for the relatively high frequencies but not for the relatively low frequencies. The frequency at which the first dip in S_{21} occurs in each experiment is different. The same experiments were also conducted for the 50 ohms load and open circuit terminations; in both these cases, the two sets of experimental results agree very well with each other. In conclusion, it is impossible to precisely predict the coupling between the single wire and the twisted wire pair at very low termination impedance at relatively low frequencies because of the sensitivity to the exact twist. Nevertheless, the UTLM simulation is able to predict the coupling for very low termination impedance at relatively high frequencies and for low and high termination impedances over the whole frequency range, very well.
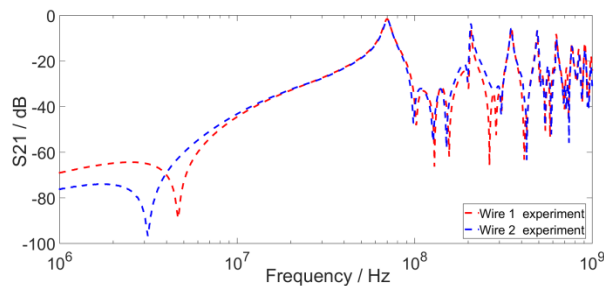
Fig. 5. The coupling between the single wire and the twisted wire pair with a short circuit termination; results from different experiments.

V. CONCLUSION

The Unstructured Transmission Line Modelling (UTLM) method has been successfully applied to model the coupling between a single wire and a pair of twisted wires by explicitly meshing the complete wire geometry. This is enabled by the wide dynamic range of mesh size that can be used within UTLM. The method shows powerful capability in meshing wires with small diameters within a large space. The accuracy of the UTLM method for this class of problems has been validated for the first time by comparing the simulated results for coupling between wires with experimental ones. The close agreement between the simulated and experimental results confirms that the UTLM method is a very useful and powerful tool that can be used for modelling complex wire structures.

ACKNOWLEDGMENT

REFERENCES

[1] P. Sewell, J. Wykes, T. Benson, C. Christopoulos, D. Thomas, and A. Vukovic, "Transmission-line modeling using unstructured triangular meshes," *IEEE Trans. on Microwave Theory and Techniques*, vol. 52, no. 5, pp. 1490-1497, 2004.

[2] P. Sewell, T. M. Benson, C. Christopoulos, D. W. P. Thomas, A. Vukovic, and J. G. Wykes, "Transmission line modeling (TLM) based upon unstructured tetrahedral meshes," *IEEE Trans. on Microwave Theory and Techniques*, vol. 53, pp. 1919-1928, 2005.

[3] C. Christopoulos, *The Transmission-Line Modeling Method TLM*. New York: IEEE Press, 1995.

[4] P. S. Duxbury, J. Wlodarczyk, and R. A. Scaramuzza, "The implementation and benefits of Octree staggered meshing in a TLM based EM simulation package," *2004 RF and Microwave Conference*, 2004.

[5] A. Cangellaris and D. Wright, "Analysis of the numerical error caused by the stair-stepped approximation of a conducting boundary in FDTD simulations of electromagnetic phenomena," *IEEE Trans. on Microwave Theory and Techniques*, vol. 39, pp. 1518-1525, 1991.

[6] P. Sewell, T. Benson, C. Christopoulos, D. Thomas, A. Vukovic, and J. Wykes, "Implicit element clustering for tetrahedral transmission line modeling (TLM)," *IEEE Trans. on Microwave Theory and Techniques*, vol. 57, no. 8, pp. 2005-2014, 2009.

[7] X. Meng, P. Sewell, A. Vukovic, Z. Zhang, and T. Benson, "Experimental benchmarking of unstructured transmission line modelling method (UTLM) simulations of explicitly meshed wiring," in *Computational Electromagnetics International Workshop (CEM) 2015*, pp. 1-2, 2015.

[8] P. Sewell, Y. K. Choong, and C. Christopoulos, "An accurate thin-wire model for 3-D TLM simulations," *IEEE Trans. on Electromagnetic Compatibility*, vol. 45, no. 2, pp. 207-217, 2003.

[9] C. R. Paul and M. B. Jolly, "Sensitivity of crosstalk in twisted-pair circuits to line twist," *IEEE Trans. on Electromagnetic Compatibility*, vol. 24, no. 3, pp. 359-364, 1982.

Four-Stage Split-Step 2D FDTD Method with Error-Cancellation Features

Theodoros T. Zygiridis [1], Nikolaos V. Kantartzis [2], and Theodoros D. Tsiboukis [2]

[1] Department of Informatics and Telecommunications Engineering
University of Western Macedonia, Kozani, 50100, Greece
tzygiridis@uowm.gr

[2] Department of Electrical and Computer Engineering
Aristotle University of Thessaloniki, Thessaloniki, 54124, Greece
kant@auth.gr, tsibukis@auth.gr

Abstract ─ We develop a methodology that enables the proper introduction of high-order spatial operators in an unconditionally-stable, split-step, finite-difference time-domain scheme. The proposed approach yields spatial approximations that guarantee better balancing of space-time errors, compared to standard fourth-order expressions. The latter are not as efficient as expected, due to their unmatched order with the scheme's second-order temporal accuracy. Our technique treats the dispersion relation as an error descriptor, derives spatial formulae that change with the cell shape and time-step size, and rectifies the performance over all frequencies.

Index Terms ─ Finite-difference time-domain (FDTD) methods, numerical-dispersion relation, split-step approaches, unconditionally-stable methods.

I. INTRODUCTION

Among the diverse advances of Yee's finite-difference time-domain (FDTD) method [1,2], approaches featuring unconditional stability [3] belong to the most significant contributions. Numerical schemes such as the alternating-direction implicit [4] and the locally one-dimensional FDTD techniques [5] are free from constraints governing their temporal sampling density, which is an advantageous property in many electromagnetic simulations. Other solutions based on split-step procedures [6-8] also exhibit similar behavior, and have been the subject of various studies [9,10].

In the context of the aforementioned implicit methods, the improvement of temporal accuracy can be a computationally expensive task, as it commonly requires the increase of the intermediate stages for the successive update of field components. On the other hand, direct incorporation of high-order spatial operators is a simpler and more straightforward approach towards performance upgrade, although it too augments the algorithm's complexity. Since the combination of accurate spatial approximations with low (first or second)

temporal order usually impedes the full exploitation of high-order operators' potential, amending techniques may be applied for further error mitigation. The implementation of constant-valued correctional coefficients, calculated in diverse ways, is a popular practice in this category of useful concepts [11,12].

This paper's purpose is to efficiently incorporate four-point spatial approximations into a two-dimensional (2D) four-stage split-step FDTD (SS-FDTD) method, aiming at a balanced treatment of space-time errors. Our approach exploits the scheme's dispersion relation to represent the inherent discretization errors. By using the estimator's Taylor polynomial, improvement over all frequencies is facilitated, while its trigonometric expansion leads to accuracy correction irrespective of propagation direction. The resulting unconditionally-stable algorithm performs better than its counterpart with standard high-order operators, verifying the optimal use of computational resources.

II. MODIFIED 4-STAGE SS-FDTD METHOD

The considered SS-FDTD scheme has second-order temporal accuracy, and the time-stepping is performed according to the following splitting approach:

$$\left([\mathbf{I}] - \tfrac{\Delta t}{4}[\mathbf{A}]\right)[\mathbf{u}]^{n+1/4} = \left([\mathbf{I}] + \tfrac{\Delta t}{4}[\mathbf{A}]\right)[\mathbf{u}]^{n}$$
$$\left([\mathbf{I}] - \tfrac{\Delta t}{4}[\mathbf{B}]\right)[\mathbf{u}]^{n+1/2} = \left([\mathbf{I}] + \tfrac{\Delta t}{4}[\mathbf{B}]\right)[\mathbf{u}]^{n+1/4}$$
$$\left([\mathbf{I}] - \tfrac{\Delta t}{4}[\mathbf{B}]\right)[\mathbf{u}]^{n+3/4} = \left([\mathbf{I}] + \tfrac{\Delta t}{4}[\mathbf{B}]\right)[\mathbf{u}]^{n+1/2} \quad , \tag{1}$$
$$\left([\mathbf{I}] - \tfrac{\Delta t}{4}[\mathbf{A}]\right)[\mathbf{u}]^{n+1} = \left([\mathbf{I}] + \tfrac{\Delta t}{4}[\mathbf{A}]\right)[\mathbf{u}]^{n+3/4}$$

where $[\mathbf{I}]$ is the 3×3 unitary matrix, $[\mathbf{u}] = [E_x \ E_y \ H_z]^\mathrm{T}$ is the vector with the three field components in 2D, Δt is the time increment, and $[\mathbf{A}],[\mathbf{B}]$ are derivative matrices:

$$[\mathbf{A}] = \begin{bmatrix} 0 & 0 & \tfrac{1}{\varepsilon}D_y \\ 0 & 0 & 0 \\ \tfrac{1}{\mu}D_y & 0 & 0 \end{bmatrix}, \ [\mathbf{B}] = -\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & \tfrac{1}{\varepsilon}D_x \\ 0 & \tfrac{1}{\mu}D_x & 0 \end{bmatrix}. \tag{2}$$

In its conventional form [8], the methodology relies on

standard second-order approximations of the derivatives. Here, we adopt four-point symmetric expressions,

$$D_x u\big|_{i,j}^n = \frac{1}{\Delta x}\sum_{m=1}^{2} C_m^x \left(u\big|_{i+\frac{2m-1}{2},j}^n - u\big|_{i-\frac{2m-1}{2},j}^n\right), \quad (3)$$

$$D_y u\big|_{i,j}^n = \frac{1}{\Delta y}\sum_{m=1}^{2} C_m^y \left(u\big|_{i,j+\frac{2m-1}{2}}^n - u\big|_{i,j-\frac{2m-1}{2}}^n\right), \quad (4)$$

whose final form is determined via an analytical procedure that aims at suitable error cancellation.

The basic element of the proposed optimization approach is the scheme's numerical dispersion relation, and its utilization as a means to express the inherent dispersion errors. The dispersion relation is obtained after introducing plane-wave forms in (1),

$$[\mathbf{u}] = [\mathbf{u}_0]e^{\mathrm{j}(\omega t - \tilde{k}_x x - \tilde{k}_y y)}, \quad (5)$$

and requiring the existence of non-trivial solutions for the resulting system ($\tilde{k}_x = \tilde{k}\cos\theta$, $\tilde{k}_y = \tilde{k}\sin\theta$, $\tilde{k}$ is the numerical wavenumber). In essence, the condition,

$$\det\left(e^{\mathrm{j}\omega\Delta t}[\mathbf{I}] - [\mathbf{L}_{\mathcal{A}}][\mathbf{L}_{\mathcal{B}}]^2[\mathbf{L}_{\mathcal{A}}]\right) = 0, \quad (6)$$

is obtained, where

$$[\mathbf{L}_{\mathbf{U}}] = \left([\mathbf{I}] - \tfrac{\Delta t}{4}[\mathbf{U}]\right)^{-1}\left([\mathbf{I}] + \tfrac{\Delta t}{4}[\mathbf{U}]\right), \quad (7)$$

and matrices $[\mathcal{A}]$, $[\mathcal{B}]$ are derived from $[\mathbf{A}]$, $[\mathbf{B}]$, respectively, after replacing the D_x, D_y operators with:

$$X = -\frac{2\mathrm{j}}{\Delta x}\sum_{m=1}^{2} C_m^x \sin\left(\tfrac{2m-1}{2}\tilde{k}_x\Delta x\right), \quad (8)$$

$$Y = -\frac{2\mathrm{j}}{\Delta y}\sum_{m=1}^{2} C_m^y \sin\left(\tfrac{2m-1}{2}\tilde{k}_y\Delta y\right). \quad (9)$$

The resulting dispersion equation has the form:

$$\cos(\omega\Delta t) = \frac{\alpha_{\mathrm{num}}(\tilde{k},\omega,\theta)}{\alpha_{\mathrm{den}}(\tilde{k},\omega,\theta)}, \quad (10)$$

where

$$\alpha_{\mathrm{num}} = \left(16 + c_0^2\Delta t^2 X^2\right)^2\left(16 + c_0^2\Delta t^2 Y^2\right)^2$$
$$+ 64c_0^2\Delta t^2\left(X^2 + Y^2\right)\left(256 + c_0^4\Delta t^2 X^2 Y^2\right), \quad (11)$$

$$\alpha_{\mathrm{den}} = \left(16 - c_0^2\Delta t^2 X^2\right)^2\left(16 - c_0^2\Delta t^2 Y^2\right)^2. \quad (12)$$

Of crucial importance is the definition of the error formula that is used to represent the discretization flaws. As we are interested in combating the inaccuracies pertinent to the phase velocity, we define,

$$\Lambda(\omega,\theta) = \cos(\omega\Delta t) - \frac{\alpha_{\mathrm{num}}(k,\omega,\theta)}{\alpha_{\mathrm{den}}(k,\omega,\theta)}, \quad (13)$$

which practically describes the deviation from the numerical dispersion relation, once the numerical wave-vector has been replaced by its exact value $k = \omega/c_0$. Now, the determination of the optimum spatial operators is reduced to the following problem: find suitable coefficients C_1^x, C_2^x, C_1^y, C_2^y, so that the magnitude of

$\Lambda(\omega,\theta)$ is rendered as close to zero as possible, for all frequencies ω and propagation angles θ.

In order to satisfy – to the best possible degree – the aforementioned requirements, the Taylor-series of (13) with respect to the spatial increment is exploited. Specifically, we are working on the expression:

$$\Lambda(\omega,\theta) = \delta^{(2)}(\theta)(k\Delta x)^2 + \delta^{(4)}(\theta)(k\Delta x)^4 + \ldots. \quad (14)$$

This expansion effectively isolates the dependences on frequency and propagation direction, which significantly facilitates their separate treatment. Specifically, accuracy improvement irrespective of frequency is now possible, by cancelling the corresponding δ coefficients, which do not depend on ω. If it was possible to accomplish $\delta^{(2)} = \delta^{(4)} = \ldots = 0$, a totally error-free FDTD scheme would be devised. Apparently, this is an observation of merely theoretical interest, since the discretization error can be controlled only to a certain degree in practice. In our case, we are proceeding with the manipulation of the $\delta^{(2)}$ and $\delta^{(4)}$ coefficients.

Starting from the second-order term, we find that:

$$\delta^{(2)} = \frac{Q^2 R^2}{2\left(1 + R^2\right)}\left[\left(C_1^x + 3C_2^x\right)^2\tau_x^2 + \left(C_1^y + 3C_2^y\right)^2\tau_y^2 - 1\right], \quad (15)$$

where $R = \Delta y/\Delta x$, $\tau_x = \cos\theta$, $\tau_y = \sin\theta$, and Q determines the time-step size, via:

$$\Delta t = \frac{QR\Delta x}{c_0\sqrt{1 + R^2}}. \quad (16)$$

In (16), $Q = 1$ yields the well-known Yee's stability criterion. Clearly, the second-order term vanishes if:

$$C_1^u + 3C_2^u = 1, \quad u = x,y, \quad (17)$$

In essence, (17) guarantees that the spatial operators are at least second-order accurate, which is necessary so that their error matches the corresponding temporal one.

The treatment of the $\delta^{(4)}$ term is more involved, as it is expressed according to:

$$\delta^{(4)}(\theta) = \frac{(QR)^4}{24\left(1 + R^2\right)^2} - \frac{(QR)^2}{3072\left(1 + R^2\right)^2}\Big[$$
$$128\left(C_1^x + 27C_2^x\right)\tau_x^4 + 128R^4\left(C_1^y + 27C_2^y\right)\tau_y^4$$
$$+ 128R^2\left(C_1^x + 27C_2^x\right)\tau_x^4 + 128R^2\left(C_1^y + 27C_2^y\right)\tau_y^4$$
$$+ 3Q^2R^2\left(64\tau_x^4 + \tau_x^3\tau_y + 128\tau_x^2\tau_y^2 + \tau_x\tau_y^3 + 64\tau_y^4\right)\Big], \quad (18)$$

where (17) has already been partially substituted for simplicity. It is clear that $\delta^{(4)}$ cannot be eliminated, as the corresponding SS-FDTD scheme cannot accomplish fourth-order space-time accuracy. A more realistic goal is to render $\delta^{(4)}$ as small as possible, so that better overall performance is attained, compared to standard fourth-order spatial approximations. For this reason, (18) is rearranged, using the identities:

$$\tau_x^4 = \frac{3}{8} + \frac{1}{2}\cos(2\theta) + \frac{1}{8}\cos(4\theta), \quad (19)$$

$$\tau_y^4 = \frac{3}{8} - \frac{1}{2}\cos(2\theta) + \frac{1}{8}\cos(4\theta), \quad (20)$$

$$\tau_x^3\tau_y = \frac{1}{4}\sin(2\theta) + \frac{1}{8}\sin(4\theta), \quad (21)$$

$$\tau_x^2\tau_y^2 = \frac{1}{8} - \frac{1}{8}\cos(4\theta), \quad (22)$$

$$\tau_x\tau_y^3 = \frac{1}{4}\sin(2\theta) - \frac{1}{8}\sin(4\theta). \quad (23)$$

Taking (19)-(23) into account, $\delta^{(4)}$ is expressed as a finite trigonometric series, and two additional constraints can be derived from the vanishing of an equal number of terms. If the series' constant term is set equal to zero, the following equation is obtained:

$$C_1^x + 27C_2^x + R^2\left(C_1^y + 27C_2^y\right) = -\frac{4(QR)^2}{3(1+R^2)}. \quad (24)$$

The last equation is extracted from the coefficient of the $\cos(2\theta)$ term, resulting in:

$$C_1^x + 27C_2^x - R^2\left(C_1^y + 27C_2^y\right) = 0. \quad (25)$$

The solution of the system comprising (17), (24), and (25) yields the optimum spatial operators, whose final form takes into account the cell shape and the time-step size, as the resulting coefficient expressions are:

$$C_1^x = \frac{9}{8} + \frac{Q^2}{12}\frac{R^2}{1+R^2}, \quad C_2^x = -\frac{1}{24} - \frac{Q^2}{36}\frac{R^2}{1+R^2}, \quad (26)$$

$$C_1^y = \frac{9}{8} + \frac{Q^2}{12}\frac{1}{1+R^2}, \quad C_2^x = -\frac{1}{24} - \frac{Q^2}{36}\frac{1}{1+R^2}. \quad (27)$$

Consequently, the procedure followed herein concludes that the standard fourth-order operators are the most suitable choice, only if a very small time step (i.e., $Q \rightarrow 0$) is selected. This, however, is not the case in unconditionally-stable FDTD methods, and modified approximations that do not necessarily preserve the maximum order of accuracy can guarantee lower dispersion flaws.

III. ASSESSMENT OF METHODOLOGY

The stability of the numerical scheme is revealed by obtaining the eigenvalues of its amplification matrix. The latter is equal to $[\mathbf{L}_\mathcal{A}][\mathbf{L}_\mathcal{B}]^2[\mathbf{L}_\mathcal{A}]$, as defined in (6), (7), and its eigenvalues are: $\lambda_1 = 1$, and,

$$\lambda_{2,3} = \frac{w - 64v \pm j16\sqrt{wv}}{r}, \quad (28)$$

where

$$w = \left(16 + c_0^2\Delta t^2 X^2\right)^2\left(16 + c_0^2\Delta t^2 Y^2\right)^2, \quad (29)$$

$$v = -64c_0^2\Delta t^2\left(X^2 + Y^2\right)\left(c_0^4\Delta t^4 X^2 Y^2 + 256\right), \quad (30)$$

$$r = -c_0^2\Delta t^2\left(16 + c_0^2\Delta t^2 X^2\right)^2\left(16 + c_0^2\Delta t^2 Y^2\right)^2$$
$$\times\left(X^2 + Y^2\right)\left(c_0^4\Delta t^4 X^2 Y^2 + 256\right). \quad (31)$$

It can be shown that the magnitude of $\lambda_{1,2,3}$ is 1,

regardless of the time-step size. Hence, the specific SS-FDTD updates are unconditionally stable. An exemplary plot of the eigenvalues on the complex plane is given in Fig. 1 when $R = 1$, the spatial density is 40 cells per wavelength, and $Q \leq 50$. As expected, all values lie on the circumference of the unit circle.
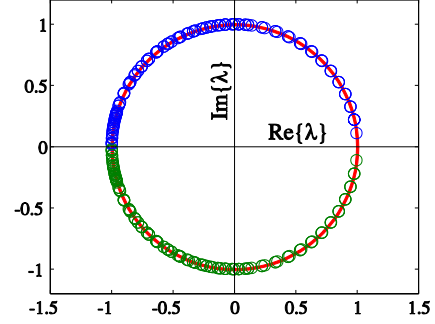
Fig. 1. Eigenvalues of the amplification matrix on the complex plane for various Δt.

Next, the algorithm's accuracy is assessed in terms of the dispersive error affecting the phase velocity, whose value $\tilde{c} = \omega/\tilde{k}$ is extracted from the dispersion relation (10), and the overall error is estimated from:

$$e_t = \frac{1}{2\pi c_0}\int_0^{2\pi}\left|c_0 - \tilde{c}\right|d\theta. \quad (32)$$

The positive impact of the modified operators on the scheme's accuracy is illustrated in Fig. 2, where the SS-FDTD method that applies standard fourth-order approximations is also depicted (time-steps five and ten times larger than Yee's stability limit are considered). The plotted curves reveal a significant error reduction that is not confined within specific frequency bands. In fact, the evidence show that a specific accuracy level can be now accomplished with twice as large a time-step, thanks to the sophisticated design of the algorithm, ensuring better utilization of computational resources.
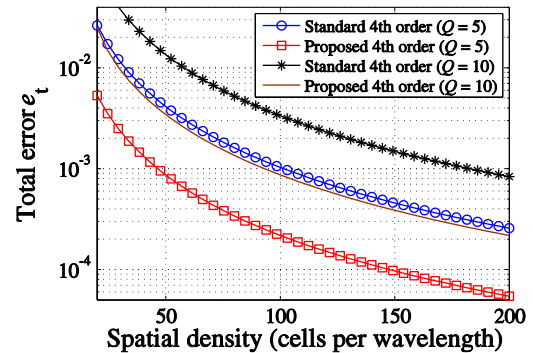
Fig. 2. Error e_t versus mesh density for different Δt.

IV. NUMERICAL RESULTS
The performance of the modified SS-FDTD algorithm

is evaluated considering an 8 cm × 6 cm cavity, bounded by perfectly conducting walls. First, we perform two sets of simulations, one for the TE_{11} mode at 3.123 GHz and one for the TE_{21} mode at 4.504 GHz. The computational space comprises 200×150 cells and tests are conducted for different time-step magnitudes. In essence, the maximum L_2 error of H_z is recorded for a time period equivalent to 2000 iterations when $Q = 2$. The results are displayed in Table 1, where the standard fourth-order operators and the proposed ones are compared. It is verified that significant error cancellation is accomplished regardless of the time-step size. Specifically, accuracy is improved by 3.5 times in the first case, and by 4.2 times in the second case, verifying the potential of the new derivative approximation.

Using the same configuration as previously, the second test pertains to the detection of the structure's first 24 resonant frequencies. Now, an 80×60 grid is used, and simulations for 32768 time-steps with $Q = 5$ are performed. Figure 3 displays the absolute errors in the frequencies of the detected modes, and the superior spectral properties of the proposed SS-FDTD method are clearly illustrated. Specifically, the average error of the standard scheme is 55.07 MHz, which is suppressed to only 7.31 MHz by the modified operators.

Table 1: Maximum L_2 errors for different time-step sizes, in the cavity problems with single-mode support

	TE_{11} Mode		TE_{21} Mode	
Q	Standard	Proposed	Standard	Proposed
2	$1.05 \cdot 10^{-3}$	$2.94 \cdot 10^{-4}$	$3.11 \cdot 10^{-3}$	$7.31 \cdot 10^{-4}$
4	$4.19 \cdot 10^{-3}$	$1.18 \cdot 10^{-3}$	$1.24 \cdot 10^{-2}$	$2.92 \cdot 10^{-3}$
6	$9.42 \cdot 10^{-3}$	$2.65 \cdot 10^{-3}$	$2.80 \cdot 10^{-2}$	$6.59 \cdot 10^{-3}$
8	$1.67 \cdot 10^{-2}$	$4.72 \cdot 10^{-3}$	$4.95 \cdot 10^{-2}$	$1.17 \cdot 10^{-2}$
10	$2.61 \cdot 10^{-2}$	$7.38 \cdot 10^{-3}$	$7.71 \cdot 10^{-2}$	$1.84 \cdot 10^{-2}$
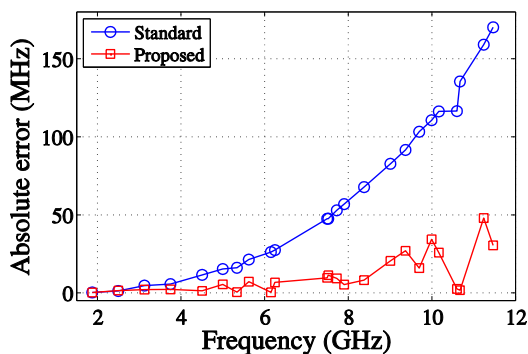
Fig. 3. Absolute error in detecting the resonant frequencies of a rectangular cavity.

V. CONCLUSION

We have successfully remedied the accuracy of an unconditionally stable SS-FDTD method, by deriving modified spatial operators with three-cell stencils. The form of these approximations is determined via a design procedure that balances space-time errors over all frequencies better than standard formulae. The modified scheme outperforms its conventional counterpart, as it guarantees similar error levels with larger time-steps.

REFERENCES

[1] K. S. Yee, "Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media," *IEEE Trans. Antennas Propag.*, vol. AP-14, pp. 302-307, 1966.

[2] A. Taflove and S. C. Hagness, *Computational Electrodynamics: The Finite-Difference Time-Domain Method*. Artech House, Norwood, MA, ed. 3, 2005.

[3] G. Sun and C. W. Trueman, "Accuracy of three unconditionally-stable FDTD schemes," *ACES J.*, vol. 18, no. 4, pp. 41-47, Nov. 2003.

[4] F. Zheng, Z. Chen, and J. Zhang, "A finite-difference time-domain method without the Courant stability conditions," *IEEE Microw. Guided Wave Lett.*, vol. 9, no. 11, pp. 441-443, Nov. 1999.

[5] J. Shibayama, M. Muraki, J. Yamauchi, and H. Nakano, "Efficient implicit FDTD algorithm based on locally one-dimensional scheme," *Electron. Lett.*, vol. 41, no. 19, pp. 1046-1047, Sept. 2005.

[6] J. Lee and B. Fornberg, "A split step approach for the 3-D Maxwell's equations," *J. Comp. Appl. Math.*, vol. 158, no. 2, pp. 485-505, Sept. 2003.

[7] W. Fu and E. L. Tan, "Development of split-step FDTD method with higher-order spatial accuracy," *Electron. Lett.*, vol. 40, no. 20, pp. 1252-1253, Sept. 2004.

[8] M. Kusaf and A. Y. Oztoprak, "An unconditionally stable split-step FDTD method for low anisotropy," *IEEE Microw. Wireless Comp. Lett.*, vol. 18, no. 4, pp. 224-226, Apr. 2008.

[9] A. Grande, J. A. Pereda, A. Serroukh, I. Barba, A. C. Cabeceira, and J. Represa, "Reinterpreting four-stage split-step FDTD methods as two-stage methods," *IEEE Trans. Antennas Propag.*, vol. 61, no. 11, pp. 5818-5821, Nov. 2013.

[10] D. Y. Heh and E. L. Tan, "Further reinterpretation of multi-stage implicit FDTD schemes," *IEEE Trans. Antennas Propag.*, vol. 62, no. 8, pp. 4407-4411, Aug. 2014.

[11] W. Fu and E. L. Tan, "A parameter optimized ADI-FDTD method based on the (2,4) stencil," *IEEE Trans. Antennas Propag.*, vol. 54, no. 6, pp. 1836-1842, June 2006.

[12] Q.-F. Liu, W.-Y. Yin, Z. Chen, and P.-G. Liu, "An efficient method to reduce the numerical dispersion in the LOD-FDTD method based on the (2,4) stencil," *IEEE Trans. Antennas Propag.*, vol. 58, no. 7, pp. 2384-2393, July 2010.

A 3-D Polynomial-Chaos FDTD Technique for Complex Inhomogeneous Media with Arbitrary Stochastically-Varying Index Gradients

Georgios G. Pyrialakos [1], Theodoros T. Zygiridis [2], and Nikolaos V. Kantartzis [1]

[1] Department of Electrical and Computer Engineering
Aristotle University of Thessaloniki, Thessaloniki, 54124, Greece
pyrialak@auth.gr, kant@auth.gr

[2] Department of Informatics and Telecommunications Engineering
University of Western Macedonia, Kozani, 50100, Greece
tzygiridis@uowm.gr

Abstract ─ An enhanced finite-difference time-domain algorithm featuring the polynomial chaos representation is introduced in this paper for problems with stochastic uncertainties. Focusing on the solution of the governing partial differential equations, the new 3-D method uses the Karhunen-Loève expansion to effectively decorrelate random input parameters denoted by stochastic processes. So, the space dimension is seriously reduced and high accuracy levels are attained, even for media with abrupt and fully unknown statistical variations. These profits are verified via a detailed numerical study.

Index Terms ─ Advanced FDTD methods, polynomial chaos, random media, stochastic process, uncertainties.

I. INTRODUCTION

The assessment of stochastic uncertainties, inherent in electromagnetics, has been of pivotal significance, so leading to different numerical schemes. Amid them the Monte Carlo (MC) approach [1], albeit accurate, has proven time-consuming, due to its excessively many realizations and slow convergence to the desired result. Recently, efficient techniques have been presented [2-6], such as the generalized polynomial chaos finite-difference time-domain (GPC-FDTD) algorithm [7, 8]. Usually, uncertainties emerge from discrete stochastic variables or processes that vary in an unknown way. Being difficult to model, the latter cannot be directly plugged in the GPC-FDTD technique, as they require an infinite number of correlated random variables.

To overcome such an issue, this paper develops a 3-D GPC-FDTD methodology for complex materials with arbitrary statistically-varying index gradients. The novel algorithm utilizes an orthogonal field expansion over the space of random parameters, so minimizing the error for both the mean value and variance. While it is primarily optimized for the extraction of the first two moments, the approximated relation between the output and stochastic parameter can provide with more valuable information like the computation of high-order statistical moments or maxima/minima, useful in electromagnetic compatibility (EMC) applications [9, 10]. To decrease the dimension of space spanned by the input parameters and decorrelate them, the Karhunen-Loève scheme is employed. It transforms the infinite product space of random inputs to be described by a new base that can be safely truncated. The resulting variables are, also, uncorrelated, which for Gaussian processes is equivalent to independence; an ample claim for the GPC applicability. Numerical outcomes certify our method, accelerated via graphics processor units (GPUs), and reveal its superiority.

II. PROPOSED METHODOLOGY

A. Generalized polynomial chaos expansion

The GPC method expands all fields in a summation of orthogonal, under an inner product, basis functions over space $\Omega = \cup \Omega_i$ ($i = 1, 2, \ldots, D$), spanned by all D random variables ω_i (each defined in Ω_i), on condition that they are statistically independent. The orthogonality is satisfied with respect to the inner product:

$$\left\langle \Theta_a(\boldsymbol{\omega}), \Theta_b(\boldsymbol{\omega}) \right\rangle \triangleq \int_{\Omega} \Theta_a(\boldsymbol{\omega}) \Theta_b(\boldsymbol{\omega}) w(\boldsymbol{\omega}) d\boldsymbol{\omega} = \left\langle \Theta_a(\boldsymbol{\omega}), \Theta_a(\boldsymbol{\omega}) \right\rangle \delta_{a,b}, \tag{1}$$

with $\Theta_{a,b}(\boldsymbol{\omega})$ the basis functions, $\boldsymbol{\omega} = [\omega_1, \omega_2, \ldots, \omega_D]^{\mathrm{T}}$ the vector formed by ω_i, $\delta_{a,b}$ the Kronecker's delta, and $w(\boldsymbol{\omega})$ the distribution function of $\boldsymbol{\omega}$. Note that, in the case of a single random variable ω_i, the most common $w(\omega_i)$ are related to well-known polynomials like the Hermite (Gaussian distribution), the Jacobi (beta distribution), and the Legendre (uniform distribution) polynomials. However, in the multivariate case, $\Theta_{a,b}(\boldsymbol{\omega})$ are generally unknown, unless a statistical independence between the random variables is guaranteed [8]. Only then $\Theta_{a,b}(\boldsymbol{\omega})$ may be expressed as the product of the prior polynomials related to their known distributions, i.e.,

$$\Theta_a(\omega) = \prod_{i=1}^{D} \Theta_{a_i}(\omega_i), \qquad (2)$$

where $\Theta_{ai}(\omega_i)$ is a univariate basis in ω_i of polynomial order a_i. So, any electric/magnetic component $F = \{E_x, E_y, E_z, H_x, H_y, H_z\}$ in the FDTD domain is written as:

$$F\big|_{i,j,k}^{n}(\omega) = \sum_{a=0}^{P} f^a\big|_{i,j,k}^{n} \Theta_a(\omega), \qquad (3)$$

with f^a the corresponding coefficients and P the number of polynomials, calculated for the highest order N, by:

$$P = \binom{N+D}{N} = \frac{(D+N)!}{D!N!}. \qquad (4)$$

Note that the GPC technique can not directly treat cases, where stochastic processes affect the output, since the latter are described by an infinite series of correlated random variables indexed by some physical coordinate. In contrast, the FDTD discretization can hardly offer any benefit, since one ends up with a very large (although not infinite) number of correlated random variables.

B. Karhunen-Loève stochastic representation

To overcome these issues, we launch the Karhunen-Loève (KL) expansion [8] for both the dimension reduction and decorrelation of the stochastic processes. Let $Y_x(\omega)$ be a stochastic process varying over coordinate $x \in [p, q]$ bounded domain, with a covariance function of $C(x_1, x_2)$. The KL expansion of $Y_x(\omega)$ reads:

$$Y_x(\omega) = m\{Y_x(\omega)\} + \sum_{i=0}^{\infty} \sqrt{\lambda_i} \psi_i(x) Y_i(\omega), \qquad (5)$$

where $m\{Y_x(\omega)\}$ is the mean value of $Y_x(\omega)$ and $Y_i(\omega)$ are centered, uncorrelated random variables of unit variance. Eigenfunctions $\psi_i(x)$ and their respective eigenvalues λ_i are determined via the eigenvalue problem:

$$\int_{[p,q]} C(x,l)\psi_i(l)\,dl = \lambda_i\,\psi_i(x). \qquad (6)$$

Note that for Gaussian processes, $Y_i(\omega)$ are generated as independent random variables. Also, when the process is stationary, i.e., its covariance can be written as $C(x_1, x_2) = C(x_2 - x_1)$, the solution of (6) is equivalent to the Fourier transform of $C(x)$, assuming it is periodic outside $[p, q]$. A key trait of the KL expansion (and a motive for its choice in our method) is the decay of λ_i as i increases. So, it is possible to describe the entire stochastic process with only a small truncated series of $Y_i(\omega)$.

C. Enhanced GPC-FDTD update equations

The update equations of the 3-D algorithm are extracted by plugging (3) into the leapfrog formulas [9]. The stochastic process (randomness source) $Y_x(\omega)$, in the KL expansion (5), is the relative electric permittivity $\varepsilon_r(\omega)$, where x can be any coordinate. By replacing $Y_i(\omega)$ with ω_i in (5), defining Ω, and truncating the infinite sum up to a K (KL truncation limit), we get:

$$M = m\{\varepsilon_r(\omega)\}\varepsilon_0 + \sum_{i=0}^{K} \sqrt{\lambda_i}\,\psi_i(x)\omega_i, \qquad (7)$$

which is a very accurate approximation. For instance, the E_z coefficients are given by:

$$\sum_{a=0}^{P} e_z^a\big|_{i,j,k+1/2}^{n+1} \Theta_a(\omega) = R_1 \sum_{a=0}^{P} e_z^a\big|_{i,j,k+1/2}^{n} \Theta_a(\omega)$$

$$+R_2 \left[\begin{array}{l} \dfrac{1}{\Delta x} \sum_{a=0}^{P} \left(h_y^a\big|_{i+1/2,j,k+1/2}^{n+1/2} - h_y^a\big|_{i-1/2,j,k+1/2}^{n+1/2} \right)\Theta_a(\omega) \\[2mm] -\dfrac{1}{\Delta y} \sum_{a=0}^{P} \left(h_x^a\big|_{i,j+1/2,k+1/2}^{n+1/2} - h_x^a\big|_{i,j-1/2,k+1/2}^{n+1/2} \right)\Theta_a(\omega) \end{array} \right], \qquad (8)$$

with $R_1 = (2M - \sigma'\Delta t)/(2M + \sigma'\Delta t)$, $R_2 = 2\Delta t/(2M + \sigma'\Delta t)$, and σ' the losses. Due to (3), only the respective e^a coefficients are involved. To derive the update equation for every a, we use a Galerkin process, which takes the inner product, as in (1), on both sides of (8) with the respective basis function. Thus, one arrives at:

$$e_z^a\big|_{i,j,k+1/2}^{n+1} = \frac{1}{\theta}\sum_{b=0}^{P} e_z^b\big|_{i,j,k+1/2}^{n} \left\langle R_1\Theta_a(\omega), \Theta_b(\omega)\right\rangle$$

$$+\frac{1}{\theta}\sum_{b=0}^{P}\left[\begin{array}{l} \dfrac{1}{\Delta x}\left(h_y^b\big|_{i+1/2,j,k+1/2}^{n+1/2} - h_y^b\big|_{i-1/2,j,k+1/2}^{n+1/2} \right) \\[2mm] -\dfrac{1}{\Delta y}\left(h_x^b\big|_{i,j+1/2,k+1/2}^{n+1/2} - h_x^b\big|_{i,j-1/2,k+1/2}^{n+1/2} \right) \end{array} \right] \qquad (9)$$

$$\cdot\left\langle R_2\Theta_a(\omega), \Theta_b(\omega)\right\rangle,$$

for $\theta = \langle \Theta_a(\omega), \Theta_a(\omega)\rangle$, while similar expressions hold for the other components. Note that the stability of the algorithm is specified by the usual Courant condition, which yields flexible time-steps and is proven remarkably efficient (also stated in [7]), as no late-time instabilities were observed in any of our simulations. Consequently and along with the scheme's enhanced dispersion behavior, cases with challenging variations are precisely handled.

All inner products in (9) hold only in the region with the random medium and reduce to $\delta_{a,b}$ elsewhere. Their evaluation is conducted (prior the FDTD update) for each coordinate in the region and is trivial for up to second order of the KL limit K. For higher-order approximations, (1) may be split into 1-D integrals, on condition that (2) holds. Therefore, via a Taylor series expansion of ω_i in every inner product of (9), for an arbitrarily large order s_i (even up to 10 is viewed trivial), we compute:

$$I_{s_i}^{a_i,b_i} = \int_{\Omega_i} \omega_i^{s_i}\Theta_{a_i}(\omega_i)\Theta_{b_i}(\omega_i)w(\omega_i)\,d\omega_i, \qquad (10)$$

only once, and apply the Taylor expansion. For example, the second inner product in (9) is expressed as:

$$\left\langle R_2\Theta_a(\omega), \Theta_b(\omega)\right\rangle = \sum_{s_1}^{D_T}\sum_{s_2}^{D_T}\cdots\sum_{s_K}^{D_T} T_{s_1, s_2,\ldots, s_K}\prod_{d=1}^{D} I_{s_d}^{a_d,b_d}, \qquad (11)$$

with T_s the respective Taylor coefficients and D_T the maximum Taylor series order. As a consequence, extra accuracy can be consistently accomplished.

III. GPU/CUDA IMPLEMENTATION

For its acceleration, the 3-D algorithm is parallelized for GPUs via the CUDA platform [11, 12], where different code parts are optimized for enhanced thread concurrency. As the coalescing mechanism is critical, all GPC coefficient matrices are accessed throughout the kernel code. Also, read-only matrices in each kernel are accessed through the texture buffer to increase performance. Then, the shared memory is used to load the pre-calculated inner products for all threads related to spatial coordinates that exhibit material stochasticity. For the 3-D convolution perfectly matched layers (CPMLs) [9], diverse kernels are initialized for each side and field component, executed concurrently by different streams. Hence, the proper grid and block alignment for every kernel is separately fulfilled. To evade errors at mesh corners, we use advanced atomic operators [6]. The acceleration of GPU codes compared to their conventional CPU (serialized) realizations, exceeds the promising value of 50 times.

IV. NUMERICAL RESULTS

The new technique is validated via 3-D setups terminated by 8-cell CPMLs. We examine a z-directed wire current source that illuminates an infinite planar dielectric surface at the $y = 0$ plane, featuring a non-uniform (toward x direction) stochastic permittivity (Fig. 1 (a)). The distribution function follows the Gaussian norm, hence Hermite polynomials are used as the basis functions. Also, the statistical moments of a domain point occupied by the random medium are $m\{\varepsilon_r\} = 3$ for the mean value and $\sigma\{\varepsilon_r\} = 0.02 \, m\{\varepsilon_r\}$ for the standard deviation. Our correlation function is defined as $C(x - t) = e^{-(|x - t|/\alpha)}$, with α the correlation length, while higher a translate to larger variations between neighboring points. The problem is divided into $317 \times 317 \times 317$ cells, whose size is fixed and equal to the one tenth of the wavelength corresponding to the central frequency of the Gaussian excitation pulse. Also, the time increment is set at the level of 100-150 psec via the Courant condition, while the MC-FDTD scheme gives the reference solution. It generates 5000 different stochastic processes for the dielectric scattering surface. The determined confidence intervals, considered from a 10000 realization reference, do not exceed the 10% and 1% mark for 1000 and 5000 simulations respectively. Any choice beyond this level results in a marginal improvement. Figure 1 (a) depicts three stochastic processes for $\alpha = 0.5$ and Fig. 1 (b) shows the first five eigenfunctions $\psi_i(x)$ produced by the eigenvalue problem (6). The weighted addition of the infinite series of polynomials gives exactly the spatial variation of the dielectric slab. Thus, a higher K order, including up to the Kth eigenfunction, accounts for larger spatial variations. As these play a decreasingly serious role due to the nature of the electromagnetic wave solution the truncation is safely justified.
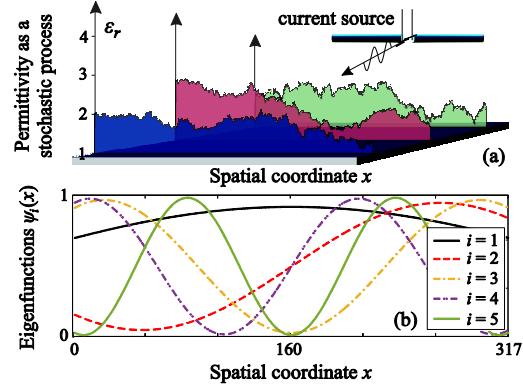
Fig. 1. (a) Perspective 3-D view of the simulation setup with the randomly generated stochastic processes, and (b) eigenfunctions $\psi_i(x)$ for $C(x - t) = e^{-(|x - t|/\alpha)}$ and $\alpha = 0.5$.

We, next, solve the problem for a KL limit up to $K = 3$ ($\alpha = 0.5$, $N = 3$). The choice of N is justified as the best compromise between accuracy and computational efficiency. Also, for larger K the system burden does not seem to justify the poor increase in accuracy. Figure 2 gives the electric field variance, where the plotted curves are snapshotted at a time-step near the peak of the Gaussian pulse, along the white line of the inlet figure. A clear improvement is attained as K augments, yet the difference between the $K = 2$ and 3 curve is very small.
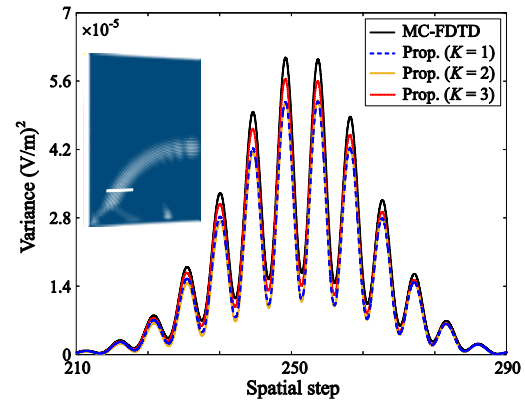
Fig. 2. Electric field variance via 5000 MC-FDTD runs and the proposed method for $\alpha = 0.5$ and $N = 3$.

Finally, Fig. 3 presents the variance of the electric field for $\alpha = 0.5$, yet with a prefixed $K = 4$ and a variety of GPC orders N. Apparently, convergence is slower in this case, revealing that the KL truncation limit has a more substantial impact compared to that of N. In contrast, it seems that the opposite situation holds for higher correlation length values of the random medium dielectric permittivity, where it is better to optimize N in an effort to accomplish the desired approximation.

The main benefit of our algorithm is that via the GPC-KL formulation, one gets a similar approximation

order as with the MC-FDTD approach. While the latter requires thousands of FDTD solutions, the complexity of the former (analogous to *P*) is lower by two magnitude orders. Thus, we can reach a sufficient convergence by analyzing simulations with ascending *K* and *L* orders and keep the efficiency over an exact MC study. It is true, however, that memory can be up to two or even five to ten times higher, in relation with *K* and *L*. So, when memory is not an issue, the GPC-KL method is a powerful alternative. The GPU/CUDA implementation, while it does not add to the theoretical analysis, it is essential for the results assessment in rational times and so it is fully preferred over CPU for both methods.
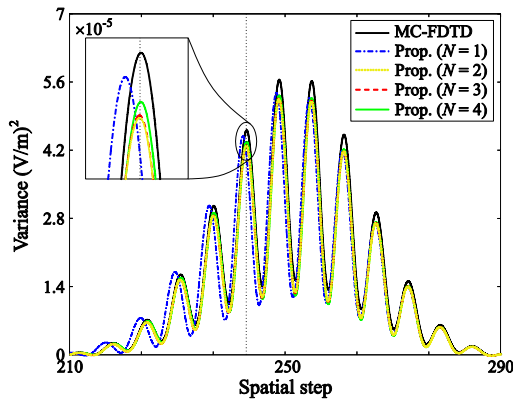
Fig. 3. Electric field variance via 5000 MC-FDTD runs and the proposed method for $\alpha = 0.5$ and $K = 4$.

V. CONCLUSION

The rigorous modeling of inhomogeneous materials with complex statistical index gradients, is presented in this paper via a consistent 3-D GPC-FDTD method. The 3-D technique utilizes the efficient Karhunen-Loève expansion to handle the arbitrary stochastic processes. Results prove the profits of the proposed concept.

REFERENCES

[1] F. Hastings, J. Schneider, and S. Broschat, "A Monte Carlo FDTD technique for rough surface scattering," *IEEE Trans. Antennas Propag.*, vol. 43, no. 11, pp. 1183-1191, Nov. 1995.

[2] M. Wong, J. Carette, A. Hadjem, and J. Wiart, "Stochastic electromagnetic modeling with uncertain dielectric properties using FDTD," in *Proc. 24th ACES Conf.*, pp. 450-455, 2008.

[3] M. Gilbert and F. Teixeira, "A small-perturbation automatic-differentiation method for uncertainty in electromagnetics," *IEEE Trans. Antennas Propag.*, vol. 60, no. 11, pp. 5305-5314, Nov. 2012.

[4] T. Tan, A. Taflove, and V. Backman, "Single realization stochastic FDTD for weak scattering waves in biological random media," *IEEE Trans. Antennas Propag.*, vol. 61, no. 2, pp. 818-828, 2013.

[5] B. T. Nguyen, C. Furse, and J. J. Simpson, "A 3-D stochastic FDTD model electromagnetic propagation in ionosphere plasma," *IEEE Trans. Antennas Propag.*, vol. 63, no. 1, pp. 304-313, 2015.

[6] G. Pyrialakos, T. Zygiridis, N. Kantartzis, and T. Tsiboukis, "GPU-based calculation of lightning-generated fields in 3D problems with statistically defined uncertainties," *IEEE Trans. Electromagn. Compat.*, vol. 57, no. 6, pp. 1556-1567, Dec. 2015.

[7] A. Austin and C. Sarris, "Efficient analysis of geometrical uncertainty in the FDTD ethod using PC for microwave circuits," *IEEE Trans. Microw. Theory Tech.*, vol. 61, no. 12, pp. 4293-4301, 2012.

[8] D. Xiu and G. Karniadakis, "The Wiener-Askey polynomial chaos for stochastic equations," *SIAM J. Sci. Comput.*, vol. 24, no. 2, pp. 619-644, 2002.

[9] A. Taflove and S. C. Hagness, *Computational Electrodynamics: The Finite-Difference Time-Domain Method*. Artech House, Boston, MA, 2005.

[10] H. W. Ott, *Electromagnetic Compatibility Engineering*. Wiley, New York, NY, 2009.

[11] V. Demir and A. Z. Elsherbeni, "Compute unified device architecture (CUDA) based finite-difference time-domain (FDTD) implementation," *ACES J.*, vol. 25, no. 4, pp. 303-314, Apr. 2010.

[12] A. Capozzoli, C. Curcio, G. D'Elia, A. Liseno, and P. Vinetti, "Fast CPU/GPU pattern evaluation of irregular arrays," *ACES J.*, vol. 2, pp. 355-272, 2010.

The Success of GPU Computing in Applied Electromagnetics

A. Capozzoli [1], O. Kilic [2], C. Curcio [1], and A. Liseno [1]

[1] Università di Napoli Federico II
Dipartimento di Ingegneria Elettrica e delle Tecnologie dell'Informazione
via Claudio 21, I 80125 Napoli, Italy
a.capozzoli@unina.it

[2] The Catholic University of America
Department of Electrical Engineering and Computer Science, Washington, DC
kilic@cua.edu

Abstract ─ In the field of electromagnetic modeling, whether it is the complex designs for engineered materials or devices and components integrated within their natural environments, there is a big drive for highly efficient numerical techniques to model the performance of complex structures. This often cannot be achieved by conventional computer systems, but rather through using the so-called high performance computing (HPC) systems that utilize hardware acceleration. We review recent General Purpose Graphics Processing Units (GPGPU) computing strategies introduced in four fields of computational electromagnetics: Finite-Difference Time-Domain (FDTD), Finite Elements Method (FEM), Method of Moments (MoM) and ElectroMagnetic Ray Tracing (EMRT).

Index Terms ─ CUDA, ElectroMagnetic Ray Tracing (EMRT), Finite-Difference Time-Domain (FDTD), Finite Elements Method (FEM), Graphics Processing Units (GPUs), Method of Moments (MoM), OpenCL, parallel programming.

I. INTRODUCTION

Electromagnetic simulators are essential tools in the analysis and the design of large and complex systems. The last two decades have witnessed dramatic improvements in both algorithms for computational electromagnetics and computing hardware. For the latter point, the use of General Purpose computing on Graphics Processing Units (GPGPU) has become increasingly prevalent. Due to their many computational cores, GPGPUs are indeed suitable for solving problems with a high degree of parallelism.

Successful applications of GPGPU computation require appropriate code implementations and optimizations, depending on whether the problem is memory bound (most of the time spent in memory transactions) or compute bound (most of the time spent

in using the GPU) [1]. Throughout the literature, there are several success stories in GPGPU computing as applied to computational electromagnetics. The purpose of this review paper is to sketch the latest GPU computing strategies adopted in four fields of particular interest; namely Finite-Difference Time Domain (FDTD), Finite Elements Method (FEM), Method of Moments (MoM) and ElectroMagnetic Ray Tracing (EMRT). For each of the mentioned representative fields, we will point out the critical aspects, which enable achieving high performance in computations. Also, we will provide relevant references, which will help the interested reader for further details. Finally, nowadays, desktop computers can easily fit four GPUs although, if more computational resources are required, multiple GPUs can be clustered together or heterogeneous systems can be used for large scale simulations. How multi-GPU and heterogeneous systems help increasing the computational performance for the mentioned applications will also be discussed.

II. FDTD

FDTD is one of the most widely used numerical methods for electromagnetic simulations. From the computational point of view, it essentially amounts at stencil calculations. Therefore, the main issue of FDTD is the very low arithmetic intensity, which means that the attainable performance in terms of Floating Point Operations per Second (FLOPS) is limited by the memory bandwidth [2].

Typical strategies like optimizing the arithmetic instructions or hiding the latency of the global memory access by maximizing the multiprocessor occupancy are not effective. For this reason, essentially the optimization approaches below have been applied to GPU-based FDTD implementations for different GPU architectures:
1. Exploit shared memory;
2. Achieve global memory coalesced accesses;

3. Use the texture cache;
4. Use built-in arrays;
5. Properly arrange the computation in the 3rd dimension.

Concerning point #1, the calculation of field components depends, at each time step, on the value of the same component at the previous step, and on other field components at neighboring cells. Accordingly, it was proposed in [3] to use shared memory to cache all the needed field components, including those corresponding to adjacent computational tiles. In this way, it is possible to significantly reduce data read redundancy. The use of shared memory also enables to limit uncoalesced accesses, as for point #2, see [4].

Regarding point #3, texture memory buffers data in a suited cache, optimized for two-dimensional spatial locality. This leads to performance gains when threads read locations that are spatially close, as in FDTD [4]. However, this benefit appears to be less relevant for latest architectures due to their newly available caching mechanisms.

Concerning point #4, built-in arrays have two, three or four components accessible which allow to best exploit global memory bandwidth. They are used to minimize the number of access operations by maximizing the number of bytes simultaneously transferred [4].

Finally, a very important point in 3D FDTD is the organization of the computation in the third dimension. An efficient solution has been proposed in [3] and a discussion of this topic, in particular, on different solutions proposed in the literature has been recently provided in [5]. An approach to reduce thread divergence when applying Convolutional Perfectly Matched Layer (CPML) boundary conditions has been also proposed in [6].

Compared to a typical implementation on multicore CPUs, an optimized parallelization on GPUs reaches a speedup of the order of ten times. By properly overlapping computation and communication, high parallelization efficiencies (~75%) can be achieved in these cases [7].

III. FEM

The Finite Element Method (FEM) is one of the most advanced and powerful methods for solving Maxwell's equations. Although often used in computational electromagnetics, GPU research on FEM has not been yet as popular as for other numerical methods. Solving Maxwell's equations using FEM essentially consists of three phases [8]:

(i) Local Assembly: For each element e in the domain, an $N \times N$ matrix, $\underline{M}_e$ (*local matrices*), and an N-length vector, $\underline{b}_e$ (*local vectors*), are computed, where N is the number of nodes per element. The computation of $\underline{M}_e$ and $\underline{b}_e$ usually involves the evaluation of integrals over the element using Gaussian quadrature. Since meshes are typically unstructured, gathering the data associated with each element forces highly irregular memory accesses.

(ii) Global Assembly: The matrices $\underline{M}_e$ and the vectors $\underline{b}_e$ are used to form a global matrix $\underline{M}$ and global vector $\underline{b}$ by assembling the contributions of the elements together. Typically, $\underline{M}$ is very sparse, although its sparsity depends on the connectivity of the mesh. The Compressed Sparse Row (CSR) format is often used to reduce the storage requirement of the matrix and to eliminate redundant computations.

(iii) Solution of the Relevant Linear System: The sparse system $\underline{M}\,\underline{x} = \underline{b}$ is solved for $\underline{x}$.

There are different possible ways of parallelizing the first two steps. Unfortunately, until now, there is no definite answer on which is the most promising approach. Different techniques are discussed in [8] that are fairly general and relevant to many types of computations on unstructured meshes. A range of possible implementations is presented and recommendations to potential implementers are given. In particular, three possibilities have been considered depending on what each thread is assigned to:

1. Assembly by non-zero elements (each thread is assigned to a different non-zero global matrix element);
2. Assembly by rows (each thread is assigned to a different row of the global matrix);
3. Assembly by elements (each thread is assigned to a different finite element).

Some results have been published for electromagnetic problems in [9] using OpenCL and in [10] using CUDA. A speedup of 19 has been observed for the former case against a multi-core CPU implementation, while a speedup between 87 (matrix assembly) and 51 (solution of the linear system) has been reported for the latter case.

IV. MOM

Method of Moments is another powerful tool used widely in computational electromagnetics. Radiation and scattering problems can be solved numerically using various formulations of the MoM (e.g., EFIE, CFIE, etc.), which is a well-established full-wave analysis based on meshing the geometry into coalescent triangles. The technique employs the expansion of the surface currents of the mesh into a set of basis functions, such as the well-known Rao-Wilton-Glisson (RWG), [11]. The series expansion results in a linear system as expressed as $[V] = [Z].[I]$, where V represents the source function, I is the unknown current, and Z is the impedance matrix. The size of the linear system; i.e., $N \times N$, depends on the number of non-boundary edges in the triangular mesh, N. In the conventional MoM approach, first the

impedance matrix is computed. Then it is inverted, and the unknown currents are calculated. The source vector is computed based on the geometry and the excitation fields at each triangle, [11].

The direct solution of MoM by a matrix inversion presents a big challenge as the object size increases. This is due to the computational complexity, $O(N^3)$, and storage requirements, $O(N^2)$ of MoM. While one way to address the complexity problem is the use of iterative solvers, MoM remains computationally expensive for electrically large objects. The Fast Multipole Method (FMM), which was first introduced by Rokhlin [12] as an augmentation to MoM, reduces the computational complexity for such problems to $O(N_{it}N^2)$ without a significant loss of accuracy. In FMM, the N edges in the mesh are classified into M localized groups, such that each group supports approximately N/M edges. The groups are then categorized as near and far, based on their spatial proximity, allowing the system matrix to be split into, Z_{near} and Z_{far} components, which describe the near and far interactions among the edges. A few authors have applied FMM for electromagnetic problems using a single GPU for small size problems [13], or a GPU cluster for larger problems [14], [15].

Further enhancements have evolved to handle larger problems, such as FMM-FFT, which applies FFT at the translation and multipole expansion stages of FMM, which reduces the complexity to $O(NlogN)$ for two-dimensional rough surfaces, [16] and to $O(N^{4/3} log2/3N)$ for three-dimensional objects, [17]. Recently, FMM-FFT was implemented on a multi-node GPU cluster to demonstrate significant acceleration in computation time while preserving the scalabilty of FMM, [18]. However, FMM-FFT still suffers from the limitation of the GPU memory to solve for larger problems. Another such attempt to enhance FMM for larger scale problems is by introducing a multi-level tree structure of MLFMA, which reduces the computational complexity of MoM to $O(NlogN)$.

V. RAY TRACING

Geometrical Optics (GO) is appealing for scenes with electrically large objects as it provides approximate solutions to Maxwell's equations. In such cases, GO can benefit from the use of data structures inherited by computer graphics, as the Binary Bounding Volume Hierarchies (BBVH), to properly handle the intersections between rays and scene objects.

Ray tracing for GO involves two main steps: searching for the intersections between rays and geometric primitives (for example, triangles) discretizing the object surfaces, and electromagnetic field transport. The first step can be the most time consuming, and must be properly managed. A simple brute force approach would be unfeasible due to the large number of intersection tests to be issued.

This intersection problem can be faced by introducing objects of simple geometry helping in determining if the ray intersects the generic primitive or not, as well as organizing primitives and objects into proper (usually binary) tree hierarchies to reduce the number of intersection tests. Typically, such objects are Axis Aligned Bounding Boxes (AABB). An AABB encloses a group of geometrical primitives or even other bounding volumes. The leaf nodes contain the primitives while the inner nodes enclose the bounding volume of its child nodes. With such a hierarchy, a tree-search algorithm is used to find the nearest object that is hit by a ray. Generally, two schemes are the most popular to construct the hierarchy, namely, *spatial subdivision* and *object partitioning*.

With spatial subdivision, space is recursively split. Each primitive is placed into all leaf nodes to which it overlaps and straddling primitives are copied in multiple nodes. Subdividing space with axis aligned planes leads to the so called KD-tree [19].

On the other side, a binary object partitioning scheme recursively subdivides the primitive list in two non-empty and disjoint sub-lists. For each sub-list, the minimum bounding volumes containing all the sub-list primitives is computed. The bounding volumes may partially overlap and the accelerating structure associated to object partitioning scheme is called BVH [20]. Unlike KDtree, each primitive is stored only once.

Object partitioning and spatial subdivision can work together resulting in a hybrid scheme known as Split Bounding Volume Hierarchy (SBVH) [20, 21], see also [22]. Recently, the benefits and the drawbacks of the above schemes have been analyzed with reference to their GPU implementations [22]. It has emerged that:

- The most critical drawback of KD-tree is the high number of primitive duplicates and the tree depth.
- Besides leading to high memory consumption (which is a problem by itself in GPU computing), primitive duplicates and tree depth are responsible of a larger (as compared to BVH) number of inner-node traversal steps, leaf visits and ray-primitive intersection tests.
- BVH, unlike KD-tree, poorly adapts to arbitrary scenes with very varying density. SBVH has shown to be a very satisfactory compromise.

With SBVH, it has recently shown how thousands of millions of rays per second can be traced on a Kepler K20c card [23].

VI. CONCLUSION

We have reviewed recent GPGPU computing strategies introduced in five fields of computational electromagnetics: FDTD, FEM, MoM and EMRT. The purpose has been to provide new Researchers in this field with initial guidelines on the dealt with topics. At present, research in GPU accelerated FEM for electromagnetics surprisingly appears to have been

overlooked in the literature.

REFERENCES

[1] P. Micikevicius, "Identifying performance limiters," *GTC Technology Conf.*, 2011.

[2] K.-H. Kim, K. H. Kim, and Q.-H. Park, "Performance analysis and optimization of three-dimensional FDTD on GPU using roofline model," *Computer Phys. Commun.*, vol. 182, no. 6, pp. 1201-1207, June 2011.

[3] P. Micikevicius, "3D finite difference computation on GPUs using CUDA," *Proc. of 2nd Workshop on General Purpose Processing on GPUs*, Washington, DC, USA, pp. 79-84, Mar. 8, 2009.

[4] D. De Donno, A. Esposito, L. Tarricone, and L. Catarinucci, "Introduction to GPU computing and CUDA programming: a case study," *IEEE Antennas Prop. Mag.*, vol. 52, no. 3, pp. 116-122, June 2010.

[5] M. Livesey, J. F. Stack Jr., F. Costen, T. Nanri, N. Nakashima, and S. Fujino, "Development of a CUDA implementation of the 3D FDTD method," *IEEE Antennas Prop. Mag.*, vol. 54, no. 5, pp. 186-195, Oct. 2012.

[6] J. I. Toivanen, T. P. Stefanski, N. Kuster, and N. Chavannes, "Comparison of CPML implementations for the GPU-accelerated FDTD solver," *Progr. Electromagn. Res.*, vol. 19, pp. 61-75, 2011.

[7] R. Shams and P. Sadeghi, "On optimization of finite-difference time-domain (FDTD) computation on heterogeneous and GPU clusters," *J. Parallel Distrib. Comput.*, vol. 71, no. 4, pp. 584-593, Apr. 2011.

[8] C. Cecka, A. J. Lew, and E. Darve, "Assembly of finite element methods on graphics processors," *Int. J. Numer. Meth.*, vol. 85, no. 5, pp. 640-669, Feb. 2011.

[9] A. Dziekonski, P. Sypek, A. Lamecki, and M. Mrozowski, "Finite element matrix generation on a GPU," *Progr. in Electromagn. Res.*, vol. 128, pp. 249-265, 2012.

[10] Z. Fu, T. J. Lewis, R. M. Kirby, and R. T. Whitaker, "Architecting the finite element method pipeline for the GPU," *J. Comput. Appl. Math.*, vol. 256, pp. 195-211, Feb. 2014.

[11] S. M. Rao, D. R. Wilton, and A. W. Glisson, "Electromagnetic scattering by surfaces of arbitrary shape," *IEEE Trans. Antennas Prop.*, vol. AP-30, no. 3. pp. 409-418, May 1982.

[12] R. Coifman, V. Rokhlin, and S. Wandzura, "The fast multipole method for the wave equation: A pedestrian prescription," *IEEE Antennas Prop. Mag.*, vol. 35, no. 3, pp. 7-12, June 1993.

[13] K. Xu, D. Z. Ding, Z. H. Fan, and R. S. Chen, "Multilevel fast multipole algorithm enhanced by GPU parallel technique for electromagnetic scattering problems," *Microw. Opt. Technol. Lett.*, vol. 52, pp. 502-507, 2010.

[14] Q. Nguyen, V. Dang, O. Kilic, and E. El-Araby, "Parallelizing fast multipole method for large-scale electromagnetic problems using GPU clusters," *IEEE Antennas Wireless Prop. Lett.*, vol. 12, pp. 868-871, 2013.

[15] V. Dang, Q. Nguyen, and O. Kilic, "Fast multipole method for large-scale electromagnetic scattering problems on GPU cluster and FPGA accelerated platforms," *Applied Comp. Electromag. Soc. Journal*, Special Issue, vol. 28, no. 12, pp. 1187-1198, 2013.

[16] R. L. Wagner, J. Song, and W. C. Chew, "Monte Carlo simulation of electromagnetic scattering from two-dimensional random rough surfaces," *IEEE Trans. Antennas Prop.*, vol. 45, no. 2, pp. 235-245, 1997.

[17] C. Waltz, K. Sertel, M. A Carr, B. C. Usner, and J. L. Volakis, "Massively parallel fast multipole method solutions of large electromagnetic scattering problems," *IEEE Trans. Antennas Prop.*, vol. AP-55, no. 6, pp. 1810-1816, 2007.

[18] V. Dang, Q. Nguyen, and O. Kilic, "GPU cluster implementation of FMM-FFT for large-scale electromagnetic problems," *IEEE Antennas Wireless Prop. Lett.*, vol. 13, pp. 1259-1262, 2014.

[19] Y. Tao, H. Lin, and H. Bao, "GPU-based shooting and bouncing ray method for fast RCS prediction," *IEEE Trans. Antennas Prop.*, vol. 58, no. 2, pp. 494-502, Feb. 2010.

[20] T. Aila and S. Laine, "Understanding the efficiency of ray traversal on GPUs," *Proc. of the Conf. on High Performance Graphics*, Saarbrucken, Germany, pp. 145-150, June 25-27, 2009.

[21] I. Wald and V. Havran, "On building fast KD-Trees for ray tracing, and on doing that in O(NlogN)," *Proc. of the IEEE Symposium on Interactive Ray Tracing*, Salt Lake City, UT, pp. 61-69, Sept. 18-20, 2006.

[22] A. Breglia, A. Capozzoli, C. Curcio, and A. Liseno, "GPU-based shooting and bouncing ray method for fast RCS prediction," *IEEE Antennas Prop. Mag.*, vol. 57, no. 5, pp. 159-176, Oct. 2015.

[23] A. Breglia, A. Capozzoli, C. Curcio, and A. Liseno, "Why does SBVH outperform KD-tree on parallel platforms?," *Proc. of the IEEE/ACES Int. Conf. on Wireless Inf. Tech. and Syst. and Appl. Comput. Electromagn.*, Honolulu, HI, pp. 1-2, Mar. 13-18, 2016.

Benefits and Challenges of GPU Accelerated Electromagnetic Solvers from a Commercial Point of View

Ulrich Jakobus

Altair Development S.A. (Pty) Ltd.
Stellenbosch, 7600, South Africa
jakobus@altair.com

Abstract ─ This paper discusses the benefits but also challenges of GPU accelerated electromagnetic solvers from a commercial point of view, namely using FEKO as example. Specifically, the effects of some of the complex interdependencies between different components are presented. It is shown that despite the advances made in the field of GPGPU computing, and impressive speedups for parts of a program or simplified problems, there are a number of factors to consider before these techniques can be applied to a commercial product that is expected to be robust and, most importantly, to always give trustworthy results for a wide variety of problems.

Index Terms ─ Commercial Solvers, CUDA, FDTD, FEKO, FEM, GPGPU, GPU Acceleration, MoM, RL-GO, SBR.

I. INTRODUCTION

In the field of computational electromagnetics (CEM), a wide range of numerical techniques can be used to simulate a variety electromagnetic radiation and scattering problems. One of the primary reasons that such a wide variety of methods exists, is that no single method performs best for all problem types [1]. Thus, one of the first challenges in solving an electromagnetic problem is to select the method that is best or at least reasonably suited to the problem of interest.

Even with the optimal method selected, there is still the matter of the available computational resources to consider. It may then be that the desired solution takes hours, days, or even weeks to compute. One of the ways in which an attempt has been made to increase the computational power at disposal – thereby decreasing the time required for a solution – has been to make use of graphics processing units (GPUs) to perform general purpose computational tasks, and not just the graphics-related tasks for which they were originally designed for. This practice, called general purpose GPU (GPGPU) computing, has seen a remarkable increase of late, both in terms of hardware capability, as well as the ease with which these devices can be programmed [2].

The most common way of programming such devices is using the Compute Unified Device Architecture (CUDA) by NVIDIA. This couples a genuinely programmable hardware architecture with programming tools that can be used by any developer with a knowledge of C/C++. Previously, GPGPU programming involved convincing a GPU to do what one wanted by rewriting computational routines as graphics programs. Since its inception, CUDA's hardware/software combination has evolved to such an extent that the latest generation of devices can be found in the fastest supercomputers in the world, with a much more powerful set of software features available as well.

There has been considerable development and a large number of papers were published on the GPU acceleration of CEM methods, for example the Method of Moments (MoM) [3] and [4], the Finite Element Method (FEM) [5] and [6], and the method of Shooting and Bouncing Rays (SBR) [7] and [8]. More general advances such as in GPU based dense linear algebra methods can be found, e.g., in [9]. The focus of this paper is not to add to this (we have done so earlier, e.g., in [10] or [11]), but instead to present an alternate perspective on these advances. That is to say the use of GPU technology as well as the challenges related to it are considered from the point of view of a commercial CEM software. To this end, the software package FEKO [12] is taken as an example. The motivation for this is that quite often such advances are considered from a purely academic standpoint, and this leads to a number of short-comings and challenges being overlooked.

Section II gives a short introduction on the FEKO solution kernel and the various CEM methods that are supported by it. This serves as background for a discussion on the difficulties associated with the GPU acceleration of a commercial CEM software package such as FEKO in Section III, and a short discussion of GPU accelerated solvers that exist in FEKO or are under development in Section IV. The paper is concluded in Section V, where a discussion on future paths to facilitate further GPU acceleration is included.

II. THE FEKO SOLUTION KERNEL

As already mentioned, a number of CEM methods exist which have their own strengths and weaknesses, and which can solve various problems of interest with varying degrees of success. It is thus important that a commercial CEM code such as FEKO implements a number of these methods to allow it to be competitive for a large selection of target application areas.

Figure 1 shows the various solution techniques available in FEKO for the solution of RF/microwave problems. Two factors influencing the choice of solution method – the electrical size of the problem being considered and the complexity of the materials being simulated – are indicated on the axes. The possibility of hybridizing various methods exists, and this allows for the solution of more complex problems by selecting the best solution method for different regions of the same problem with full bi-directional coupling between them.
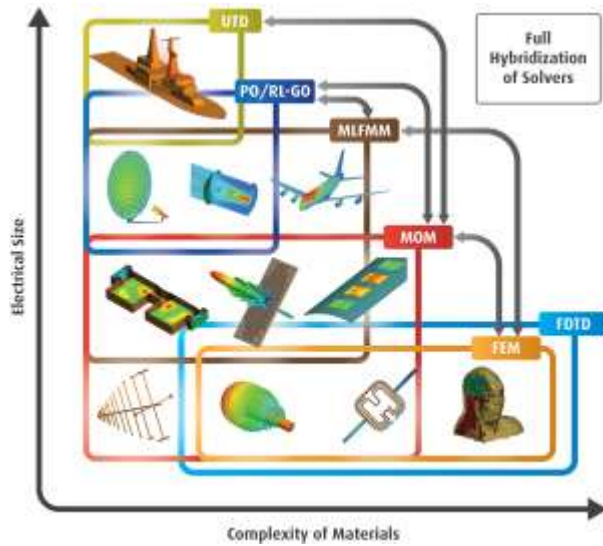
Fig. 1. A diagram depicting the various computational methods in FEKO. The hybridization that exists between some of the methods is also shown by green arrows.

III. CHALLENGES IN GPU ACCELERATION

In any software development, it is required that the available resources be allocated to maximize the delivered value in the software project. How value is determined is specific to each project, and may also differ greatly between the academic and commercial environments. In the commercial environment, for example, the number of customers with capable hardware demanding or being able to use GPU acceleration directly influences the relative value of GPU accelerated extensions when compared to other feature extensions. Academic development may, on the other hand, place a high importance on novelty for use in academic publications.

A. Versatility, reliability, and reproducibility

Many academic publications on the topic of accelerated CEM codes consider a small number of examples to illustrate the applicability or performance improvements of a specific method. These examples are often simple or canonical problems, which may play to the strengths of the method being considered, and also may not exceed the resources – such as available memory – of the GPU being used for acceleration.

In the commercial setting, there is no such control over which examples are being considered, and customers expect accurate results for a wide variety of problems. This not only imposes heavy resource requirements for additional validation and verification of the accelerated methods, but also in the detection of possible problem cases at run-time (such as running out of GPU memory and then switching automatically to block based algorithms or switching the computations on the fly back to the CPU), and handling these in a well-rounded and user friendly way.

B. Variety of CEM methods

The various computational methods included in the FEKO solution kernel and discussed in Section II have their own strengths and weaknesses when it comes to the solution of CEM problems. In addition, each of these methods present its own challenges in parallelization in general (MPI, OpenMP, etc.), and in GPU computing specifically.

Take the Methods of Moment (MoM) and the Finite Element Method (FEM) as examples. These are both matrix-based methods which require the construction, and (for driven problems) the subsequent solution of a linear system of equations. It is also possible to formulate certain classes of problems in each method as generalized eigenvalue problems.

At this point it may seem as if these two methods would be amenable to similar approaches when considering them for GPU acceleration. The situation is, however, that the linear system which results as part of a MoM computation is dense, whereas that associated with the FEM is a sparse system. Although GPU tools exit for the solution of both types of systems, the difference in performance of dense and sparse computation on a GPU means that the realized speedup will differ significantly. Furthermore, the effect of the other phases in the solution process (e.g., matrix fill) must also be taken into consideration and will be discussed in Section IV.

C. Software and design decisions

Another important factor regarding the adoption of GPU acceleration in an existing commercial CEM package are design and development decisions such as the language of implementation and low-level program

flow, which if – if not selected carefully – may not map well to massively parallel architectures such as GPUs.

CUDA was already mentioned as programming language to support NVIDIA cards. In the OpenMP 4.0 standard, for example, provision has been made for the use of accelerators. OpenMP is a directive-based, open standard which provides a portable means to parallelize code over a number of threads. The inclusion of the concept of accelerators and the associated operations, means that the importance of such technologies has been recognized. Furthermore, since the directives are platform agnostic, acceleration would in theory not be limited to a particular set of devices – such as NVIDIA GPUs when using CUDA – but the same code could be used to run on multi-core CPUs, GPUs by other vendors, and other accelerator technologies such as Intel's Xeon Phi coprocessors. There is also OpenCL, kind of being in the middle between CUDA and OpenMP. In FEKO, all three techniques (OpenCL, OpenMP, and CUDA) are being explored and partially used, but all the following GPU discussions refer to CUDA specifically.

Considering that many of the GPGPU programming tools are centered on C/C++ implementations, the options for the acceleration of for instance FORTRAN based routines generally involve rewriting large portions of code in C/C++, or switching to FORTRAN compilers that do support GPU computing. Any rewriting introduces the risk of introducing new bugs, increasing the need for proper tuning, testing and software verification.

In terms of switching compilers, there are also a number of factors to consider. One of the biggest problems is the loss of productivity – possibly for a whole development team – due to changes required in build processes and utilities, the introduction of unforeseen bugs caused by incompatible compiler options, and bugs in the compilers themselves.

IV. GPU ACCELERATION IN FEKO
A. The Method of Moments

As discussed in Section III, the MoM requires the assembly and solution of a dense linear system with other steps followed like near or far field calculations. The run-time for the assembly of the matrix is quadratic in terms of the number of unknowns, whereas that of the solution of the linear system is cubic. The post-processing is typically linear in terms of the number of unknowns and linear in terms of the number of far field directions/near field observation points etc. It follows that as the problem size gets larger, the matrix solution phase will dominate the overall run-time.

The matrix solution phase can be isolated and accelerated using libraries such as MAGMA [9] or cuSOLVER (available as part of CUDA since version 7.0). Unfortunately, even though it can be accelerated by up to an order of magnitude, the total simulation acceleration is significantly less, with the matrix assembly phase now dominating the run-time. Even though considerable speedups can be attained for this matrix fill phase in simplified MoM code [3], a considerable amount of development resources need to be invested for a FEKO implementation due to the complex nature of the code (many different basis functions, higher order on curvilinear meshes, Sommerfeld integrals for planar Green's functions etc.).

B. The Finite Element Method

Another matrix-based method implemented in FEKO is the FEM. In contrast to the MoM, the matrices are sparse, but many of the same challenges present themselves when the GPU acceleration of the method is considered.

Here, the phases of the solution process which contribute most significantly to the total simulation time are the construction of the relevant preconditioner and the subsequent solution of the sparse linear system. FEKO uses by default iterative solvers for a single right hand side which – with the right preconditioners – provide according to our experience faster solution times than direct sparse solvers and in particular use less memory.

For the solution of FEM linear system, a simple iterative solver can be expected to show a 2-5x performance improvement when running on a GPU, but for most problem sizes where the amount of GPU memory is not a limitation, this translates into a simulation speedup of only 50% as the other phases start dominating.

Further acceleration is hampered by the sheer number of preconditioning options available in a software such as FEKO. In addition, differences in matrix representation and the lack of complex value support in available third-party libraries make the use of a standalone approach – as was done with the MoM matrix solution – problematic.

C. Ray launching Geometrical Optics

Along Uniform Theory of Diffraction (UTD) and Physical Optics (PO), the Ray Launching Geometrical Optics (RL-GO) solver – which is sometimes referred to as Shooting and Bouncing Rays (SBR) – is ideal for the analysis of electrically large and complex objects. It is inherently parallel and is well suited to GPU acceleration. As an initial proof of concept, we were able to accelerate the calculation of the intersections of rays with geometry in FEKO by at least an order of magnitude when using CUDA.

However, this was handwritten CUDA code. It is not possible to simply run the RL-GO code through a GPU aware compiler and obtain an accelerated implementation with similar performance. Furthermore, the complexity and recursive nature of the code means

that GPU specific limits such as smaller stack size must be addressed as well.

D. Finite Difference Time Domain Method

In much the same way that the RL-GO solver is algorithmically well suited to GPU acceleration, the Finite Difference Time Domain (FDTD) method lends itself well to such parallelization. Much of this stems from the fact that the same simple update equations are applied to each voxel in each time step with (almost) no communication required between adjacent updates. As is indicated in Fig. 1, one advantage of the acceleration of the FDTD over the RL-GO solver in FEKO is that there is as yet no hybridization of FDTD with other methods and thus, less complexity to be considered.

The FDTD solver implemented in FEKO makes used of GPU acceleration to provide roughly an order of magnitude speedup for certain problems. One disadvantage of such a speedup is that from a user's perspective, the relative performance of post-processing phases such as the calculation of far fields is significantly lower.

For both CPU and GPU based FDTD solvers, the measured performance is greatly affected by the problem setup, which includes factors such as user-requested near fields or the far fields already mentioned. If these are in the frequency domain, for example, then additional costly computations are required during every simulation time step.

V. CONCLUSION

In this paper, a discussion on the challenges associated with the GPU acceleration of the commercial CEM software package FEKO was presented. This showed that although a method may be promising theoretically, its application in commercial software generally requires the allocation of significant development resources, with at this stage not always the necessary demand from the market.

As examples, the acceleration of the MoM, FEM, and RL-GO were considered, and although certain phases of the computational process can be accelerated significantly, the total simulation speedup is limited. The further acceleration of these methods is hampered by the complexity of the numerical algorithms, e.g., through hybridization. As illustrated, for FDTD, the situation is different.

REFERENCES

[1] D. B. Davidson, *Computational Electromagnetics for RF and Microwave Engineers*, 2nd ed., Cambridge: Cambridge University Press, 2011.

[2] D. B. Kirk and W. W. Hwu, *Programming Massively Parallel Processors – A Hands-on Approach*, Burlington: Morgan Kaufmann, 2010.

[3] E. Lezar and D. B. Davidson, "GPU-accelerated methods of moments by example: monostatic scattering," *IEEE Antennas and Propagation Magazine*, vol. 52, no. 6, pp. 120-135, Dec. 2010.

[4] M. J. Inman, A. Z. Elsherbeni, and C. J. Reddy, "CUDA based GPU solvers for method of moment simulations," *Annual Review of Progress in Applied Computational Electromagnetics*, Tampere, Finland, Apr. 2010.

[5] E. Lezar and D. B. Davidson, "GPU-based Arnoldi factorization for accelerating finite element eigenanalysis," *International Conference on Electromagnetic in Advanced Applications (ICEAA)*, Torino, Italy, Sept. 2009.

[6] A. Dziekonski, A. Lamecki, and M. Mrozowski, "On fast iterative solvers with GPU acceleration for finite elements in electromagnetics," *10th International Workshop on Finite Elements for Microwave Engineering*, Mill Falls, NH, USA, Oct. 2010.

[7] K. E. Spagnoli, *An Electromagnetic Scattering Solver Utilizing Shooting and Bouncing Rays Implemented on Modern Graphics Cards*, ProQuest, 2008.

[8] Y. Tao, H. Lin, and H. Bao, "GPU-based shooting and bouncing ray method for fast RCS prediction," *IEEE Trans. Antennas Propagat.*, vol. 58, no. 2, pp. 494-502, 2010.

[9] ICL, University Tennessee, Knoxville, "MAGMA: Matrix Algebra on GPU and Multicore Architectures," 2015. [Online]. Available: http://icl.cs.utk.edu/magma/index.html

[10] E. Lezar and U. Jakobus, "GPU accelerated electromagnetic simulations with FEKO," *International Supercomputing Conference*, Hamburg, June 2012.

[11] E. Lezar, U. Jakobus, and S. Kodiyalam, "GPU related advances in the FEKO electromagnetic solution kernel," *International Conference on Electromagnetics in Advanced Applications (ICEAA)*, Torino, Italy, Sept. 2013.

[12] Altair Development S.A. (Pty) Ltd, "FEKO – Field Computations Involving Bodies of Arbitrary Shape," 2016. [Online]. Available: www.altairhyperworks.com/feko

GPU Acceleration of Nonlinear Modeling by the Discontinuous Galerkin Time-Domain Method

Huan-Ting Meng and Jian-Ming Jin

Department of Electrical and Computer Engineering
University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA
meng2@illinois.edu, j-jin1@illinois.edu

Abstract — A discontinuous Galerkin time-domain (DGTD) algorithm is formulated and implemented to model the third-order instantaneous nonlinear effect on electromagnetic fields due the field-dependent medium permittivity. The nonlinear DGTD computation is accelerated using graphics processing units (GPUs). Two nonlinear examples are presented to show the different Kerr effects observed through the third-order nonlinearity. With the acceleration using MPI + GPU under a large cluster environment, the solution times for nonlinear simulations are significantly reduced.

Index Terms — Computational electromagnetics, DGTD, GPU acceleration, Kerr effect, nonlinear electromagnetics, third-order nonlinearity.

I. INTRODUCTION

Nonlinear phenomena in electromagnetics generally involve changes in the material properties due to the presence of electromagnetic fields. The changes in the material properties in turn modify the state of the original electromagnetic fields in the medium. Since the material properties and the contained fields interact with each other constantly, it is most natural to describe and model these interactions in the time domain, where at each time instant the changes in the fields induce nonlinear modifications on both the material properties and the fields themselves.

The nonlinear Kerr effect [1] is one of the most studied and exploited optical effects. It describes the third-order interaction between the electric field and the permittivity of the material, which produces a variety of nonlinear phenomena [1], [2], such as third-harmonic generation (THG), self-phase modulation (SPM), self-focusing, and frequency mixing. Much investigation has been carried out for the simulation of the nonlinear optical effects using the finite-difference time-domain (FDTD) algorithms [3], due to their straightforward implementation.

This work is focused on the modeling of the third-order Kerr instantaneous nonlinearity using the discontinuous Galerkin time-domain (DGTD) algorithm.

The nonlinear DGTD algorithm possesses many advantages of the linear DGTD algorithms over nonlinear FDTD algorithms, including the flexibility in complex geometry modeling, reduced phase shifts, and the ease to achieve higher order accuracy and convergence. To speed up the computation, the MPI + GPU framework developed in [4] is adapted to accelerate the nonlinear DGTD algorithm.

II. FORMULATION

For a general third-order nonlinear medium, the relative permittivity can be written as:

$$\varepsilon_r = \varepsilon_r(E) = \varepsilon_{r,\text{L}} + \varepsilon_{r,\text{NL}} = \varepsilon_{r,\text{L}} + \chi^{(3)} E^2, \qquad (1)$$

where $\varepsilon_{r,\text{L}}$ and $\varepsilon_{r,\text{NL}}$ are the linear and nonlinear parts of the relative permittivity, respectively, $\chi^{(3)}$ is the third-order nonlinear polarization coefficient, and E is the magnitude of the time-varying electric field. Here we focus on the derivation of the DGTD algorithm to model a nonlinear, lossless, and non-dispersive medium to update the electric field since the updating equation for the magnetic field has no nonlinear components and thus is identical to that in a linear medium. Testing Ampere's law using the Galerkin method, substituting in the expansion of the fields, and applying the central flux, the equation after taking the time derivative on $\bar{D}$ for element e becomes:

$$[S_e]\{h\} + \iiint_{V_e} \left[\varepsilon_0 \frac{\partial \varepsilon_r}{\partial t} \bar{N}_i^e \cdot \bar{N}_j^e \right] dV \{e\}$$
$$+ \iiint_{V_e} \left[\varepsilon_0 \varepsilon_r \bar{N}_i^e \cdot \bar{N}_j^e \right] dV \frac{\partial \{e\}}{\partial t} = [F_{eh}]\{h^+ - h\}, \qquad (2)$$

where

$$S_e(i,j) = \iiint_{V_e} \frac{1}{\mu_\infty} (\nabla \times \bar{N}_i^e) \cdot (\nabla \times \bar{N}_j^e) \, dV, \qquad (3)$$

and $\{e\}$ and $\{h\}$ are the electric and magnetic field solution vectors and $\bar{N}_i^e$ and $\bar{N}_j^e$ are vector basis functions. The terms associated with the boundary conditions are omitted for simplicity. Since the time-varying permittivity is embedded in the mass matrix of

the DGTD algorithm, the volume integration pertaining to the electric field is now split into two terms by the product rule, where for the nonlinear medium, both the relative permittivity and the electric field are functions of time. Discretizing Equation (2) in the time domain using central difference gives:

$$\iiint_{V_e} \left[\varepsilon_0 \frac{\varepsilon_r^{n+1} - \varepsilon_r^n}{\Delta t} \vec{N}_i^e \cdot \vec{N}_j^e \right] dV \left(\frac{\{e\}^{n+1} + \{e\}^n}{2} \right)$$

$$+ \iiint_{V_e} \left[\varepsilon_0 \frac{\varepsilon_r^{n+1} + \varepsilon_r^n}{2} \vec{N}_i^e \cdot \vec{N}_j^e \right] dV \left(\frac{\{e\}^{n+1} - \{e\}^n}{\Delta t} \right) \quad (4)$$

$$= \{b\}^{n+1/2},$$

where ε_r^{n+1} is the field-dependent nonlinear permittivity at the future time step, ε_r^n is the converged permittivity at the current time step, and

$$\{b\}^{n+1/2} = [F_{eh}](\{h^+\}^{n+1/2} - \{h\}^{n+1/2}) - [S_e]\{h\}^{n+1/2}. \quad (5)$$

After rearranging the terms, Equation (4) can be cast into a field-marching form as:

$$[M_e]^{n+1}\{e\}^{n+1} - [M_e]^n\{e\}^n = \{b\}^{n+1/2}, \quad (6)$$

where

$$[M_e]^{n+1} = \frac{\varepsilon_0}{\Delta t} \iiint_{V_e} \left[\varepsilon_r^{n+1} \vec{N}_i^e \cdot \vec{N}_j^e \right] dV, \quad (7)$$

and

$$[M_e]^n = \frac{\varepsilon_0}{\Delta t} \iiint_{V_e} \left[\varepsilon_r^n \vec{N}_i^e \cdot \vec{N}_j^e \right] dV. \quad (8)$$

Due to the variation of the field magnitude at each time step, $\varepsilon_r(E)$ of each element changes with time, and therefore the mass matrix $[M_e]^{n+1}$ has to be reassembled at every time step. Note that, we have recovered the original expression for $[M_e]$ as in the linear DGTD algorithm, albeit with a field- and time-dependent permittivity. The dependency of $\{e\}^{n+1}$ in $[M_e]^{n+1}$ renders Equation (6) a nonlinear equation.

At each time marching step n, the fixed-point method is employed to solve Equation (6), where $\{b\}^{n+1/2}$ is computed with the initial guess $\{e\}_0^{n+1} = \{e\}^n$ and $[M_e]_0^{n+1} = [M_e]^n$. At the kth iteration step, the mass matrix $[M_e]_{k-1}^{n+1}$ is inverted to update the field solution $\{e\}_k^{n+1}$. The updated solution is in turn used to update the mass matrix $[M_e]_k^{n+1}$ using Equation (7). If the norm of the residual $\{r\}_k^n$ of Equation (6) is smaller than a predefined threshold, then the nonlinear iteration is converged, and the equation can be marched to the next time step $n+1$. Otherwise it continues with the $(k+1)$th iteration step.

III. GPU IMPLEMENTATION

Because of the necessity to solve nonlinear equations in each time step, the nonlinear DGTD computation is very time-consuming. This computation can be effectively accelerated by exploiting the power of graphics processing units (GPUs). The GPU implementation for the nonlinear DGTD algorithm is similar to the approach described in [4], employing the same coalesced memory accessing pattern and thread/block allocation. Since the electric field update processes that are not related to $\{e\}^{n+1}$ are similar to the ones found in [6], here we focus on the parallelization of the computation related to $\{e\}^{n+1}$, which includes the assembly of the nonlinear mass matrix $[M_e]^{n+1}$ and the inversion of this mass matrix.

To assemble the nonlinear mass matrix, note that each mass matrix entry is numerically integrated through quadrature, where the contribution from each weighted quadrature point is summed. Due to the presence of nonlinearity, ε_r on each quadrature point changes during each iteration step, while the other constituting terms in equation (7) remain identical. To parallelize the assembly of the mass matrix, the constituting matrices at each quadrature point are pre-calculated and stored, and then summed together at each iteration step by first multiplying with the updated ε_r. The proposed parallelization strategy and the memory access pattern are shown in

Fig. 1, with each of the total *numTets* elements parallelized over its *numTetDofs* unknowns using CUDA threads. Each threadblock is assigned with a calculated number of elements to utilize all warps [4]. At each iteration step, the mass matrices are assembled by looping through *numQuads* quadrature points and summing their contribution, which is completely parallelizable.
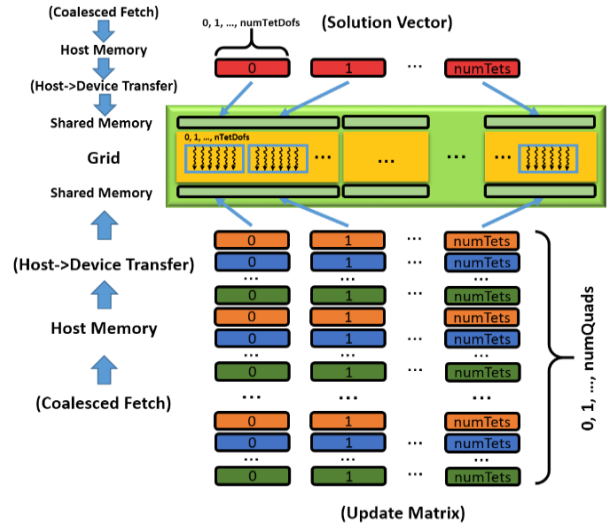
Fig. 1. Parallelization and memory access pattern for the assembly of the nonlinear mass matrices.

To invert the nonlinear mass matrix, we parallelize the standard non-pivoting element-level Gaussian elimination on the GPU. Each *numTetDofs threads* for an element loops over each elemental matrix rows and

reduce them into row echelon form. Although the elimination is only semi-parallelizable, the batch processing of the elimination process for the nonlinear elements somewhat provides a decent speedup. Note that, the mass matrix has a small condition number, and therefore can be easily inverted using the standard Gaussian elimination without partial pivoting. This is beneficial for the GPU acceleration since the partial pivoting process involves many conditional statements and branches, which are undesirable for the parallelization on GPUs.

IV. NUMERICAL EXAMPLES

Two examples are presented here to demonstrate the self-phase modulation, the third-harmonic generation, and the self-focusing effects captured by the extended DGTD algorithm and the GPU speedup. The simulation was carried out on the XSEDE Stampede cluster with NVIDIA Tesla K20 GPUs and Xeon E5-2680 CPU threads.

A. Demonstration of the self-phase modulation and the third-harmonic generation

The first example is a coaxial waveguide with an inner and outer radius of 1 and 2 mm, respectively, and a length of 40 mm. A small section of linear medium is placed near each end for excitation and absorption of the fields, and the rest of the coaxial waveguide is filled with either a linear or nonlinear medium, with a linear permittivity of $\varepsilon_{r,L} = 1.0$ and a third-order nonlinearity coefficient of $\chi^{(3)} = $ 4e-8. The input signal is a modulated Gaussian pulse with a center frequency of 20 GHz. The number of finite elements is 110,715, and the solution marches at a time step of $\Delta t = 0.075\text{ps}$ for a total of 10,000 time steps for both the linear and nonlinear cases. Mixed first-order basis functions are used for the computation. The time-domain response for the two cases is shown in

Fig. 2. It can be observed that with a linear medium, the shape of the output signal is identical to the input, whereas with a nonlinear medium the output signal steepens and forms shock waves, showing the self-steepening effect [1].

The frequency-domain response for the output signal is shown in

Fig. 3. For the linear case, we have retained the frequency profile of the original input Gaussian pulse centered at 20 GHz. For the nonlinear case, the third-harmonic effect generates harmonics at odd multiples of the original 20 GHz signal at 60 GHz, 100 GHz, 140 GHz, and so on. In addition, the self-phase modulation effect broadens the input bandwidth, where the leading and the trailing edges shift to lower and higher frequencies, respectively [1]. This result is validated using COMSOL. Table 1 gives the average per-step CPU

and GPU timing for the simulation. The lower speedup as comparing to [4] is in large due to the uneven nonlinearity encountered by the different elements, which correlates to thread idling in a warp, and the semi-serial nature of the Gaussian elimination process. This thread idleness effectively lowers the number of FLOPS as well as the overall bandwidth.
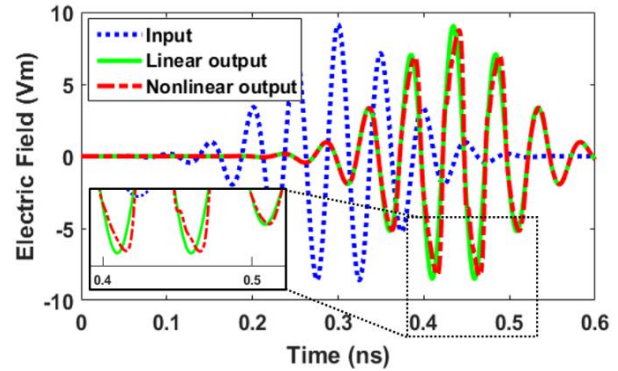
Fig. 2. Time-domain response of the electric field for a coaxial waveguide filled with a section of linear or nonlinear medium.
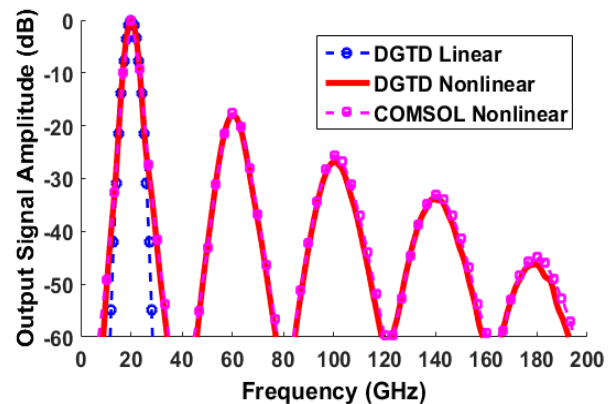
Fig. 3. Frequency-domain response of the output signal for a coaxial waveguide filled with a section of linear or nonlinear medium.

Table 1: Average per-step timing comparison for the simulation of a nonlinear coaxial waveguide

# MPI	1	2	4	8
	CPU Time per Step (ms)			
Marching	1,482.00	741.61	369.71	183.15
Comm.	0	35.51	33.74	38.79
Per-Step	1,482.00	777.12	403.45	221.94
	GPU Time per Step (ms)			
Marching	47.21	23.73	12.00	6.14
Comm.	0	2.94	1.53	4.57
Per-Step	47.21	26.67	13.52	10.71
Speedup	31.39	29.14	29.83	20.72

B. Demonstration of the self-focusing effect

The second example demonstrates the self-focusing effect through beam-shaped field propagation in a $1\text{mm}\times1\text{mm}\times3\text{mm}$ bulk medium. The linear relative permittivity is $\varepsilon_{r,L} = 1.0$ and the third-order nonlinearity coefficient is $\chi^{(3)} = 8$. The excitation is a tapered TEM sine wave at 300 GHz, launched through a square aperture with a dimension of a half of the excitation wavelength. The number of finite elements is 664,039, and the solution marches at a time step of $\Delta t = 0.01\text{ps}$ for a total of 5,000 time steps, where mixed first-order basis functions are used for the simulation. The field profiles in the bulk medium at various times for both linear and nonlinear cases are shown in Fig. 4. In the nonlinear medium, the specific electric field generates a strong nonlinearity, which results in a maximum instantaneous relative permittivity of $\varepsilon_r = 8.27$, or a 727% change to the linear relative permittivity. As can be seen, due to nonlinearity, the field experiences pulse compression which shortens the duration of each pulse. This effect is due to self-phase modulation. As the field propagates along the bulk medium, the wave is naturally diffracted in the linear medium, where the magnitude of the field decreases significantly after a couple of wavelengths. In the nonlinear medium, the intensity of the field modifies the surrounding medium into a self-induced waveguide, which counteracts natural diffraction and preserves the magnitude of the propagating wave for a longer distance in the medium.

(a)   (b)
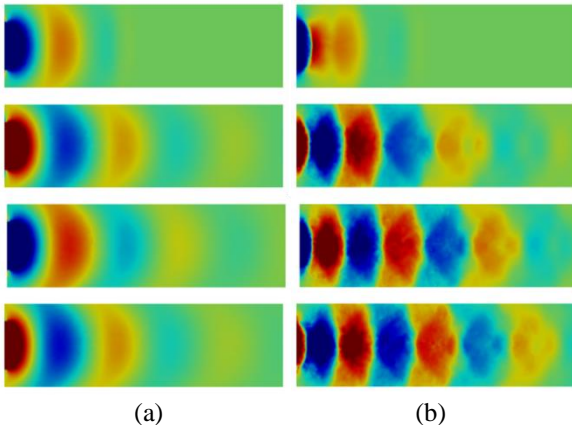
Fig. 4. Time-domain field profile for wave propagation in a: (a) linear and (b) nonlinear medium at 5, 20, 25, and 50ns, respectively.

Table 2 shows the GPU average per-step timing. Since different elements experience different levels of nonlinearity at different times due to the propagation of the field, the CUDA threads for a converged element will idle and wait for the rest of the elements in the same GPU to synchronize before completing the kernel (a single time step). This results in some MPI nodes having to idle

and wait for the others to iteratively converge before moving onto the next time step together. This idling time is taking into account in the average communication time, which is significantly longer for the fixed-point method due to the large differences in the number of iterations between different regions at any particular moment. Due to the high nonlinearity of the example, it is impractical to analyze the CPU performance. However, it is expected that higher speedup can be achieved comparing to the previous example, due to the increasing number of elements [4].

Table 2: Average GPU per-step timing (in ms) for the wave propagation in a bulk medium

# MPI	1	2	4	8
Volume	569.93	287.23	142.95	72.41
Surface	10.01	5.04	2.55	1.31
Comm.	0	35.51	33.74	38.79
Per-Step	1,482.00	777.12	403.45	221.94

V. CONCLUSION

The DGTD algorithm was extended to model the instantaneous third-order Kerr-type nonlinearity. The resulting computationally intensive DGTD algorithm was accelerated with GPUs based on the parallelization framework from our prior work. Numerical examples demonstrated that the DGTD simulation was able to capture various nonlinear phenomena and the GPU acceleration was able to achieve a good speedup for this computationally intensive simulation.

REFERENCES

[1] R. W. Boyd, *Nonlinear Optics*. Burlington, MA: Academic Press, 2008.
[2] B. Saleh and M. Tech, *Fundamentals of Photonics*. New York, NY: Wiley, 2013.
[3] R. M. Joseph and A. Taflove, "FDTD Maxwell's equations models for nonlinear electrodynamics and optics," *IEEE Trans. Antennas Propag.*, vol. 45, pp. 364-374, Mar. 1997.
[4] H.-T. Meng and J.-M. Jin, "Acceleration of the dual-field domain decomposition algorithm using MPI-CUDA on large-scale computing systems," *IEEE Trans. Antennas Propag.*, vol. 62, no. 9, pp. 4706-4715, Sept. 2014.

Multilevel Inverse-Based Factorization Preconditioner for Solving Sparse Linear Systems in Electromagnetics

Yiming Bu [1,2], Bruno Carpentieri [3], Zhaoli Shen [1,2], and Tingzhu Huang [2]

[1] Institute of Mathematics and Computer Science, University of Groningen, Groningen, 9712 CP, The Netherlands
yangyangbu@126.com, z.shen@rug.nl

[2] School of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu, 611731
China tingzhuhuang@126.com

[3] School of Science and Technology, Nottingham Trent University, Burton Street, Nottingham NG1 4BU, UK
bruno.carpentieri@ntu.ac.uk

Abstract ─ We introduce an algebraic recursive multilevel approximate inverse-based preconditioner, based on a distributed Schur complement formulation. The proposed preconditioner combines recursive combinatorial algorithms and multilevel mechanisms to maximize sparsity during the factorization.

Index Terms ─ Approximate inverse preconditioners, computational electromagnetics, Krylov subspace methods, sparse matrices.

I. INTRODUCTION

We consider multilevel approximate inverse-based factorization preconditioners for solving systems of linear equations;

$$Ax = b, \qquad (1)$$

where $A \in \mathbb{C}^{n \times n}$ is a typically large nonsymmetric sparse matrix arising from finite difference, finite element or finite volume discretization of systems of partial differential equations in electromagnetism applications. Approximate inverse methods directly approximate A^{-1} as the product of sparse matrices, so that the preconditioning operation reduces to forming one (or more) sparse matrix-vector product(s). Due to their inherent parallelism and numerical robustness, this class of methods are receiving renewed consideration for iterative solutions of large linear systems on emerging massively parallel computer systems. In practice, however, some questions need to be addressed. First of all the computed preconditioner could be singular. In the second place, these techniques usually require more CPU-time to compute the preconditioner than Incomplete LU factorization (ILU)-type methods. Third, the computation of the sparsity pattern of the approximate inverse can be problematic, as the inverse of a general sparse matrix is typically fairly dense. This leads to prohibitive computational and storage costs.

In this paper we present experiments with an algebraic recursive multilevel inverse-based factorization preconditioner that attempts to remedy these problems. The solver, proposed in [1], uses recursive combinatorial algorithms to preprocess the structure of A and to produce a suitable ordering of the unknowns of the linear system that can maximize sparsity in the approximate inverse. An efficient tree-based recursive data structure is generated to compute and apply the multi-level approximate inverse fast and efficiently. We assess the effectiveness of the sparse approximate inverse to reduce the number of iterations of Krylov methods for solving matrix problems arising from electromagnetism applications, also against other popular solvers in use today.

II. THE MULTILEVEL FRAMEWORK

We divide the solution of the linear system into the following five distinct phases:

1) a *scale phase*, where the matrix A is scaled by rows and columns so that the largest entry of the scaled matrix has magnitude smaller than one;
2) a *preorder phase*, where the structure of A is used to compute a suitable ordering that maximizes sparsity in the approximate inverse factors;
3) an *analysis phase*, where the sparsity preserving ordering is analyzed and an efficient data structure is generated for the factorization;
4) a *factorization phase*, where the nonzero entries of the preconditioner are actually computed;
5) a *solve phase*, where all the data structures are accessed for solving the linear system.

A. Scale phase

Prior to solving the system, we scale it by rows and

columns to reduce its condition number. We replace system (1) with:

$$D_1^{1/2} A y = D_1^{1/2} b, \quad y = D_2^{1/2} x, \tag{2}$$

where the $n \times n$ diagonal scaling matrices have the form:

$$D_1(i, j) = \begin{cases} \dfrac{1}{\max\limits_i |a_{ij}|}, & \text{if } i = j \\ 0, & \text{if } i \neq j \end{cases},$$

$$D_2(i, j) = \begin{cases} \dfrac{1}{\max\limits_j |a_{ij}|}, & \text{if } i = j \\ 0, & \text{if } i \neq j \end{cases}.$$

For simplicity, we still refer to the scaled system (2) as $Ax = b$.

B. Preorder phase

We describe this step using standard notation of graph theory. First, we compute the undirected graph $\Omega(\tilde{A})$ associated with the matrix;

$$\tilde{A} = \begin{cases} A, & \text{if } A \text{ is symmetric,} \\ A + A^T, & \text{if } A \text{ is unsymmetric.} \end{cases}$$

Then, $\Omega(\tilde{A})$ is partitioned into p non-overlapping subgraphs Ω_i of roughly equal size by using the multilevel graph partitioning algorithms available in the Metis package [2]. For each partition Ω_i we distinguish two disjoint sets of nodes: *interior nodes* that are connected only to nodes in the same partition, and *interface nodes* that straddle between two different partitions; the set of interior nodes of Ω_i form a so called *separable* or *independent cluster*. After renumbering the vertices of Ω one cluster after another, followed by the interface nodes as last, and permuting A according to this new ordering, a block bordered linear system is obtained, with coefficient matrix of the form:

$$\tilde{A} = P^T A P = \begin{pmatrix} B & F \\ E & C \end{pmatrix} = \begin{pmatrix} B_1 & & & F_1 \\ & \ddots & & \vdots \\ & & B_p & F_p \\ E_1 & \cdots & E_p & C \end{pmatrix}. \tag{3}$$

In (3), each diagonal block B_i corresponds to the interior nodes of Ω_i; the blocks E_i and F_i correspond to the interface nodes of Ω_i; the block C is associated to the mutual interactions between the interface nodes. In our multilevel scheme we apply the same block downward arrow structure to the diagonal blocks of $\tilde{A}$ recursively, until a maximum number of levels is achieved or until the blocks at the last level are sufficiently small and easy to factorize. As an example, in Fig. 1 (a) we show the structure of the general sparse matrix *rdb2048* from Tim Davis matrix collection [3] after three reordering levels.
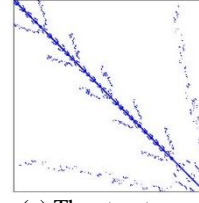
(a) The structure of rdb2048 after permutation
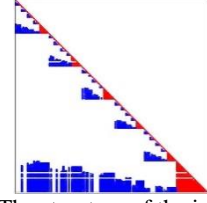
(b) The structure of the inverse factor (In red are displayed the entries actually stored)

Fig. 1. Structure of the multilevel inverse-based factorization for the matrix rdb2048.

C. Analysis phase

The data format for storing the block bordered form (3) of $\tilde{A}$ is defined, allocated and initialized using a tree structure. The root is the whole graph Ω and the leaves at each level are the independent clusters of each subgraph. In other terms, each node of the tree corresponds to one partition Ω_i or equivalently to one block B_i of $\tilde{A}$. The information stored at each node are the entries of the off-diagonal blocks E and F of $B_i's$ father, and those of the block C of B_i after its permutation, except at the last level of the tree where we store the entire block B. These blocks are stored in sparse format.

D. Factorization phase

In this phase, we compute the approximate inverse factors $\tilde{L}^{-1}$ and $\tilde{U}^{-1}$ of $\tilde{A}$, which have the following form:

$$\tilde{L}^{-1} \approx \begin{pmatrix} U_1^{-1} & & & W_1 \\ & \ddots & & \vdots \\ & & U_p^{-1} & W_p \\ & & & U_S^{-1} \end{pmatrix},$$

$$\tilde{U}^{-1} \approx \begin{pmatrix} L_1^{-1} & & & \\ & \ddots & & \\ & & L_p^{-1} & \\ G_1 & \cdots & G_p & L_S^{-1} \end{pmatrix},$$

where $B_i = L_i U_i$, and

$$W_i = -U_i^{-1} L_i^{-1} F_i U_S^{-1}, \quad G_i = -L_S^{-1} E_i U_i^{-1} L_i^{-1}, \tag{4}$$

and L_S, U_S are the triangular factors of the Schur complement matrix:

$$S = C - \sum_{i=1}^p E_i B_i^{-1} F_i.$$

During the factorization, fill-in may occur in $\tilde{L}^{-1}$ and $\tilde{U}^{-1}$ but only within the nonzero blocks. Additional sparsity is gained by applying the arrow structure (3) to the diagonal blocks recursively. This can be seen in Fig.

1 (b). For computing the factorization we only need to invert explicitly the last level blocks and the small Schur complements at each reordering level. The blocks W_i, G_i do not need to be assembled. They may be applied using Eq. (4). For the *rdb2048* problem in Fig. 1 (b), we display in red the entries that we actually stored for computing the exact multilevel inverse factorization; these are only 34% of the nonzeros of A.

E. Solve phase

In the solve phase, the multilevel factorization is applied at every iteration step of a Krylov method for solving the linear system. Notice that the inverse factorization of $\tilde{A}$ may be written as:

$$(PAP^T)^{-1} = \begin{pmatrix} U^{-1} & W \\ 0 & U_S^{-1} \end{pmatrix} \times \begin{pmatrix} L^{-1} & 0 \\ G & L_S^{-1} \end{pmatrix}, \qquad (5)$$

where $W = -U^{-1}L^{-1}FU_S^{-1}$, $G = -L_S^{-1}EU^{-1}L^{-1}$, and L_S, U_S are the inverse factors of the Schur complement matrix $S = C - EB^{-1}F$.

From Eq. (5), we obtain the following expression for the exact inverse:

$$\begin{pmatrix} B^{-1} + B^{-1}FS^{-1}EB^{-1} & -B^{-1}FS^{-1} \\ -S^{-1}EB^{-1} & S^{-1} \end{pmatrix}. \qquad (6)$$

We can derive preconditioners from Eq. (6) by computing approximate solvers $\tilde{B}^{-1}$ for B and $\tilde{S}^{-1}$ for S. Hence, the preconditioner M has the form:

$$M = \begin{pmatrix} \tilde{B}^{-1} + \tilde{B}^{-1}F\tilde{S}^{-1}E\tilde{B}^{-1} & -\tilde{B}^{-1}F\tilde{S}^{-1} \\ -\tilde{S}^{-1}E\tilde{B}^{-1} & \tilde{S}^{-1} \end{pmatrix}.$$

III. NUMERICAL EXPERIMENTS

We show some preliminary results with the proposed Algebraic Multilevel Explicit Solver (AMES) for solving a set of matrix problems arising from electromagnetics applications [3]. We summarize the list of problems in Table 1. In our experiments, we choose ILUPACK [7] as the local solver in AMES to invert the diagonal blocks at the last level, and the Schur complements at each level. Notice that in this case the entries of the inverse factors are not computed explicitly, and the application of the preconditioner is carried out through a backward and forward substitution procedure. We solve the right preconditioned system $AMy = b$, $x = My$ instead of (1), using restarted GMRES [4] preconditioned by AMES. We compare AMES against two other popular algebraic preconditioners for linear systems, that are the Algebraic Recursive Multilevel Method (ARMS) by Saad and Suchomel [5] and the Sparse Approximate Inverse preconditioner (SPAI) by Grote and Huckle [6], at roughly equal memory costs.[1] We use the zero vector as initial

guess in our code, and we terminate the solution process when the norm of residual is below 10^{-12} or the iterations count exceeds 5000. For the performance comparison, we report on the memory ratio $\frac{nnz(M)}{nnz(A)}$, number of iterations (*Its*), and time costs for performing the preordering phase (t_p), the factorization phase (t_f) and the solving phase (t_s). The experiments are run in double precision floating point arithmetic in Fortran95, on a PC equipped with an Intel(R) Core(TM) i5-3470 running at 3.20 GHz and with 8 GB of RAM and 6144 KB of cache memory.

Table 1: Set and characteristics of test matrix problems

Matrix Problem	Size	$nnz(A)$	Field
dw2048	2,048	10,114	Square dielectric waveguide
dw8192	8,192	41,746	Square dielectric waveguide
utm3060	3,060	42,211	Uedge test matrix
utm5940	5,940	83,842	Uedge test matrix
2cubes_sphere	101,492	874,378	FEM electromagnetics

A. Varying number of reduction levels in AMES

We consider the *dw2048*, *dw8192* and *2cubes_sphere* problems for these experiments. Increasing the number of levels may help reduce the number of iterations at similar memory cost. In our experiments, varying the number of levels n_{lev} from 1 to 3 for a given problem, we tuned the dropping threshold to keep roughly the same memory cost in each run, and then we studied the effect on convergence. The results of our experiments, reported in Table 2, show that using more levels enabled us to reduce the number of iterations at similar memory ratio. However, the computing time for the preordering phase (t_p) and the solution cost per iteration tend to increase with the n_{lev}. We conclude that a small number of reduction levels is recommended to use in AMES.

Table 2: Performance of AMES with varying numbers of reduction levels

Matrix	n_{lev}	$\frac{nnz(M)}{nnz(A)}$	*Its*	t_p (sec)	t_f (sec)	t_s (sec)	t_{tot} (sec)
dw2048	1	2.37	24	0.023	0.025	0.008	0.056
	2	2.33	22	0.029	0.021	0.011	0.061
	3	2.38	17	0.030	0.021	0.027	0.078
dw8192	1	3.22	87	0.067	0.109	0.312	0.488
	2	3.27	82	0.083	0.128	0.417	0.628
	3	3.28	78	0.092	0.141	0.744	0.977
2cubes_ sphere	1	0.31	12	1.271	3.691	0.310	5.272
	2	0.31	12	1.503	2.552	0.598	4.653
	3	0.31	11	2.333	1.829	1.200	5.362

---

[1] We choose a combination of parameters for AMES, and tune the dropping threshold for ARMS and SPAI to obtain similar memory cost.

B. Varying the number of reduction levels for the Schur complement

The Schur complement matrix S relative to the block C in (3) typically preserves a good deal of sparsity that can be exploited during the factorization by reordering S in a multilevel nested dissection structure, similarly to what is done to the upper leftmost block B. We have implemented this idea at the first permutation level, using ILU factorization as local solver for the reduced Schur complement matrix. We denote by AS_{lev} the number of reduction levels used for the Schur complement. We consider again the *dw2048*, *dw8192* and *2cubes_sphere* problems in these experiments. For a certain test problem, we vary AS_{lev} keeping all the other parameters constant, and we tune the drop tolerance in the ILU factorization to have similar memory costs. The value $AS_{lev} = 0$ means that only the diagonal blocks of the upper-left block B are permuted. Clearly, the max value of AS_{lev} is limited by the size of Schur complement. From Table 3, we see that simultaneous permutation of both the diagonal blocks of B and of the Schur complement S can make the AMES solver more robust to some extent. However, the implementation cost increases and thus, although useful, this option is problem dependent. In our experiments of the coming sections, we select the value for the parameter AS_{lev} that minimizes the total solution cost.

Table 3: Performance of AMES with varying numbers of reduction levels

Matrix	AS_{lev}	$\dfrac{nnz(M)}{nnz(A)}$	Its	t_p (sec)	t_f (sec)	t_s (sec)	t_{tot} (sec)
dw2048	0	2.37	24	0.023	0.025	0.008	0.056
	1	2.37	12	0.023	0.027	0.005	0.055
	2	2.37	12	0.024	0.031	0.012	0.067
dw8192	0	3.22	87	0.067	0.109	0.312	0.488
	1	3.26	21	0.067	0.164	0.057	0.288
	2	3.26	18	0.073	0.156	0.060	0.289
2cubes_ sphere	0	0.31	12	1.271	3.691	0.310	5.272
	1	0.31	11	1.277	3.974	0.334	5.585
	2	0.31	11	1.288	4.016	0.350	5.654
	3	0.31	11	1.298	3.985	0.355	5.638

C. Comparing AMES against other preconditioners

From Table 4, we can clearly see that the AMES preconditioner shows a good potential of reducing the number of iterations against other state-of-the-art preconditioning techniques at similar memory costs. This result demonstrates the overall good efficiency of the fill reducing strategies implemented in the preconditioner on the selected electromagnetic problems. One exception is the 2cubes_sphere problem, which has favourable properties for the SPAI method. The good decay of the entries away from the diagonal makes this problem suitable for SPAI. The AMES method still remains competitive. However, the pre-processing and solution costs for setting up and applying the multilevel recursive scheme do not pay off in this case.

Table 4: Performance comparison of the multilevel approximate inverse preconditioner against other iterative solvers

Matrix	Method	$\dfrac{nnz(M)}{nnz(A)}$	Its	t_p (sec)	t_f (sec)	t_s (sec)	t_{tot} (sec)
dw2048	AMES	2.37	12	0.023	0.027	0.005	0.055
	ARMS	2.39	670	0	0.009	0.081	0.090
	SPAI	2.37	2239	0	0.094	0.367	0.461
dw8192	AMES	3.26	21	0.067	0.164	0.057	0.288
	ARMS	3.37	+5000	0	0.040	+10.89	+10.93
	SPAI	3.33	+5000	0	0.836	+4.841	+5.677
Utm3060	AMES	2.79	125	0.077	0.145	0.366	0.588
	ARMS	2.93	402	0	0.030	0.763	0.793
	SPAI	2.88	+5000	0	3.131	+3.095	+6.226
Utm5940	AMES	3.50	267	0.147	0.409	2.738	3.294
	ARMS	3.51	1150	0	0.077	5.085	5.162
	SPAI	3.51	+5000	0	11.76	+11.02	+22.78
2cubes_ sphere	AMES	0.31	12	1.271	3.691	0.310	5.272
	ARMS	0.32	68	0	0.262	0.986	1.248
	SPAI	0.32	8	0	3.269	0.153	3.422

IV. CONCLUSIONS

In this paper we used recursive combinatorial techniques to remedy two typical drawbacks of explicit preconditioning, that are lack of robustness and high construction cost. The numerical experiments show that these strategies can improve the performance of conventional approximate inverse methods, yielding iterative solutions that can compete favourably against other popular solvers in use today.

REFERENCES

[1] Y. Bu, B. Carpentieri, Z. Shen, and T.-Z. Huang, "A hybrid recursive multilevel incomplete factorization preconditioner for solving general linear systems," *Applied Numerical Mathematics*, vol. 104, pp. 141-157, 2016.

[2] G. Karypis and V. Kumar, "A fast and high quality multilevel scheme for partitioning irregular graphs," *SIAM J. Sci. Comput.*, vol. 20, pp. 359-392, 1999.

[3] T. Davis, *Sparse Matrix Collection*, (1994). Available at the URL: http: //www.cise.ufl.edu/ research/sparse/matrices

[4] Y. Saad, *Iterative Methods for Sparse Linear Systems*. SIAM Publications, 2nd edition, 2003.

[5] Y. Saad and B. Suchomel, "ARMS: An algebraic recursive multilevel solver for general sparse linear systems," *Numer. Linear Algebra Appl.*, vol. 9, no. 5, pp. 359-378, 2002.

[6] M. Grote and T. Huckle, "Parallel preconditionings with sparse approximate inverses," *SIAM J. Sci. Comput.*, vol. 18, pp. 838-853, 1997.

[7] M. Bollhoefer, Y. Saad, and O. Schenk, *ILUPACK - Preconditioning Software Package*, 2010. Available online at the URL: http://ilupack.tu-bs.de.

Porting an Explicit Time-Domain Volume Integral Equation Solver onto Multiple GPUs Using MPI and OpenACC

Saber Feki [1], Ahmed Al-Jarro [3], and Hakan Bagci [2]

[1] KAUST Supercomputing Laboratory
[2] Division of Computer, Electrical and Mathematical Sciences and Engineering
King Abdullah University of Science and Technology (KAUST), Thuwal, 23955-6900, KSA
{saber.feki, hakan.bagci}@kaust.edu.sa

[3] Department of Electronic and Electrical Engineering
University College London, Torrington Place, WC1E 7JE, London, UK
ahmed.aljarro@ucl.ac.uk

Abstract ─ A scalable parallelization algorithm to port an explicit marching-on-in-time (MOT)-based time domain volume integral equation (TDVIE) solver onto multi-GPUs is described. The algorithm makes use of MPI and OpenACC for efficient implementation. The MPI processes are responsible for synchronizing and communicating the distributed compute kernels of the MOT-TDVIE solver between the GPUs, where one MPI task is assigned to one GPU. The compiler directives of the OpenACC are responsible for the data transfer and kernels' offloading from the CPU to the GPU and their execution on the GPU. The speedups achieved against the MPI/OpenMP code execution on multiple CPUs and parallel efficiencies are presented.

Index Terms ─ Explicit marching-on-in-time scheme, GPU, MPI, OpenACC, time-domain volume integral equation.

I. INTRODUCTION

The use of hardware accelerators, including multi and many-core architectures, has been increasing in many emerging applications of high performance computing (HPC) as they provide cost effectiveness, power efficiency, and physical density. Nevertheless, one of the limiting factors to a wider spread use of multi-core accelerators, such as GPUs, is the human-labor intensive porting process required by low-level programming models, such as CUDA [1] and OpenCL [2]. To overcome this limit, HPC research has focused on developing high-level directive based programming models, such as OpenACC [3], which provide compiler directives and clauses to annotate codes originally developed for CPUs in a manner similar to how OpenMP [4] is used on codes executed on multicore CPU architectures. This high-level approach, when carefully implemented, significantly reduces the re-programming efforts while maintaining the efficiency of the resulting codes.

In this work, we report on our recent efforts on parallelizing a fully explicit marching-on-in-time (MOT)-based time-domain volume integral equation (TDVIE) solver [5] for efficient execution on multiple GPUs. The MOT-TDVIE solvers are becoming attractive alternatives to finite difference time domain (FDTD) schemes for analyzing transient electromagnetic scattering from inhomogeneous dielectric objects [5, 6]. However, their effective use in practical problems of photonics, optoelectronics, and bio-electromagnetics, where electrically large scatterers need to be discretized with millions of degrees of freedom, relies on acceleration algorithms such as the plane-wave time domain (PWTD) method [7] and/or hardware-based acceleration [8-11].

Our recent research has focused on the latter; we developed highly scalable parallelization algorithms [8, 9] to enable the explicit MOT-TDVIE solver of [5] in analyzing scattering from electrically large structures. Additionally, we used OpenACC to enable the execution of the same solver on GPUs [10, 11]. Significant performance improvements with up to 30X and 11X speedups relative to the sequential and multi-threaded CPU codes were achieved. Furthermore, we demonstrated that the (single) GPU-accelerated MOT-TDVIE solver could leverage energy consumption gains on the order of 3X relative to its multi-threaded CPU version [10]. In this paper, we describe in detail the process of porting the same MOT-TDVIE solver onto *multi*-GPUs using MPI/OpenACC. Additionally, we present numerical results, which demonstrate that the ported code executes up to 11.2X faster on multi-GPUs than on conventional CPUs.

II. MOT-TDVIE SOLVER

A. MOT-TDVIE algorithm

Let V represent the volumetric support of dielectric scatterer with permittivity $\varepsilon(\mathbf{r})$ residing in an unbounded background medium with permittivity ε_0. The scatterer is excited by a band-limited incident electric field $\mathbf{E}_0(\mathbf{r},t)$. Upon excitation, scattered field $\mathbf{E}^{\text{sca}}(\mathbf{r},t)$ is generated. Scattered and incident fields satisfy $\mathbf{E}(\mathbf{r},t) = \mathbf{E}_0(\mathbf{r},t) + \mathbf{E}^{\text{sca}}(\mathbf{r},t)$, where $\mathbf{E}(\mathbf{r},t)$ is the unknown "total" field. One can construct a TDVIE as [5, 6]:

$$\mathbf{E}(\mathbf{r},t) = \mathbf{E}_0(\mathbf{r},t) + \left[\nabla\nabla\cdot - \partial_t^2/c_0 \right]\mathbf{A}(\mathbf{r},t), \ \mathbf{r}\in V, \quad (1)$$

where $\mathbf{A}(\mathbf{r},t)$ is given by:

$$\mathbf{A}(\mathbf{r},t) = \int_V \frac{(\varepsilon(\mathbf{r}')-\varepsilon_0)\mathbf{E}(\mathbf{r}',t-R/c_0)}{4\pi\varepsilon_0 R}dv', \ \mathbf{r}'\in V. \quad (2)$$

Here, c_0 is the speed of light in the background medium, and $R = |\mathbf{r}-\mathbf{r}'|$ is the distance between points $\mathbf{r}$ and $\mathbf{r}'$. TDVIE (1) is solved by time marching, which makes use of an explicit predictor-corrector algorithm as described next [5]. First V is discretized using N_e cubic elements. Let $\mathbf{r}_k$, $k = 1:N_e$, and Δt represent the centers of these elements and time step size. Assume n represents the index of the "current" time step. At the predictor step, first $\mathbf{A}_{k,n} = \mathbf{A}(\mathbf{r}_k,n\Delta t)$ are computed using $\mathbf{E}_{l,m} = \mathbf{E}(\mathbf{r}_l,m\Delta t)$, $l = 1:N_e$, $m = \max(1,n-N_g):n\Delta t$, in the integral given in (2). For this operation, $\mathbf{E}(\mathbf{r},m\Delta t) = \mathbf{E}_{l,m}$ is assumed within cubic element l and linear interpolation is used to approximate $\mathbf{E}(\mathbf{r}_l,n\Delta t - R_{kl}/c_0)$, where $R_{kl} = |\mathbf{r}_k-\mathbf{r}_l|$, from $\mathbf{E}_{l,m-1}$ and $\mathbf{E}_{l,m}$ for $[m-1]\Delta t < n\Delta t - R_{kl}/c_0 < m\Delta t$. Note that here $N_g = \lfloor R_{\max}/c_0\Delta t \rfloor + 2$, where $R_{\max} = \max\{R_{kl}\}$, for any $\mathbf{r}_k,\mathbf{r}_l \in V$. Then, finite differences (FD), which approximate the spatial derivative operator "$\nabla\nabla\cdot$", are applied to $\mathbf{A}_{k,n}$ to yield "predicted" samples $\mathbf{E}_{k,n}$. Differentiation "∂_t^2" in (1) is approximated using backward FD for pairs $(\mathbf{r}_k,\mathbf{r}_l)$ that satisfy $R_{kl} < 2c_0\Delta t$ and using central FD for all other pairs. At the corrector step, differentiation "∂_t^2" is *recomputed* using a central difference formula for pairs $(\mathbf{r}_k,\mathbf{r}_l)$ that only satisfy $R_{kl} < 2c_0\Delta t$. Note that use of central FD is now allowed since field samples that are not known at the predictor step (due to causality) can now be replaced by the predicted fields' samples. At the end of time step n, $\mathbf{E}_{k,m}$ are stored as part of the "history" of field samples to be used in the computation of $\mathbf{A}_{k,n+1}$.

Note that FD evaluations and corrector updates are spatially local operations while computation of $\mathbf{A}_{k,n}$, $k = 1:N_e$, is global. Samples $\mathbf{E}_{l,m}$ that satisfy the condition $[n-m]c_0\Delta t > R_{kl}$ do not contribute to $\mathbf{A}_{k,n}$ since the fields radiated from point $\mathbf{r}_l$ at time $m\Delta t$ have not yet reached point $\mathbf{r}_k$ at time $n\Delta t$. This also means that for $n \geq N_g$, all fields radiated from all points reach to all other points. Consequently, they all contribute to all samples $\mathbf{A}_{k,n}$, $k = 1:N_e$, rendering the computational cost of the integral evaluation $O(N_e^2)$ per time step for all $n \geq N_g$. As N_e increases, the cost of computing $\mathbf{A}_{k,n}$

limits the solver's applicability to electrically large problems. This limitation can be overcome by using acceleration algorithms such as the PWTD method [6-7] and/or highly scalable parallelization algorithms [8-11]. In this work, we implement and fine-tune the parallelization algorithm of [8, 9], which is originally developed for CPUs, for multi-GPUs to further increase the applicability of the MOT-TDVIE solver to electrically large problems.

B. MPI parallelization

Operations required by the MOT-TDVIE solver at each time step can be grouped into two: (i) computation of $\mathbf{A}_{k,l}$, $k = 1:N_e$, which requires access to samples $\mathbf{E}_{l,n-m}$, $l = 1:N_e$, $m = 1:\min(n-1,N_g)$ and (ii) computation of samples $\mathbf{E}_{k,n}$ by applying FD to $\mathbf{A}_{k,n}$. The parallelization scheme used here, first, ensures the even distribution of the memory via application of the graph-based partitioning scheme to the distribution of the points $\mathbf{r}_k$, $k = 1:N_e$, representing the discretization of V. This results in an unstructured partitioning of the points $\mathbf{r}_k$ [9]. In this partitioning, each process stores only $\mathbf{E}_{k,n}$ and $\mathbf{E}_{l,n-m}$ that belong to the partition assigned to it. The computational load of step (i) is distributed using a one-way pipeline communication strategy, so-called the "rotating tiles" paradigm [8]. The test tiles (partitions that contain test points) are initially same as the source titles (partitions that contain source points) at the beginning of the rotation but they are rotated among the processors during the computation of $\mathbf{A}_{k,l}$. When a processor receives a test tile, it first adds the contribution from the source titles it stores to $\mathbf{A}_{k,l}$ associated with the received test tile, then it passes the (updated) tile to its "neighboring" processor. At the end of a full rotation all contributions to $\mathbf{A}_{k,l}$ are computed. It is noted here that, this strategy eliminates the need for globally executed collective routines such as MPI_Reduce [8]. The computational load of step (ii) is distributed using the same grouping of the test points provided by the graph-partitioning algorithm. This reduces the communication costs associated with spatial FD computations by ensuring that the data communication only happens between points residing on the boundary of any two partitions [9].

III. PORTING TO MULTIPLE GPUS

The state-of-the-art GPU-nodes can include up to 8 K80 GPUs, which is essentially equivalent to having 16 independent GPUs. On the other hand, OpenACC standard, as a stand-alone programming model, provides very limited support for code development on multiple devices. Therefore, one typically relies on using OpenACC/OpenMP together with the MPI standard to port codes onto a cluster of nodes equipped with multiple GPUs/multicore CPUs. In this work, OpenACC is used to accelerate the time marching loop of the MOT-TDVIE

solver. Both memory- and compute-bound operations are executed on GPUs, which benefit from the improved memory bandwidth and higher flop rate, respectively. However, because the amount of compute-bound operations is significantly higher that memory-bound operations, benefits from increased memory bandwidth might be considered negligible. The main advantages of OpenACC over CUDA are the significantly increased programming efficiency and code portability on different hardware platforms. More specifically, OpenACC offers an easy way to port codes onto accelerators using simple descriptive compiler directives. Additionally, the same OpenACC-annotated code can be compiled on different hardware platforms, including the host itself (multicore CPU architecture) as well as any other accelerator supported by the OpenACC standard. In contrast, the CUDA programming model is more tedious to implement and can be used on only NVIDIA GPUs.

The code is designed such that the number of MPI processes spawn on each node is equal to the number of GPUs per node. Each MPI process is assigned to a GPU using the runtime API function *acc_set_device_num* to set the GPU target to the MPI rank modulo the number of GPUs per node, as shown in the pseudo code in Fig. 1. The data directive *#pragma acc data* is applied to the outermost time loop in order to minimize data transfers between the host and the device. Input and output arrays are annotated with clauses *present_or_copyin* and *present_or_copyout*, respectively. However, the arrays needed for the MPI communications, which are of very limited memory size, are copied in and out at each iteration so that they are accessible to the MPI routines. Each enclosed code block in the MOT-TDVIE solver is annotated with *#pragma acc kernels* and offloaded to the assigned GPU. The code blocks implementing the computation of $\mathbf{A}_{k,l}$ consist of two nested loops yielding a quadratic computational complexity. The second loop is further annotated with *#pragma acc loop reduction* and the associated variables to further optimize the sum operation of all source contributions. The OpenACC standard offers the ability to further tune loop execution using the *gang* and *vector* clauses, which can be used to modify the number of blocks of threads and threads per block to be executed, respectively. Since there are only two nested loops in the kernels of the parallel MOT-TDVIE solver, values assigned to these two parameters by the compiler already result in good performance improvements. Having said that, tuning these parameters in the presence of three or more nested loops may significantly increase the performance. Indeed, this was demonstrated for the serial version of the code, with structured grid, when executed on single GPUs. The tuning of these two parameters improved the acceleration performance by up to 23X [10]. For some of the loops that are not parallelized by the compiler due to perceived false data dependencies, the code block is annotated with

the loop pragma accompanied with the independent clause to avoid unnecessary synchronization between the loop iterations in absence of data dependencies. Note that, the code design using multi-GPU kernels allows for MPI synchronizations and communications to take place between the compute kernels as necessary. That is the case with the rotating tiles communications implemented to compute $\mathbf{A}_{k,l}$, and the halo cells exchange communications implemented to compute $\mathbf{E}_{k,n}$ using FD.

```
// Get number of MPI processes = # of GPUs
MPI_Init(&argc, &argv);
MPI_Comm_size(MPI_COMM_WORLD, &size);
MPI_Comm_rank(MPI_COMM_WORLD, &rank);
// Assign each MPI process to a GPU
acc_set_device_num(rank%ngpus,acc_device_nvidia);

#pragma acc data present_or_copyin(input
arrays) present_or_copyout(output arrays)

for (int t=0;  t<nt; ++t) {  // time loop
  for(rot=0; rot<=size; ++rot){
    // MPI communication for rotating tiles
    MPI_Sendrecv();
    MPI_Barrier();
    #pragma acc kernels
    // Spatio-temporal convolutions
    for (int k=0; k<Ne; ++k){
      #pragma acc loop reduction
      for (int l=0; l<Ne; ++l){
          A[t][k] = A[t-tkl][l] + ...
  } }
  MPI_Barrier();
    // Loops with no data dependencies
    #pragma acc kernels
    #pragma acc loop independent
    for(i=0; i<ni; ++i){

    }
  } // end rotation

  // MPI communication for Halo Exchange
  MPI_Sendrecv();
  MPI_Barrier();
  #pragma acc kernels
  // spatial finite difference operations
  for (int k=0; k<Ne; ++k){
    B[t][k] = A[t][k] + ....
  }
} // end time loop
```

Fig. 1. Pseudo code for the implementation of the MOT-TDVIE solver using MPI and OpenACC.

IV. NUMERICAL EXPERIMENTS

The test bed used for performance evaluation consists of a system of two nodes connected using an Infiniband FDR high-speed network. Each node is a dual socket CPU system hosting four NVIDIA Kepler K20c GPUs. Each socket is an eight-core Sandy Bridge

Intel(R) Xeon(R) CPU E5-2650.

In our performance evaluation, as shown in Fig. 2, a significant speedup ranging from 7.4X to 11.2X is recorded comparing the MPI and OpenMP implementation on 16 cores SandyBridge to the MPI and OpenACC implementation on four K20c GPUs. It is also observed that as N_e increases, higher speedup is achieved. This is due to the fact that the GPUs are supplied with larger computational loads; therefore, taking better advantage of its computational capacity. It has been shown before that the MPI implementation demonstrated a great scalability on large super-computers [8-9]. Figure 3 shows the parallel efficiency of the MPI and OpenACC implementation executed on two and eight GPUs, which ranges from 82% to 94%. Another advantage of using NVIDIA GPUs is their energy efficiency as the simulation consumed 2.4X less energy on GPUs than on CPUs. For all of the above, the GPUs are identified as the preferred computing platform in our overall performance analyses of the explicit MOT-TDVIE solver.
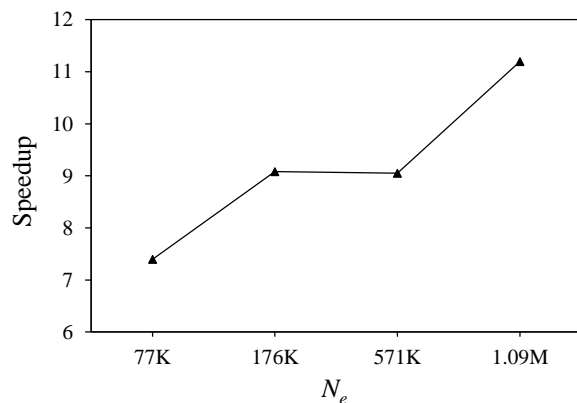

Fig. 2. Performance speedup of MPI and OpenACC on four K20c GPUs compared to MPI and OpenMP on 16 cores SandyBridge CPU.
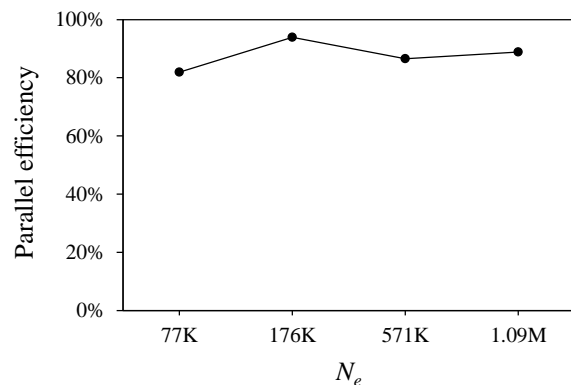

Fig. 3. Parallel efficiency of the MPI and OpenACC implementation scaling from two to eight GPUs.

V. CONCLUSION

The porting of the explicit MOT-TDVIE solver using MPI and OpenACC to multi-GPUs resulted in a highly efficient implementation. The simulations executed on multi-GPUs were faster by up to an order of magnitude compared to those executed on CPUs (using the MPI and OpenMP version of the code). The OpenACC API has the advantage of easily porting the MPI code to multi-GPU environment; therefore, increases the developer productivity while keeping the legacy of the original CPU code. Furthermore, the parallelization allows the explicit TDVIE solver to efficiently simulate transient electromagnetic wave interactions on electrically large structures discretized using a large number of spatial elements on GPUs.

REFERENCES

[1] CUDA, www.nvidia.com, 2016.
[2] OpenCL, www.khronos.org/opencl, 2016.
[3] OpenACC, www.openacc-standard.org, 2016.
[4] OpenMP, www.openmp.org, 2016.
[5] A. Al-Jarro, M. A. Salem, H. Bagci, T. M. Benson, P. Sewell, and A. Vukovic, "Explicit solution of the time domain volume integral equation using a predictor-corrector scheme," *IEEE Trans. Antennas Propag.*, vol. 60, no. 11, pp. 5203-5214, 2012.
[6] N. T. Gres, A. A. Ergin, E. Michielssen, and B. Shanker, "Volume-integral-equation-based analysis of transient electromagnetic scattering from three-dimensional inhomogeneous dielectric objects," *Radio Sci.*, vol. 36, no. 3, pp. 379-386, May 2001.
[7] Y. Liu, A. Al-Jarro, H. Bagcı, and E. Michielssen, "Parallel PWTD-accelerated explicit solution of the time domain electric field volume integral equation," *IEEE Trans. Antennas Propag.*, vol. 64, no. 6, pp. 2378-2388, 2016.
[8] A. Al-Jarro, M. Cheeseman, and H. Bagci, "A distributed-memory parallelization of the explicit time-domain volume integral equation solver using a rotating tiles paradigm," in Proc. *28th Int. Review of Progress in Appl. Comp. Electromagn.*, 2012.
[9] A. Al-Jarro and H. Bagci, "An unstructured mesh partitioning scheme for efficiently parallelizing an explicit time domain volume integral equation solver," in Proc. *29th Int. Review of Progress in Appl. Comp. Electromagn.*, 2013.
[10] S. Feki, A. Al-Jarro, A. Clo, and H. Bagci, "Porting an explicit time-domain volume-integral-equation solver on GPUs with OpenACC," *IEEE Antennas Propag. Mag.*, vol. 56, pp. 265-277, 2014.
[11] S. Feki, A. Al-Jarro, and H. Bagci, "Multi-GPU-based acceleration of the explicit time domain volume integral equation solver using MPI-OpenACC," in Proc. *IEEE Int. Symp. Antennas Propag. and USNC/URSI National Radio Sci.*, Meet., 2013.

Parallel Realization of Element by Element Analysis of Eddy Current Field Based on Graphic Processing Unit

Dongyang Wu, Xiuke Yan, Renyuan Tang, Dexin Xie, and Ziyan Ren

Department of Electrical Engineering
Shenyang University of Technology, Liaoyang, Liaoning 110870, China
shineast_521@163.com, yanxke@126.com, sgdtds@sina.com, xiedx2010@163.com, and rzyhenan@163.com

Abstract — The element by element parallel finite element method (EbE-PFEM) applied to engineering eddy current problem is presented in this paper. Unlike classical finite element method (FEM), only element matrix is needed to store for EbE method. Thereby more storage memory saved. Element by element conjugated gradient (EbE-CG) method is used to solve the equations which are discretized from elements level. Considering the ill-conditioned character of system equations, highly parallel Jacobi preconditioned (JP) method is used to accelerate the convergence. Besides, the process of dealing with boundary condition based on EbE theory is introduced. To validate the method, a 2D eddy current problem in complex frequency domain is used. The numerical analysis is carried out on the graphic processing units (GPU) with a compute unified device architecture (CUDA) parallel programming model to accelerate the convergence. And the results demonstrate that the JP method and GPU platform are effective in solving eddy current field with improved convergence.

Index Terms — Eddy current filed, element by element method, graphic processing unit, parallel computing.

I. INTRODUCTION

Due to the computer resource requirements of classical FEM for solving the electromagnetic problems, the parallel finite element method (PFEM) has become increasingly popular in recent years. Element by element (EbE) method [1] is a PFEM which can execute the parallelism on the elements level. The advantage of EbE method compared to classical FEM is that it does not need assembling and storing system matrix. Its key idea is to decouple the element solution by directly solving element equations instead of whole equations. The solving process is executed in parallel, and only intermittent communication is needed. Initially, EbE method was used for heat conduction problem and then expanded to the field of mechanics. More recently however, with the development of general purpose on graphic processing unit (GPGPU), EbE method has received increasing attention as it is very suitable for parallel processing and with the GPU[2]-[4] being a multi-core device, parallel processing at element level on different cores can be achieved. Some good results have been obtained with electrostatic problem, as in [5], [6].

In author's previous work, firstly EbB-CG method is directly used to solve 2D eddy current problem parallelly on the GPU, and 3.4 times speed up rate achieved compared with that of serial calculation with CPU [7]. Furthermore, TEAM problem 7 is taken as an example to validate the EbE method and GPU are effective for 3D linear eddy current problem, and the results have a good agreement with experiment data [8]. The purpose of this paper is to broaden the JP method to 2D eddy current analysis with two different medium in solving domain, and a comparing analysis is fulfilled between EbE-CG method and EbE-JPCG method.

II. EBE METHOD AND GPU IMPLEMENTATION

A. Node connection matrix

The key function of node connection matrix (NCM) is to transit the node information between local variables and global variables.

Now, assume x is global solution vector (GSV), x^e is the local elements solution vector (LESV), $x^{(e)}$ is global elements solution vector (GESV), E is the total number of elements, Q is NCM. Then consider three type operations of NCM as below:

$$Qx = x^{(e)}, \qquad (1)$$

where $x^{(e)} = \left(x^{(1)}, x^{(2)}, \ldots, x^{(E)}\right)^T$, this operation achieves the alternation from GSV to GESV according to the node number of each element:

$$Q^T x^e = x, \qquad (2)$$

where $x^e = \left(x^1, x^2, \ldots, x^E\right)^T$, this operation achieves the summation of LESV which have the same node number. This process alternates the LESV to GSV:

$$QQ^T x^e = x^{(e)}. \qquad (3)$$

Equation (3) achieves the alternation from LESV to GESV.

NCM also can be operated with the element matrix K^e, and the relationship between system matrix K and the element matrix K^e can be given as follows:

$$K = Q^T K^e Q. \tag{4}$$

Equations (1) to (4) provide the theoretical foundation for fulfilling the parallel EbE technique.

B. EbE-CG method

For the traditional FEM, the system matrix K and right hand side (RHS) vector b must be assembled from the element matrix K^e and element RHS b^e, while for the EbE-PFEM, considering (1)-(4) the assemble process can be deduced as follows:

$$b = Q^T b^e = Kx = Q^T K^e Qx = Q^T K^e x^{(e)}. \tag{5}$$

As shown in (5), the product of assembling the element vector is equivalent with the product of assembling element matrix. So, we can solve the element equations parallelly as below:

$$K^e x^{(e)} = b^e. \tag{6}$$

As we know, CG method mainly contains two types of inner product calculations, i.e., (r,r) and (p,Ap) which can be calculated by EbE method as follows:

$$(r,r) = r^T r = (r^e)^T QQ^T r^e = \sum (r^e)^T r^{(e)}, \tag{7}$$

where $r^{(e)} = r^e \oplus \sum_{j \in adj(e)} r^j$, r is the global residual, r^e is the local element residual, Q is the NCM. $r^{(e)}$ is the sum of r^e and r^j which are relative with r^e. So this process needs the solution information of adjacent nodes. The calculation of (p,Ap) is similar with (r,r).

C. Dealing with boundary condition

It is not necessary to assemble the system matrix for EbE method, so the boundary condition (BC) has to be applied on the elements level. Now, taking an example of 2D with triangular subdivision (Fig. 1), and assume the value of first kind BC is U_0.
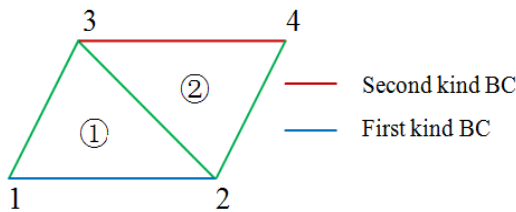

Fig. 1. Partial subdivision of 2D model.

Based on traditional FEM idea, we can get the element matrix equation of ①, as described in (8):

$$\begin{bmatrix} K_{11}^1 & K_{12}^1 & K_{13}^1 \\ K_{21}^1 & K_{22}^1 & K_{23}^1 \\ K_{31}^1 & K_{32}^1 & K_{33}^1 \end{bmatrix} \begin{bmatrix} x_1^{(1)} \\ x_2^{(1)} \\ x_3^{(1)} \end{bmatrix} = \begin{bmatrix} b_1^1 \\ b_2^1 \\ b_3^1 \end{bmatrix}. \tag{8}$$

Differ from classical FEM, the element matrix and right hand side vector must be modified with weights simultaneously. Taking element ① as an example, we can get the modified element Equation (9):

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \dfrac{K_{22}^1}{K_{22}^1 + K_{22}^2} & 0 \\ 0 & 0 & K_{33}^{(1)} \end{bmatrix} \begin{bmatrix} x_1^{(1)} \\ x_2^{(1)} \\ x_3^{(1)} \end{bmatrix} = \begin{bmatrix} U_0 \\ \dfrac{K_{22}^1}{K_{22}^1 + K_{22}^2} U_0 \\ b_3^1 \end{bmatrix}. \tag{9}$$

In contrast to first kind BC, the second kind BC (node 3 and 4) can be applied on elements directly. For the 2D eddy current problem, the current density is easily applied to the elements level during the element analysis of RHSV.

III. NUMERICAL EXPERIMENT

In this work, a conductor in an open slot of motor is taken as an example to analyze the skin effect. Two models are considered to verify the validity of the proposed method. Model I is shown in Fig. 2, it is a current-carrying conductor in an open slot, for which the analytical solution is available [9], and the domain contains only one conducting medium. And its mathematical model is shown as below:

$$\begin{cases} \dfrac{\partial^2 A}{\partial y^2} = j\omega\sigma\mu_0 A - \mu_0 J_e = p^2 A - \mu_0 J_e & in\,\Omega \\ \dfrac{\partial A}{\partial n} = 0 & (on\,AB, CD, BD) \\ \dfrac{\partial A}{\partial y} = \mu_0 H_x = -\dfrac{\mu_0 I_m}{b} & (on\,AC) \end{cases} \tag{10}$$

where A is vector magnetic potential, ω is angular frequency, σ electrical conductivity, μ_0 is magnetic conductivity, J_e is electrical current density, Ω is solving domain, H_x is tangential component of magnetic field intensity, I_m is magnitude of current and b is width of open slot. And the analytical solution of current density (J) is shown as follows:

$$J = -j\omega\sigma A + J_e = \dfrac{pI_m}{b\,\mathrm{sh}ph} \cdot \mathrm{ch}py. \tag{11}$$

Additional, in order to validate the proposed method for eddy current problem with different mediums, Model II is established in this paper (as shown in Fig. 3). For Model II, there is 1 mm width air gap surrounding the conductor, for which the condition number of its system matrix becomes greater than that of Model I, and convergence of solving the equations also becomes worse. Both of two models are under the complex excited current $I_m = (10000 + j0)$ A.

To test the accelerating performance of proposed method on different computation scales, Model I and

Model II have been meshed into three different sizes, shown in Table 1 and Table 2. The mesh of Model II in size B is shown in Fig. 4, and its magnetic field distribution is shown in Fig. 5. Furthermore, the convergence of equations solved using CG and JPCG has been researched. Both of EbE-CG and EbE-JPCG methods are implemented with CPU and GPU separately.
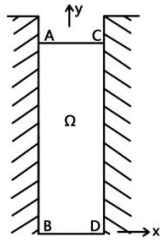

Fig. 2. Current-carrying conductor in an open slot with air surrounded.
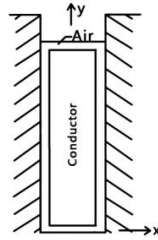
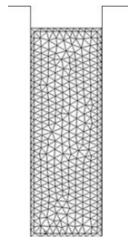Fig. 3. Current-carrying conductor in an open slot.

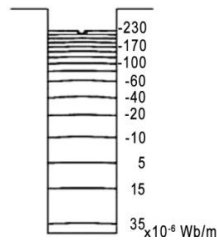Fig. 4. The mesh of Model II in size B.

Fig. 5. Magnetic field distribution of Model II.

All the numerical computations are carried out on a server with NVIDIA GTX 660 GPU clocked at 1.0 GHz with 960 cores and 2G DDR5 global memory, and an Intel Xeon E3-1230 CPU 3.3 GHz with 8G global memory. Programming is in C++, and compiled by Visual Studio 2010 and CUDA 5.5.

To reduce the communication cost between CPU and GPU, the whole elements information is transferred to GPU global memory initially. The solving process is operated parallelly on GPU until computation results meet the convergence criterion, then result data is transferred from GPU to CPU. The GPU calculation is fulfilled on different blocks, and the threads on the same block are parallel running. But different block cannot communicate. However, during the CG iteration process, some kinds of steps such as the calculation of $r^{(e)}$ need the information of other relative elements which are not in the same block. To overcome this, if the nodes on the boundary of memory block, the node information is stored on both sides concurrently. A little more memory needed, but high parallelism obtained. For other steps, all the read and write instructions for threads within same warp (a cluster of threads) are operated in the aligned and coalesced way to improve parallel performance.

The calculation results are shown in Table 1 and Table 2. Table 3 is shown the comparison of memory required. Figure 6 is the current density comparison between analytical and numerical solution of Model I. From Fig. 6, we can see that the result calculated using the proposed correlates well with analytical solution, which validates the method.
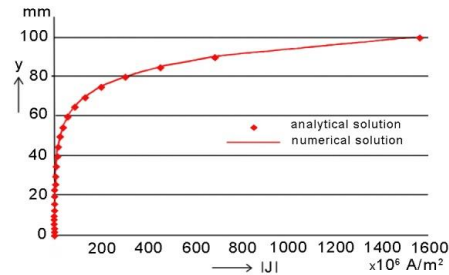
Fig. 6. Comparison of current density with EbE-CG method (Model I).

Table 1: The numerical results of Model I

Mesh Size	Node	Element	Iterations		CPU Time (ms)		GPU Time (ms)	
			CG	JPCG	CG	JPCG	CG	JPCG
A	90	138	56	33	78	62	23	18
B	342	594	100	46	485	359	87	65
C	1080	1953	175	68	1549	1231	239	173

Table 2: The numerical results of Model II

Mesh Size	Node	Element	Iterations		CPU Time (ms)		GPU Time (ms)	
			CG	JPCG	CG	JPCG	CG	JPCG
A	580	683	85	73	2578	1927	753	557
B	905	1511	134	96	6987	5125	1215	843
C	1384	2235	201	137	9768	7254	1441	935

Table 3: Comparison of memory required (Model II)

Mesh Size	Memory Required (kB)		Memory Saved (%)
	EbE	FEM	CG
A	32	79	59.5
B	72	145	50.3
C	107	218	50.9

The distribution of current density in Model II is shown in Fig. 7, which also shows that accurate results can be obtained using EbE-JPCG to eddy current problem with different medium. From the results shown in Table 1 and Table 2, we can see that the convergence of equations solving using JPCG is better than that using CG. For the same model, the GPU processor is faster than CPU due to its high parallelism.

Figure 8 shows three different mesh size level's speed up rate comparison of EbE-CG and EbE-JPCG methods which are fulfilled on GPU for Model II. Figure 9 shows the speed up rate comparison of EbE-JPCG

method fulfilled on GPU for two models.

Both EbE-CG method and EbE-JPCG method are applied to Model II which contains two materials. As shown in Table 2, the time consumed is much more than Model I, however, results indicate overall improved convergence and processing time with increasing mesh size as shown in Figs. 8-9.
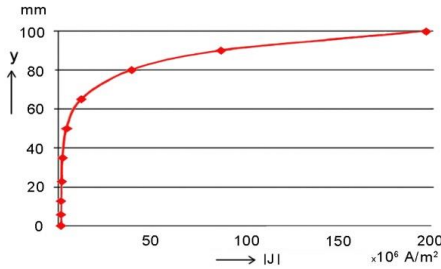

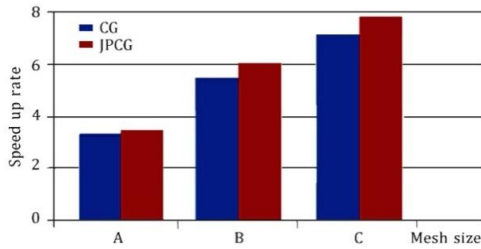Fig. 7. Distribution of current density for EbE-JPCG method (Model II).

Fig. 8. Comparison of speed up rate for EbE-CG method and EbE-JPCG method.
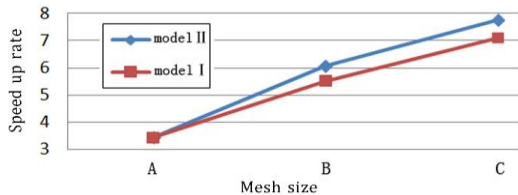

Fig. 9. Comparison of speed up rate for EBE-JPCG method implementation on GPU (Model II).

IV. CONCLUSION

The EbE-JPCG technique and GPU parallel computing platform applied to eddy current problems are the main contributions of this work. This paper presents a comparative analysis of the performance of EbE-CG method and EbE-JPCG method which are fulfilled on CPU and GPU. As shown in Table 1, Table 2 and Fig. 8, EbE-JPCG method converges more quickly than the EbE-CG method. As well, GPU acceleration becomes more effective with increasing mesh size. The numerical results demonstrate that JP method is effective for EbE method and parallel computing. As shown in Table 3, EbE method can save approximately 50% memory space, it is an important contribution for GPU platform which just has a few GB memory. Another contribution of this paper is to provide basis for solving of 3D eddy current problem, as in [8]. The future work currently in progress includes applying the EbE technique and GPU parallel platform to 3D eddy current losses calculation of large power transformer. Considering its serious ill-conditioned, JP method will be ineffective. So a new improved JP method which is also convenient for parallel EbE implementation is included in the ongoing work.

REFERENCES

[1] T. J. R. Hughus, I. Levit, and J. Winget, "An element-by-element solution algorithm for problems of structural and solid mechanics," *Computer Methods in Applied Mechanics and Engineering*, vol. 36, pp. 241-254, 1983.

[2] A. F. P. Camargos and V. C. Silva, "Performance analysis of multi-GPU implementations of Krylov-subspace methods applied to FEA of electromagnetic phenomena," *IEEE Transactions on Magnetics*, vol. 51, no. 3, March 2015.

[3] T. Okimura, T. Sasayama, and N. Takahashi, "Parallelization of finite element analysis of nonlinear magnetic fields using GPU," *IEEE Transactions on Magnetics*, vol. 49, no. 5, May 2013.

[4] O. Bottauscio, M. Chiampi, J. Hand, et al., "A GPU computational code for eddy-current problems in voxel-based anatomy," *IEEE Transactions on Magnetics*, vol. 51, no. 3, March 2015.

[5] D. M. Fernández, M. M. Dehnavi, and W. J. Gross, "Alternate parallel processing approach for FEM," *IEEE Transactions on Magnetics*, vol. 48, no. 2, pp. 399-402, 2012.

[6] I. Kiss, S. Gyimóthy, Z. Badics, et al., "Parallel realization of the element-by-element FEM technique by CUDA," *IEEE Transactions on Magnetics*, vol. 48, no. 2, pp. 507-510, 2012.

[7] D. Y. Wu, R. Y. Tang, and D. X. Xie, "*Element by element finite element method applied to eddy current problem*," *CEFC'2014*, France, May 25-28, 2014.

[8] D. Y. Wu, X. K. Yan, R. Y. Tang, D. X. Xie, and Z. W. Chen, "GPU acceleration of EBE method for 3-D linear steady eddy current field," *ICEMS'2015*, Thailand, October 25-28, 2015.

[9] D. X. Xie, "Finite element method applied to the calculation of skin effect of current-carrying conductor in an open slot," *HIET Journal*, vol. 4, no. 1, pp. 6-23, 1981. (In Chinese).

GPU-based Electromagnetic Optimization of MIMO Channels

Alfonso Breglia, Amedeo Capozzoli, Claudio Curcio, Salvatore Di Donna, and Angelo Liseno

Università di Napoli Federico II, Dipartimento di Ingegneria Elettrica e delle Tecnologie dell'Informazione
via Claudio 21, I 80125 Napoli, Italy
a.capozzoli@unina.it

Abstract — Strategies to accelerate MIMO channel capacity optimization on GPUs are outlined. The optimization scheme is dealt with by properly facing the main computational issues. In particular, the propagation environment is described by ultrafast Geometrical Optics (GO), singular values are computed by a very fast scheme and the optimizer is a parallel version of the differential evolutionary algorithm. The unknowns are given proper representations to reduce the number of optimization parameters.

Index Terms — CUDA, differential evolutionary, Geometrical Optics, GPU, MIMO channel, optimization, singular values.

I. INTRODUCTION

Multiple Input - Multiple Output (MIMO) [1, 2] is a wireless communication technology using multiple antennas in both transmission and reception modalities to increase the channel capacity (CC) over that of a conventional SISO (Single Input - Single Output). CC is a crucial parameter to assess the performance of a communication system, and the problem arises of how defining the most convenient configuration of transmitting (TX) and receiving (RX) antennas to maximize it for a given SNR (Signal to Noise Ratio) and propagation environment.

Different approaches have been proposed to optimize the MIMO CC, see [3, 4] for two representative examples. As it appears from [3, 4], besides signal processing factors (e.g., modulation), two critical aspects emerge when optimizing the performance of a MIMO channel: one is the electromagnetic environment, since the electromagnetic propagation influences the properties of the channel matrix, and the other is the optimization scheme itself. Since both aspects pose a significant computational question, the issue thus arises of how computationally addressing the problem of optimizing a MIMO channel, by firstly determining the most convenient computational resources and algorithms to be exploited and that could make the challenge feasible. Then, how much a MIMO channel can be improved, how

the CC depends on the accuracy of the employed electromagnetic model and what can be obtained by a modeling grasping only the essential aspects of the problem should be pointed out, giving general guidelines at the design stage. These points have been up to now overlooked throughout the literature. Our purpose is facing the first point, namely, how much accelerated analysis and optimization can make the goal viable. This entails understanding how to push the performance of both, the employed algorithms and computational resources. Several computational key points should be then considered, since each performance can degrade the problem to unfeasibility:

1. Properly choosing and accelerating a global optimizer since a local optimizer is typically not enough to find the best solutions;
2. Properly choosing and accelerating the approach to compute the MIMO channel matrix;
3. Being the CC related to the singular values (SVs) of the channel matrix, accelerating their computation depending on the problem size (conventional - 4x4, 6x6 - MIMO vs. massive MIMO [5]);
4. Properly representing the unknowns, to manage only the essential optimization parameters;
5. Properly exploiting massively parallel computing platforms as Graphics Processing Units (GPUs).

The paper is organized as follows. The MIMO CC is briefly recalled in Section II, just to provide a formal introduction. Section III addresses points 1) and 4), Section IV point 2) and Section V points 3) and 5). Finally, Sections VI and VII present numerical results and conclusions, respectively.

II. CHANNEL MODEL

Let us consider a narrowband, flat-fading channel, whose CC depends on the distribution of the SVs of the channel scattering matrix [1, 2]. Indeed, given N_{TX} transmitting and N_{RX} receiving antennas embedded in a complex 3D deterministic electromagnetic scenario, the MIMO channel can be described by its complex, N_{TX} x N_{RX} matrix $\underline{\underline{H}}$ [1]. The generic element h_{ij} of $\underline{\underline{H}}$ can be expressed as:

$$h_{ij} = h(\underline{r}_i^{TX}, \underline{r}_j^{RX}) = \sum_{m=1}^{M} G_m(\underline{r}_i^{TX}, \underline{r}_j^{RX}), \qquad (1)$$

where $\underline{r}_i^{TX}$ represents the position of the i-th transmitting antenna, $\underline{r}_j^{RX}$ represents the position of the j-th receiving antenna, $M(i, j)$ is the number of relevant multi-paths between $\underline{r}_i^{TX}$ and $\underline{r}_j^{RX}$ and $G_m(\underline{r}_i^{TX}, \underline{r}_j^{RX})$ is proportional to the voltage induced on the j-th receiving antenna by the m-th multipath originated at the i-th transmitting antenna.

Under the hypothesis of narrowband, flat-fading channel, AWGN noise at the receivers and equally distributed power among the transmitters [1], the Shannon CC, say C, expressed in bit/s/Hz can be calculated by first normalizing $\underline{\underline{H}}$ to its Frobenius norm as [6]:

$$\underline{\underline{\tilde{H}}} = \underline{\underline{H}} / \sqrt{\frac{\sum_{i=1}^{N_{TX}} \sum_{j=1}^{N_{RX}} |h_{ij}|^2}{N_{TX} N_{RX}}}, \qquad (2)$$

and then computing,

$$C = \sum_{k=1}^{r} \log_2 \left(1 + \frac{SNR \sigma_k^2}{N_{TX}}\right), \qquad (3)$$

where r is the rank of $\underline{\underline{\tilde{H}}}$ and σ_k^2 is its k-th SV.

The approach is illustrated in Fig. 1, where the flow-chart boxes correspond to the titles of the Sections III-V.
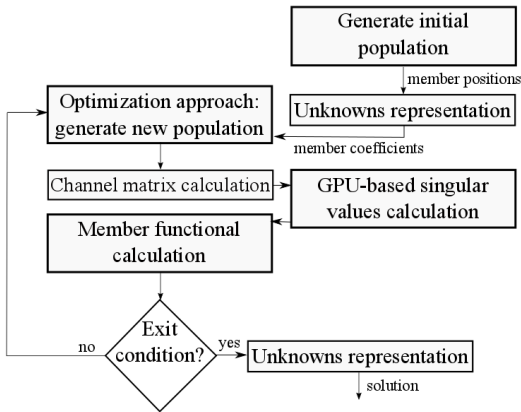

Fig. 1. Flow chart of the approach.

III. THE OPTIMIZATION APPROACH

To enable a satisfactory exploration of the objective functional landscape, the global optimization approach should be chosen to exhibit good convergence properties and to profit of the massive parallelization. In this sense, a less "complex", but massively parallelizable algorithm should be preferred to a more "involved", but "more sequential" scheme. The "genetic-like" differential evolutionary approach has been then chosen as the global optimization scheme due to its main features matching both the above mentioned requirements [7].

The approach exploits a population of N_p members, each member being represented by an array of D values, where D coincides with the number of optimization unknowns. At the g-th iteration (generation), the k-th member of the population is denoted by the D-dimensional array $\underline{p}_{k\,g} = (p_{k\,g}^{(1)}, p_{k\,g}^{(2)}, ..., p_{k\,g}^{(D)})$. The initial population is randomly generated, accounting for some physical constraints enforced by the problem. Starting from the initial population, the algorithm generates a new one by first defining new arrays as (mutation):

$$\underline{m}_{k\,g+1} = \underline{p}_{k_1\,g} + F(\underline{p}_{k_2\,g} - \underline{p}_{k_3\,g}), \qquad (4)$$

where k_1, k_2 and k_3 belong to $\{1, 2, ..., N_p\}$ and three indices mutually different and different also from k, and $F \in [0,2]$ is a user defined real and constant factor representing a scale factor of the differential variation $\underline{p}_{k_2\,g} - \underline{p}_{k_3\,g}$.

Following mutation, new trial arrays $\underline{t}_{k\,g+1} = (t_{k\,g+1}^{(1)}, t_{k\,g+1}^{(2)}, ..., t_{k\,g+1}^{(D)})$ are generated as follows (crossover):

$$t_{k\,g+1}^{(i)} = \begin{cases} m_{k\,g+1}^{(i)}, & \text{if } rand(j) \leq CR \text{ or } j = randD(k) \\ p_{k\,g}^{(i)}, & \text{if } rand(j) > CR \text{ or } j \neq randD(k) \end{cases}, \qquad (5)$$

where $rand(j) \in [0,1]$, $j=1,2,...,D$, is the j-th evaluation of a random uniform generator, $CR \in [0,1]$ is a user defined crossover constant and $randD(k) \in \{1,2,...,D\}$ is a randomly chosen integer to ensure that the trial array contains at least a mutated element. Following the crossover, the cost values of $\underline{p}_{k\,g}$ and $\underline{t}_{k\,g+1}$ are compared and the one with the largest cost becomes the new population member at generation $g+1$ (selection).

The operations involved in mutation, crossover and selection and the random number generations are inherently parallel. An issue of the crossover stage is the "random" global memory access, so that particular care has been given to improve memory coalescence.

Unknowns representation

The problem concerns the optimization of the TX and RX antenna locations to maximize the MIMO CC. To profit from a reduction of the number of unknowns, both the TX and RX antennas are assumed to be located on lines and their positions are indirectly searched for by representing them by Legendre polynomials [8] as:

$$x_n = \sum_{k=0}^{K-1} c_k \psi_k(\xi_n). \qquad (6)$$

In Eq. (6), x_n is the generic antenna coordinate on the optimization line, K is the number of polynomials, ψ_k is the k-th Legendre polynomial, the ξ_n's are uniformly spaced points in [-1,1] and the c_k's become the actual unknowns to be sought for. If K is less than the involved antennas, the number of problem parameters is reduced.

Notice that the representation in Eq. (6) is also amenable to enforcement of constraints on minimum and maximum antenna spacings [9].

IV. CHANNEL MATRIX CALCULATION

The method exploited to calculate $\tilde{\underline{\underline{H}}}$ should trade off computational accuracy and speed to execute in an iterative optimization. Calculating at each generation the matrix $\tilde{\underline{\underline{H}}}$ for a large number of antenna configurations by a full wave method would be unfeasible, especially for large scenarios. Opposite to that, Geometrical Optics (GO) is very appealing to quickly provide an approximate solution to Maxwell's equations.

Nevertheless, for electrically large scenarios GO must be properly accelerated by adequate algorithmic structures capable to properly handle the intersections of the rays with the scene objects. Indeed, ray tracing involves two main steps: the search for the intersection between a ray and the geometric primitives (e.g., triangles), and the electromagnetic field transport. The first step can be definitely the most time consuming one, if not properly managed. A brute force approach would be indeed unfeasible due to the large number of intersection tests to be performed. Fortunately, the problem can be faced by tree-like structures which, if properly setup, managed and explored, can significantly reduce the computational complexity. Data structures like KD-tree and BVH (Bounding Volume Hierarchy) [10] can be effectively applied to this purpose and may profit from a high degree of parallelization. Here, the Split BVH (SBVH) scheme set up in [10] has been exploited.

V. GPU-BASED SVs CALCULATION

Computing the SVs of small or large matrices should be dealt with different approaches. Accordingly, the computational scheme to be employed differs if considering conventional or massive MIMOs. In this paper, we address the former case. Furthermore, the number of involved matrices is related to the number of population members of the differential optimizer. Then, at each generation, the SVs of a large number of small sized matrices have to be computed. This task can be efficiently and effectively performed on a GPU as in [11].

The problem of computing the SVs can be recast to the computation of the SVs of a real-valued matrix $\underline{\underline{A}}$. To this end, the approach in [7] consists of three steps. The first step amounts at reducing $\underline{\underline{A}}$ to a bidiagonal matrix, say $\underline{\underline{B}}$, as:

$$\underline{\underline{A}} = \underline{\underline{P}}\,\underline{\underline{B}}\,\underline{\underline{Q}}^T, \tag{7}$$

where $\underline{\underline{B}}$ is a N_{TX} x N_{RX} upper bidiagonal matrix, and $\underline{\underline{P}}$ and $\underline{\underline{Q}}$ are N_{TX} x N_{TX} and N_{RX} x N_{RX} orthogonal Householder matrices, respectively. The bidiagonalization

step consist of applying a sequence of Householder transformations [12] to the matrix $\underline{\underline{A}}$, which zero the elements below the diagonal and to the right of the first superdiagonal. In the second step, $\underline{\underline{B}}$ is transformed to a tridiagonal matrix $\underline{\underline{T}} = \underline{\underline{B}}^T \underline{\underline{B}}$. Finally, in the third step, the symmetric tridiagonal eigenvalue problem is solved by a bisection method based on the use of Sturm sequences by restricting the search range using the Gershgoring circle theorem [11].

The motivation for computing the tridiagonal matrix $\underline{\underline{T}}$ as above is due to the fact that the explicit formation of $\underline{\underline{T}}$ should be avoided for numerical reasons since it may introduce non-negligible relative errors, especially in the computation of the smallest SVs [12]. However, the exploited approach is meant for those applications, as the one at hand, in which the smallest SVs have very low relative weight and may be considered irrelevant.

In summary, the problem of computing the SVs is recast as a "guided" bisection, which is amenable to parallelization.

VI. NUMERICAL RESULTS

The optimizer, the ray tracer and the SVs calculation have been implemented in parallel GPU (CUDA) and multi-core CPU (C++ with OpenMP directives) languages. For the CPU case, the SVs have been achieved using the third party Eigen library.

We consider a circular cylinder with radius 10λ centered at the origin of the $Oxyz$ reference system and a plate of width 50λ, parallel to the yz plane and located at $x=30\lambda$. The cylinder and the plate are perfectly conducting and have a height of 15λ. The scene has been discretized with 95458 triangles. This example points out how much computation time can be saved by the approach and provides an answer to the points raised in Section I and a perspective to design tools.

The optimizer can position an arbitrary number of transmitting and receiving antennas on lines with arbitrary spatial orientations. Here, $N_{TX} = 4$ and $N_{RX} = 4$. The TX and RX antennas have been located on lines lying on the xy plane, parallel to the x-axis and passing by $(15\lambda, -40\lambda, 0\lambda)$ for the TX and by $(15\lambda, 40\lambda, 0\lambda)$ for the RX antenna. The antenna positions have been represented using $K=3$ and minimum and maximum spacing of $\lambda/4$ and 2λ, respectively, have been enforced to control the maximum array size and mutual coupling. The SNR has been fixed to 20 dB.

For computational convenience, the optimization has been run with a population of 1000 elements, grouped in 10 subgroups including those configurations sharing the same positions of the TX antennas and different positions of the RX ones. The optimization has been run for a number of 50 generations, with $CR=0.4$ and $F=0.7$.

The code has been run on a workstation equipped with two Intel Xeon E5-2650 2.00GHz, Eight core processors each and four NVIDIA Kepler K20c cards, but with multi-GPU disabled. Figure 2 displays an OpenGL rendering of the MIMO channel with the optimized antenna locations. The figure also depicts the GO rays connecting the TX antennas (the pink dots) with the RX ones (not appearing in the image). As it can be seen, multiple interactions have been accounted for as well as diffraction from the plate border. Diffraction from the plate corners and the cylinder ends have been neglected for simplicity. The optimized antenna positions are reported in Table 1. As it can be seen, the TX and RX antennas occupy almost symmetric locations due to the problem symmetry. The GPU code has run in about 4.5 hours, gaining a speedup of about 5 as compared to the CPU execution obtained by running 32 CPU threads. The optimized channel capacity has been 22.3 bps/Hz, a value which well agrees with the statistical distribution of channel capacities for random channels with 4 transmitting and 4 receiving antennas reported in [2, Fig. 7].
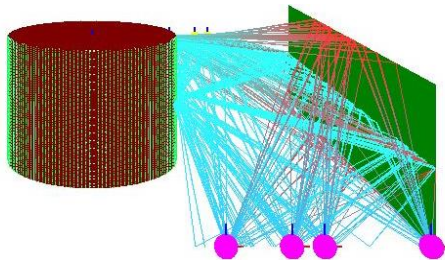

Fig. 2. OpenGL rendering of the MIMO channel with optimized antenna locations.

Table 1: Optimized TX and RX antenna positions

Antenna	x-coord.	Antenna	x-coord.
TX 1	7.0λ	RX 1	7.08λ
TX 2	12.6λ	RX 2	12.9λ
TX 3	18.0λ	RX 3	17.1λ
TX 4	20.9λ	RX 4	22.0λ

VII. CONCLUSIONS

A GPU-based approach to accelerate MIMO CC optimization has been presented using ultrafast Geometrical Optics (GO), a very fast SV calculation scheme and a parallel version of the differential evolutionary algorithm. A speedup of 5 has been achieved against a multi-core CPU implementation.

ACKNOWLEDGMENT

REFERENCES

[1] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Commun.*, vol. 6, no. 3, pp. 311-335, Mar. 1998.

[2] J. Bach Andersen, "Array gain and capacity for known random channels with multiple element arrays at both ends," *IEEE J. Selected Areas Commun.*, vol. 18, no. 11, pp. 2172-2178, Nov. 2000.

[3] U. Olgun, et al., "Optimization of linear wire antenna arrays to increase MIMO channel using swarm intelligence," *Proc. of the 2nd Europ. Conf. on Antennas Prop.*, Edinburgh, UK, pp. 1-6, Nov. 11-16, 2007.

[4] M. A. Mangoud, "Optimization of channel capacity for indoor MIMO systems using genetic algorithm," *Progr. Electromagn. Res. C*, vol. 7, pp. 137-150, 2009.

[5] E. G. Larsson, et al., "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186-195, Feb. 2014.

[6] N. Noori and H. Oraizi, "Evaluation of MIMO channel capacity in indoor environments using vector parabolic equation method," *Progr. Electromagn. Res. B*, vol. 4, pp. 13-25, 2008.

[7] R. Storn and K. Price, "Differential evolution - A simple and efficient heuristic for global optimization over continuous spaces," *J. Global Opt.*, vol. 11, no. 4, pp. 341-359, Dec. 1997.

[8] A. Capozzoli, et al., "Field sampling and field reconstruction: a new perspective," *Radio Sci.*, vol. 45, RS6004, pp. 31, 2010, doi: 10.1029/ 2009RS004298.

[9] A. Capozzoli, et al., "FFT & aperiodic arrays with phase-only control and constraints due to super-directivity, mutual coupling and overall size," *Proc. of the 30th ESA Antenna Workshop on Antennas for Earth Observ., Science, Telecomm. and Navig. Space Missions*, Noordwijk, The Netherlands, May 27-30, 2008, CD ROM.

[10] A. Breglia, et al., "Comparison of acceleration data structures for electromagnetic ray tracing purposes on GPUs," *IEEE Antennas Prop. Mag.*, vol. 57, no. 5, pp. 159-176, Oct. 2015.

[11] A. Capozzoli, et al., "Massive computation of singular values of small matrices on GPUs," *Proc. of the Int. Workshop on Comput. Electromagn.*, Izmir, Turkey, pp. 36-37, July 1-4, 2015.

[12] G. H. Golub and C. Reinsch, "Singular values decomposition and least squares solutions," *Numer. Math.*, vol. 14, pp. 403-420, 1970.

Fast and Parallel Computational Techniques Applied to Numerical Modeling of RFX-mod Fusion Device

Domenico Abate [1,2], Bruno Carpentieri [3], Andrea G. Chiariello [4], Giuseppe Marchiori [2], Nicolò Marconato [2], Stefano Mastrostefano [1], Guglielmo Rubinacci [5], Salvatore Ventre [1], and Fabio Villone [1]

[1] DIEI, Università di Cassino e del Lazio Meridionale, Loc. Folcara, 03043 Cassino (FR), Italy
s.mastrostefano@unicas.it, ventre@unicas.it, villone@unicas.it

[2] Consorzio RFX, Corso Stati Uniti 4, Padova, Italy
domenico.abate@igi.cnr.it, giuseppe.marchiori@igi.cnr.it, nicolo.marconato@igi.cnr.it

[3] Nottingham Trent University, School of Science and Technology, Burton Street, Nottingham NG1 4BU, UK
bruno.carpentieri@ntu.ac.uk

[4] DIII, Seconda Università di Napoli, Via Roma 29, Aversa (CE), Italy
andreagaetano.chiariello@unina2.it

[5] DIETI, Università di Napoli Federico II, Via Claudio 21, 80125, Napoli
rubinacci@unina.it

Abstract ─ This paper presents fast computational techniques applied to modelling the RFX-mod fusion device. An integral equation model is derived for the current distribution on the active coils of the conducting structures, and the input-output transfer functions are computed. Speed-up factors of about 200 can be obtained on hybrid CPU-GPU parallelization against uniprocessor computation.

Index Terms ─ Fusion plasma devices, GPUs, HPC, integral formulation, parallelism.

I. INTRODUCTION

Modelling fusion devices is computationally very challenging due to the electromagnetic interaction of the fusion plasma and the surrounding conducting structures, which makes the problem inherently multiphysics. The evolution of the plasma may exhibit unstable modes, thus exacerbating the aforementioned problems and requiring a feedback controller. The design of such control system requires rather accurate response model of the overall system plasma plus conductors. Therefore, fast parallel techniques are often required to make the computations affordable [1, 2]. In this paper, we analyze the RFX-mod device [3], a medium size (major radius R = 2 m, minor radius a = 0.46 m) toroidal device particularly suited to explore innovative concepts in plasma control. Passive and active conductors are very important to determine the overall properties and performances of such feedback system and therefore they should also be adequately represented in any realistic model. The main conducting structures are the vessel (needed to have the vacuum inside the machine), the shells (highly conducting sheets needed for passive stabilization), the mechanical structure, hosting the active control coils. Figure 1 shows the 3D hexahedral mesh used.

In particular, RFX-mod is equipped with a state-of-the-art control system made by 192 (4 poloidal x 48 toroidal) independently fed active coils (Fig. 1), with more than 600 magnetic sensors acquired in real time. This makes RFX-mod on the one hand very challenging for numerical modelling but on the other hand an ideal test-bed for validating the predicting capabilities of computational tools. We compute the input-output transfer functions of the system, assuming as input the currents or the voltages of the active coils and as output suitable magnetic measures [4]. The presence of an axisymmetric plasma evolving through equilibrium states is self-consistently taken into account [1].

The computer solution of such a problem is very expensive, due to the complexity of the 3D geometry and the plasma contribution. The use of High Performance Computing (HPC) cluster is mandatory. The GPU architecture has a large amount of cores designed to run a large number of execution threads at the same time; the computational model used is the single instruction,

multiple data (SIMD), where concurrent threads execute the same code (called Kernel) on different data. In the present work, we focus our attention on a hybrid multi-node system for modeling RFX-mod devices.

The paper is organized as follows. Section II describes the model, while in Section III we illustrate the computational technique. Section IV reports the results and draws the conclusions.
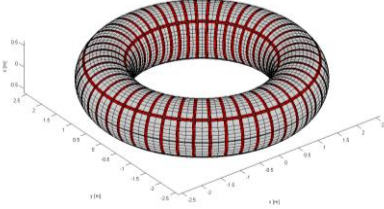

Fig. 1. Mesh used for the analysis of the problem.

II. MODEL

We consider a system of 3D conductors V_c discretized with a finite elements mesh. We use an integral formulation, which assumes as primary unknown the current density in V_c. We introduce the electric vector potential $\mathbf{T}$, such that $\mathbf{J} = \nabla \times \mathbf{T}$, and then we expanded $\mathbf{T}$ in terms of edge elements $\mathbf{N}_k$, we have:

$$\mathbf{J} = \sum_k I_k \nabla \times \mathbf{N}_k . \tag{1}$$

Imposing Ohm's law in weak form, we get [1,2,8]:

$$\underline{\underline{L}} \frac{d\underline{I}}{dt} + \underline{\underline{R}}\,\underline{I} + \frac{d\underline{U}}{dt} = \underline{V}, \tag{2}$$

$$L_{i,j} = \frac{\mu_0}{4\pi} \int_{V_c} \int_{V_c} \frac{\nabla \times \mathbf{N}_i(\mathbf{r}) \cdot \nabla \times \mathbf{N}_j(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} dV\,dV', \tag{3}$$

$$R_{i,j} = \int_{V_c} \nabla \times \mathbf{N}_i \cdot \boldsymbol{\eta} \cdot \nabla \times \mathbf{N}_j dV . \tag{4}$$

In these equations, $\underline{I}$ is the vector of degrees of freedom I_k in (1), $\underline{V}$ is the vector of externally applied voltages and $\boldsymbol{\eta}$ is the resistivity tensor. Matrix $\underline{\underline{L}}$ is a fully populated square matrix, which is the 3D analogue of mutual inductance of a system of magnetically coupled conductors; conversely, $\underline{\underline{R}}$ matrix is sparse and represents the resistance matrix of the 3D conductors. The quantity $\underline{U}$ is the magnetic flux due to plasma currents [1, 8]:

$$\underline{U} = \underline{\underline{M}}\,\underline{j}_S , \quad \underline{j}_S = \underline{\underline{S}}\,\hat{\underline{\psi}}_e , \quad \hat{\underline{\psi}}_e = \underline{\underline{Q}}\,\underline{I} , \tag{5}$$

where $\underline{j}_S$ are equivalent currents located on a coupling surface, $\underline{\underline{M}}$ is a mutual inductance matrix between the equivalent current and the 3D conducting structures, $\hat{\underline{\psi}}_e$ is the external magnetic flux, $\underline{\underline{Q}}$ is a matrix representing Biot-Savart integral [1] and $\underline{\underline{S}}$ is the plasma response matrix [8].

Combining (2)-(5), finally we get [8]:

$$\underline{\underline{L}}^* \frac{d\underline{I}}{dt} + \underline{\underline{R}}\,\underline{I} = \underline{V}, \quad \underline{\underline{L}}^* = \underline{\underline{L}} + \underline{\underline{M}}\,\underline{\underline{S}}\,\underline{\underline{Q}}, \tag{6}$$

to which we can add the expression for the magnetic field and flux perturbations $\underline{y}$ at given points, linearly related to 3D currents through a suitable matrix $\underline{\underline{C}}$ [1,8]:

$$\underline{y} = \underline{\underline{C}}\,\underline{I} . \tag{7}$$

Equations (6)-(7) represent the model; they can be easily recast in standard state space form. In the present paper, they are used to get the frequency-domain transfer functions between the inputs (voltages or currents in active coils) and the outputs (linear combinations of magnetic measurements). In doing so, the inversion of a complex matrix is required. Indeed, we split the unknowns into three subsets; the corresponding subset of indices of the various matrices are identified with the following suffix: "p" (passive structures), "m" (measurement coils), "a" (active coils), so that Equation (6) reads as:

$$\left(j\omega \underline{\underline{L}}^*_{pp} + \underline{\underline{R}}_{pp} \right)\underline{I}_p + j\omega \underline{\underline{L}}^*_{pa}\underline{I}_a = 0,$$
$$\underline{\underline{L}}^*_{mp}\underline{I}_p + \underline{\underline{L}}^*_{ma}\underline{I}_a = \underline{\Phi}_m, \tag{8}$$

where $\underline{\Phi}_m$ represent the fluxes induced at measurements coils (i.e., the output of the system). After some simple algebraic manipulations it turns out:

$$\underline{I}_p = -j\omega \left(j\omega \underline{\underline{L}}^*_{pp} + \underline{\underline{R}}_{pp} \right)^{-1} \underline{\underline{L}}^*_{pa}\underline{I}_a = \underline{\underline{H}}(j\omega)\underline{I}_a,$$
$$\underline{\Phi}_m = (\underline{\underline{L}}^*_{mp}\underline{\underline{H}}(j\omega) + \underline{\underline{L}}^*_{ma})\underline{I}_a = \underline{\underline{T}}(j\omega)\underline{I}_a. \tag{9}$$

Equation (9) can be used to evaluate $\underline{\Phi}_m$ for a given assigned unitary current flowing in excitation coil. This transfer function can be used to design the feedback control. For each frequency and each active coil, we set I_a to 1 (the known terms in the Eq. (9)) and find the Flux for all measurement coils (unknowns variables).

III. FAST TECHNIQUES

In order to speed up the overall computation we move in two directions:
- parallelize the matrix assembly phase;
- accelerate the inversion of system (9).

A. Parallel assembly strategies

Matrices $\underline{\underline{L}}$ and $\underline{\underline{Q}}$ are very expensive to assemble. For $\underline{\underline{L}}$ matrix, parallelization can be achieved grouping elements of nodes into boxes, distributing boxes among processors, and performing the element-element integration independently on each processor. The locally assembled matrix is then compressed (see [2]).

The computation of matrix $\underline{\underline{Q}}$ is the most time consuming part of the assembly algorithm. In order to reduce this cost, in the present work, parallel assembly is implemented on multi CPUs and multi GPUs environment. Here we take advantage of the fact that

Biot-Savart integral computation for elements and field points, are independent from each other. Of course, different implementations are necessary to adapt the parallel computation to the two different hardware architectures. In the following, we briefly recall the main features of the two algorithms.

In multi CPUs environment we propose a standard parallel strategy using a simple domain decomposition approach that distributes the field points equally among the processors. After the local computations, a reduction operation is required to retrieve the complete matrix from each MPI process. This strategy scales linearly with the number of the processors.

In multi GPUs environment we propose to assign to each computational thread the evaluation of a contribution of the Biot-Savart integral corresponding to a given element and a given field point. All the contributions are summed on the CPU. The algorithm is briefly summarized in the follow, see [1] for details:
1) Allocate temporary data for storing the local contribute (CPU).
2) Compute the considered element and source point from the thread and block index (GPU).
3) Compute the shape function related to the considered element (GPU).
4) Compute the local contribution (GPU).
5) Return the partial matrix to the host memory (CPU).
6) Scatter the output data on the complete matrix on the host (CPU).

The final step is due to uncoalescent memory access needed to store the results in the final matrix and possible race conditions when two different contributions are summed in the same location. The dimension of the matrix can be huge compared to the on board GPU memory (which is typically of a few GB). Step 5 involves memory transfer from GPU to Host memory, but fortunately this has no impact on the overall performance of the code. We point to [9] for more sophisticated approaches not considered here.

B. Speed up of the linear system inversion

As far as the inversion of the linear system involved in (9) is concerned, it is worth noting that $\underline{\underline{L}}_{xy}^{*}$ are fully populated submatrices of matrix $\underline{\underline{L}}^{*}$ and $\underline{\underline{R}}_{pp}$ is a sparse positive definite matrix. Using a direct solver, the cost of the inversion procedure is $O(N^3)$, N being the number of unknowns present in the passive part of the device. When geometric details are added and/or a great accuracy is required in the computation, it is easy to exceed quickly the computational resources available on a uniprocessor system. The use of powerful computing facilities can help in the search of additional speed and increase the size of the solvable problems [5].

Nevertheless, there are cases in which parallelization fails poorly. For this problem, an approximated compression technique is mandatory. The authors successfully applied these methods for the study of plasma fusion devices [2] as well as in other fields (e.g., NDT [6]). These techniques are based on an effective low-rank approximation of the submatrix representing the far interaction between well separated parts of the device. The matrix-by-vector product $\underline{\underline{L}}_{ij}^{*} I_j$ related to these parts is replaced by an accurate low cost operator (the complexity is asymptotically only $O(N)$). Finally, the inversion in (9) can be performed by an iterative method (such as the GMRES method). It is worth noting that the preconditioner (essential for any iterative solver) is $\underline{\underline{R}}_{pp}^{-1}$, which can be computed in fast and accurate way by the means of Cholesky decomposition. It is important to stress that its factorization and back substitution is very cheap using a single CPU. Moreover the preconditioner turns out to be very effective, being the number of iterations required to converge very small.

IV. RESULTS

The computational cluster used for the evaluation of the numerical performances is made by two nodes. Each node consist of 16x cores Intel Xeon CPU E5-2690 (@ 2.90 GHz processor, 20 MB L2), 128 GB RAM, 2×NVIDIA Kepler K20 (2496 cores, 6 GB VRAM).

A. 2D validation and transfer function computation

First of all, a numerical validation of the procedure is carried out. We generated a 3D mesh which fictitiously reproduce an axisymmetric geometry, so that a 2D code (CREATE_L [7]) can be used as benchmark. We computed the transfer function $\underline{T}$ defined in the previous section with the two codes, finding a very good agreement, as shown in Fig. 2. This confirms the correctness of the procedure.

In order to show the actual effect due to the presence of the plasma, we compare the results obtained with and without plasma on the full 3D mesh described above. The plasmaless computation is in fact a purely magneto-quasi-static calculation. The number of elements of the mesh is equal to 30907, the number of nodes is 81550. The number of unknowns in the passive structure (i.e., the dimension of the matrix to invert) is 22619. The results are reported in Fig. 3. Evidently, the presence of the plasma has an effect not only on the dynamical properties of the model (e.g., the phase behavior at high frequencies), but also on the static gain (amplitude at zero frequency limit). This is not surprising, since the plasma affects also the magnetostatic coupling between active coils and sensors, because it reacts to external static magnetic field perturbations, so as to reach a different equilibrium configuration and hence, modifying the whole magnetic field map in the surrounding regions.
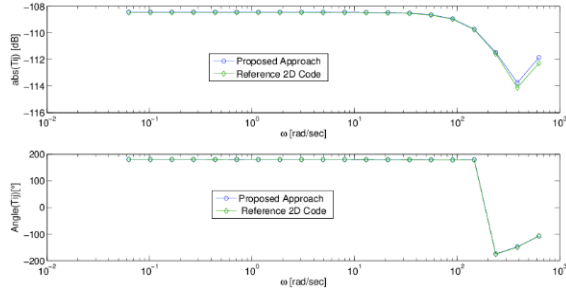
Fig. 2. Comparison of one element of the transfer function $\underline{T}$: proposed approach and reference 2D code.
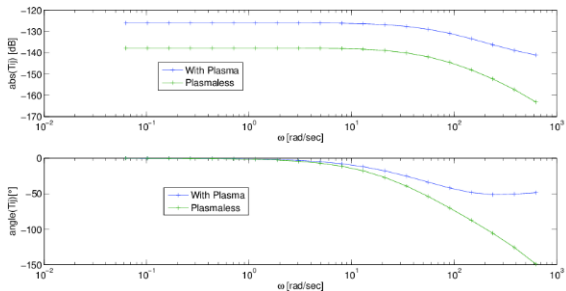

Fig. 3. Effect of plasma on the transfer function.

B. Numerical issues

Regarding the speedup of the matrix assembly, using 25 cores the time required to compute the compressed matrix $\underline{L}$ is about 90 s, the total time required to compute the plasma matrices is about 549 s (540 s of this time is due to the computation of $\underline{Q}$ matrix). In Fig. 4 we report the speedup for assembling $\underline{Q}$ matrix, defined as the assembly time required by one CPU divided by the time obtained using a parallel multi GPUs. Using standard parallel strategy (multi CPUs) the maximum achievable speed up on the proposed computational system is limited to 32.
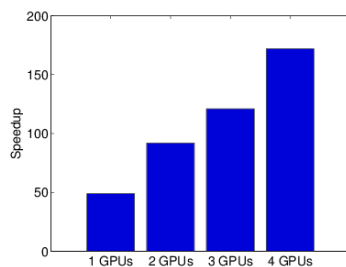

Fig. 4. Speedup for Q-matrix assembly

The time required for each single inversion is about 17.5 s. The total time for all inversions is about 7000s.

The number of iterations required by GMRES to converge increases with the excitation frequency. Without the plasma (i.e., the response due to only the passive structures) the number of iterations required by GMRES to converge is 21 at a frequency of 100 Hz and 9 at 10 Hz. If the plasma is present the number of iteration is 41 at frequency of 100 Hz and 9 at 10 Hz. This is coherent with the general expectation that the used preconditioner is more effective at lower frequencies and without plasma.

V. CONCLUSIONS

We have presented fast parallel techniques for the computation of input-output transfer functions on the RFX-mod fusion devices on hybrid architectures, featuring multiple CPUs and GPUs. The peculiarities of fusion devices make this approach particularly effective in significantly improving the performances of the computation, allowing speed-ups up to almost 200 with respect to standard computations.

REFERENCES

[1] F. Villone, A. G. Chiariello, S. Mastrostefano, A. Pironti, and S. Ventre, "GPU-accelerated analysis of vertical instabilities in ITER including three-dimensional volumetric conducting structures," *Plasma Phys. Control. Fusion*, vol. 54, no. 8, 2012.

[2] G. Rubinacci, S. Ventre, F. Villone, and Y. Liu, "A fast technique applied to the analysis of resistive wall modes with 3D conducting structures," *Journal of Comp. Phys.*, vol. 228, no. 5, pp. 1562-1572, 2009.

[3] P. Sonato, et al., *Fusion Eng. Des.*, vol. 66-68, pp. 161, 2003.

[4] F. Villone, et al., "ITER passive and active RWM analysis with the CarMa code," *38th EPS Conference*, paper P5.107, 2011.

[5] R. Fresa, G. Rubinacci, and S. Ventre, "An eddy current integral formulation on parallel computer systems," *Int. Journal for Numerical Methods in Engineering*, vol. 62, no. 9, pp. 1127-1147, 2005.

[6] G. Rubinacci, A. Tamburrino, and S. Ventre, "Fast numerical techniques for electromagnetic non-destructive evaluation," *Nondestr. Testing Eval.*, vol. 24, pp. 165-194, 2009.

[7] R. Albanese and F. Villone, "The linearized CREATE-L plasma response model for the control of current, position and shape in tokamaks," *Nucl. Fusion*, vol. 38, no. 5, pp. 723, 1998.

[8] A. Portone, et al., "Linearly perturbed MHD equilibria and 3D eddy current coupling via the control surface method," *Plasma Phys. Control. Fusion*, 50, 085004, 2008.

[9] A. Capozzoli, et al., "Speeding up aperiodic reflectarray antenna analysis by CUDA dynamic parallelism," *Proc. of the Int. Conf. on Numerical Electromagn. Model. and Opt. for RF, Microwave and Terahertz Appl.*, Pavia, Italy, pp. 1-4c, 2014.

Parallel Implementations of Multilevel Fast Multipole Algorithm on Graphical Processing Unit Cluster for Large-scale Electromagnetics Objects

Nghia Tran and Ozlem Kilic

Department of Electrical Engineering and Computer Science
The Catholic University of America, Washington, DC, 20064, USA
16tran@cua.edu, kilic@cua.edu

Abstract ─ This paper investigates solving large-scale electromagnetic scattering problems by using the Multi-level Fast Multipole Algorithm (MLFMA). A parallel implementation for MLFMA is performed on a 12-node Graphics Processing Unit (GPU) cluster that populates NVidia Tesla M2090 GPUs. The details of the implementations and the performance achievements in terms of accuracy, speed up, and scalability are shown and analyzed. The experimental results demonstrate that our MLFMA implementation on GPUs is much faster than (up to 37x) that of the CPU implementation.

Index Terms ─ Graphics Processing Unit (GPU), Multilevel Fast Multipole Algorithm (MLFMA).

I. INTRODUCTION

Over the past twenty years, various numerical techniques have been developed to reduce the computational time and memory requirements of full-wave electromagnetic models without significant loss of accuracy, including adaptive integral method (AIM) [1], impedance matrix localization (IML) [2], fast multipole method (FMM) [3], and multi-level fast multipole algorithm [4]. Compared with the others, MLFMA is among the most suitable techniques for large-scale problems. It reduces the computational complexity of the method of moments (MoM) from $O(N^2)$ to $O(NlogN)$, where N denotes the number of unknowns, whereas AIM, IML and FMM have the complexities of $O(N^{3/2}logN)$, $O(N^2logN)$, and $O(N^{3/2})$, respectively.

Recently, many authors have investigated the parallelization of MLFMA on CPU clusters [5] in solving problems of hundreds of thousands to millions of unknowns. In [6], CPU clusters were used to implement MLFMA using Open MP and MPI library to solve a billion unknowns. Multi-GPU implementation was also investigated on a single node, multi-GPU computer without using the MPI library [7]. In this paper, we demonstrate the implementation of MLFMA for electromagnetics problems on GPU clusters by using the MPI library.

We demonstrate the parallelization of MLFMA on a 12-node GPU cluster each of which is populated with an NVidia Tesla M2090 GPU. An MVAPICH2 implementation of MPI is used for cluster parallel programming. This paper is the continuation of our GPU implementation of FMM by using GPU clusters. In [9] and [10], GPU implementation for single level Fast Multipole Method (FMM) solves only the maximum problem size up to 656K unknowns on 13 nodes. In this paper, our MLFMA implementation on GPU cluster can solve up to 1.1 M unknowns. We demonstrate that the implementation of MLFMA on GPUs is faster than that of the CPU. The performance of the implementation is analyzed by using a PEC sphere.

The rest of the paper is organized such that Section II provides an overview of MLFMA. Section III presents the parallel implementation of MLFMA on GPU clusters. Experimental results are discussed in Section IV, followed by the conclusions in Section V.

II. OVERVIEW OF THE MULTILEVEL FAST MULTIPOLE ALGORITHM

The fundamental principles of MLFMA and its applications in electromagnetics have been studied in literature [4]-[5]. In this section, we provide a brief overview to help our discussion on its parallel implementation, which is presented in Section III.

MLFMA was invented based on the grouping concept to accelerate the iterative solution of the linear equation system ZI = V of the Method of Moment (MoM), where I represents the unknown currents, V depends on the incident field, and Z is the impedance matrix. The main idea of the grouping concept is shown in Fig. 1, where the M edges in the mesh of a given structure are categorized into an N-level tree structure connecting groups of different sizes from the finest (level N) to the coarsest level (level 0). Based on the groups' proximity, the impedance matrix Z can be split into two matrices, Z^{near} and Z^{far}, corresponding to near and far interactions as shown in Equation (1):

$$\sum_{m'}^{M} Z_{mm'} I_{m'} = \sum_{m'}^{M} Z_{mm'}^{near} I_{m'} + \sum_{m'}^{M} Z_{mm'}^{far} I_{m'} = V_m, \qquad (1)$$

where m and m' are observation and source edges in the

mesh, respectively.

The Z^{near} matrix comprises of interactions between edges in spatially nearby groups, and is computed and stored using the conventional MoM [8]. During the iterative solution, the near matrix is calculated by the regular sparse matrix-vector multiplications (MVMs). The remaining edges, whose parents are near, constitute the far term as shown in Fig. 1 (b). By treating the interactions between the edges that are spatially far-away using MLFMA, Z^{far} matrix does not need to be explicitly computed and stored. Instead, the far components can benefit from the fast MVMs during the iterative solution. The Z_{far} matrix is factorized into radiation, receive and translation functions, as explained in [4].
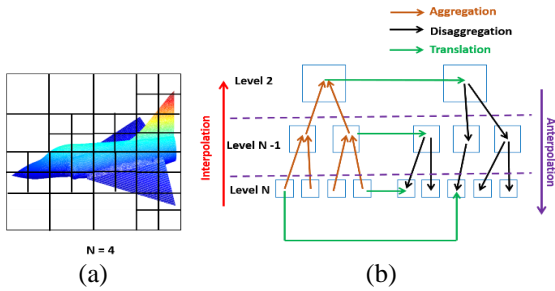

Fig. 1. MLFMA general concepts: (a) concept of the MLFMA tree, and (b) MLFMA concept of far interactions.

The far component is calculated through five main stages: aggregation, translation, and disaggregation, interpolation and anterpolation as shown in Fig. 1 (b).

In the aggregation stage, radiated fields among the groups from level N (the finest level) to level 2 are calculated. At the finest level N, the radiation functions for a group are computed by combining the radiation patterns of the basic function of all edges in this group. From level N-1 to level 2, the radiation functions for each group are computed from the combination of the radiation function of its children group of the finer level using shifting and interpolation.

In the disaggregation stage, the receive functions at each group are computed from level 2 to level N by combining the local incoming waves due to translation and the incoming waves from parent groups of the coarser level using shifting and anterpolation.

The translation stage is identical to FMM [3], and the details of interpolation and anterpolation can be found in [5].

III. PARALLELIZATION OF MLFMA ON GPU CLUSTERS

In this section, we provide an overview of our implementation on GPU. The implementation consists of pre-processing, processing and post-processing. The geometry mesh data resulting from the pre-processing step is transferred to the GPU memory, and the entire computation is performed on the GPU. The user defined

results such as radar cross section, scattered fields are post-processed on CPU.

The GPU cluster used for our implementation consists of 12 computing nodes. Each node has a dual 6-core 2.66 GHz Intel Xeon processor, 48GB RAM along with one NVidia Tesla M2090 GPU running at 1.3 GHz supported with 6GB of GPU memory. The nodes are interconnected through the InfiniBand interconnection. The cluster populates CUDA v6.0 and MVAPICH2 v1.8.1 (a well-known implementation of Message Passing Interface (MPI)).

In the processing step, the workload of the computational task is equally distributed among the computing nodes, and the inter-node communication is minimized. This is achieved by uniformly distributing the total number of groups, M, among the n computing nodes. The parallelization of the GPU cluster implementation is performed at two levels: (i) among the computing nodes using MPI library, and (ii) within the GPU per node using CUDA programming model. Within each node, the CUDA thread-block model is utilized to calculate the workload assigned to that node. We only present the far interactions in this paper, since the near field and V vector calculations implementations can be found in [9]-[10].

All CUDA kernels are implemented to calculate Z^{near} matrix, and far interactions which includes the radiation/receive functions, translation matrix, and interpolation/anterpolation matrices. In fast matrix-vector multiplication (MVM), CUDA kernel is also utilized to compute the radiated fields, translation fields and received fields in the aggregation, translation and disaggregation stages, respectively. MPI library is also used to gather results from each node in the end of MVM stage.

A. Far interactions calculations

This task comprises of five calculations: radiation, and receive functions, interpolation, anterpolation and translation matrices.

(i) Radiation and Receive Function Calculations

The first step in the far interaction calculations is the calculation of the radiation, T^E, and receive, R^E, functions for Z^{far} matrix. They are complex conjugates of each other. Thus their implementations are similar. Following the M group distribution, each node handles the calculations of K directions for M_{node} groups. Given this amount of workload per node, the CUDA kernel is launched with $M_{node}.K$ blocks such that each block implements M_{group} radiation/receive function calculations at a given direction, resulting in a total of $M_{node}.K$ blocks per node.

(ii) Translation Matrix Calculation

The second task for far interactions is the calculation of the translation matrix, T_L. The workload for the T_L calculations is also distributed across the nodes following

the group-based technique. By careful investigations, allocating a CUDA block on a single row of the matrix is the efficient way for the translation matrix calculation to save memory requirements. Each CUDA block is assigned to compute one sparse row of the T_L matrix for a given direction, and each thread computes one element in that row.

(iii) Interpolation and Anterpolation Matrices

The third task for the far interactions is the calculation of interpolation and anterpolation matrices. They are transposes of each other. Thus their implementation is similar. Each node handles the calculations of $K_{children/node}$ rows of $K_{children}$xK_{parent} interpolation matrix, where $K_{children}$ is number of directions of a finer level, and K_{parent} is number of directions of a coarser level. The CUDA kernel is launched with $K_{children/node}$ blocks per node. In each block, the maximum number of threads (1024 threads) are utilized in order to implement the full number of K_{parent} directions.

B. Fast matrix-vector multiplication

The next stage for the processing is the solution for the linear system where we employ the iterative method known as the biconjugate gradient stabilized method (BiCGSTAB). The calculation of $Z^{far}I$ comprises of five stages: aggregation, translation, interpolation, anterpolation and disaggregation, as shown in Fig. 2. Using a group-based partitioning technique, the unknown current vector I (N_{edges}x1) is distributed across the computing nodes on GPU clusters.
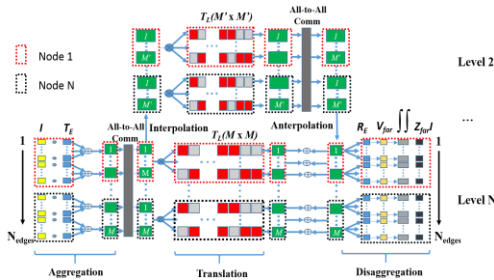

Fig. 2. Far matrix-vector-multiplication in parallel.

In the aggregation stage, at level N, each node computes the radiated fields for M_{node} groups for K directions by multiplying the unknowns I with their corresponding radiation functions, T^E, and accumulating within each group. After the aggregation step, an all-to-all communication is employed by each node to broadcast the radiated fields to all other nodes. The radiated fields from level N-1 to level 2 are computed by multiplying interpolation matrices with radiated fields of children groups at lower levels.

In the translation, the radiated fields at each direction are calculated from the sum of the multiplication of the translation matrix and the radiated fields, and the

received fields from parent groups at upper levels using anterpolation.

In the disaggregation stage, the received fields of all M group at level N are multiplied with the corresponding receive functions, and integrated over the partitioned K directions of the unit sphere. The far components of MVM are then incorporated with the near components of MVM. At the end of MVM, the partial results from all nodes are summed together and all nodes are updated.

IV. EXPERIMENTAL RESULTS

A. Accuracy

First, we verify the accuracy of our GPU implementation by calculating the radar cross section (RCS) of a 9λ diameter (corresponding to 0.27 m and 100,000 unknowns) perfect electrically conducting (PEC) sphere illuminated by an 1 GHz x-polarized normally incident field. The results are compared to Mie scattering. It can be observed in Fig. 3 that the GPU results and the analytical solutions show a very good agreement.
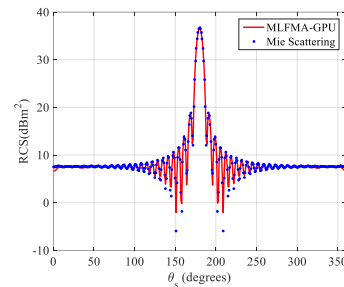

Fig. 3. RCS of a 9λ diameter PEC sphere.

B. Implementation performance on GPU cluster

In the first experiment, our GPU implementation is evaluated using the fixed-workload model (Amdahl's Law). A 22.4λ diameter PEC sphere (650K unknowns) is chosen such that it demands the use of at least 7 nodes to satisfy the required memory. Two metrics are used for the performance evaluation: speed up and scalability. The speed up is defined as the ratio of time required by multi-node GPU implementation with respect to the 7-node CPU implementation. Scalability is the normalized speedup of multiple nodes in reference to the speedup of 7 nodes. As shown in Fig. 4, the speedup factor increases from 23.7 for 7 nodes to 37 for 12 nodes. Since each node processes less workload, the GPU execution time decreases as the number of nodes increases. The inter-node communication overhead results in the difference between the speedup of total execution time and computation time. For 7 computing nodes, the speed-up for the near-field system matrix is over 86 (CPU computation time: 848s, GPU computation time: 9.5s), while the speed-up of the BICGstab iterative solution is over 22 times for 100 iterations, which is restricted by the overhead communication between computing nodes (CPU computation time: 9100s, GPU

computation time: 415.1s).

In order to investigate the scalability of this implementation, we compare how the speedup improves with increasing number of computing nodes as we keep the problem size constant, as observed in Fig. 5. The computation speedup scales similar to the theoretical linear behavior, demonstrating our efficient hardware implementation. The total speedup scales closely to the theoretical expectation demonstrating our efficiency in reducing the inter-node communication overhead.
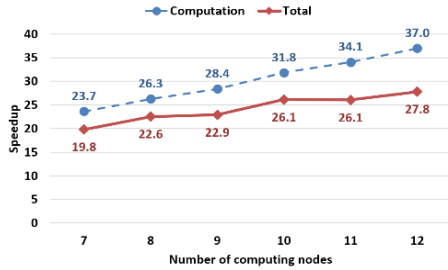

Fig. 4. Speedup analysis for the fixed-workload model (vs. 7 nodes CPU implementation, 100 iterations). Computational CPU exec time = 5573 sec, total CPU exec time = 5627 sec.
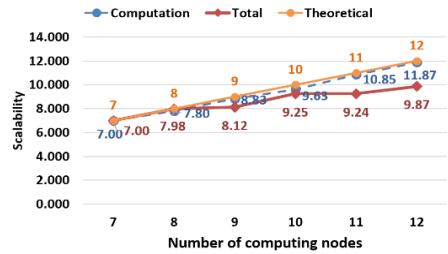

Fig. 5. Scalability analysis for the fixed-workload model.

In the second experiment, we investigate the largest problem size our GPU implementation can handle. As the number of nodes increases, the problem size is also increased so that the GPU memory in each node in fully utilized. As shown in Fig. 6, the GPU implementation can process a maximum problem size of 1.1 M unknowns with a speed up factor of 25.2.
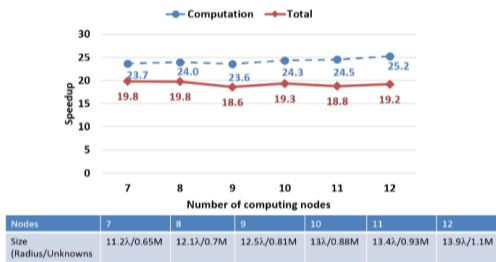

Fig. 6. Speedup analysis when the number of nodes increases along with problem size increases (vs. multi-node CPU, 100 iterations).

VI. CONCLUSION

In this paper, the GPU implementation of MLFMA for electromagnetic scattering problems up to 1.1 million unknowns using our 12-node GPU cluster is demonstrated. The maximum problem size is determined by the available on-board GPU memory. For the same degree of accuracy, the GPU implementation outperforms the CPU implementation. Moreover, the GPU implementation has a good scalability as the number of computing nodes increases.

REFERENCES

[1] E. Bleszynski, M. Bleszynski, and T. Jaroszewicz, "AIM: Adaptive integral method for solving large-scale electromagnetic scattering and radiation problems," *Radio Science*, vol. 31, no. 5, pp. 1225-1251, 1996.

[2] F. X. Canning, "The impedance matrix localization (IML) method for moment-method calculations," *IEEE Ant. Prop. Mag.*, vol. 32, no. 5, pp. 18-30, 1990.

[3] R. Coifman, V. Rokhlin, and S. Wandzura, "The fast multipole method for the wave equation: A pedestrian prescription," *IEEE Antennas Propagat. Mag.*, vol. 35, no. 3, pp. 7-12, June 1993.

[4] J. M. Song and W. C. Chew, "Multilevel fast multipole algorithm for solving combined field integral equations of electromagnetic scattering," *Microw. Opt. Tech. Lett.*, vol. 10, pp. 14-19, Sep. 1995.

[5] O. Ergul and L. Gurel, "Efficient parallelization of the multilevel fast multipole algorithm for the solution of large-scale scattering problems," *IEEE Trans. Antennas Propag.*, vol. 56, no. 8, pp. 2335-2345, Aug. 2008.

[6] X.-M. Pan, W.-C. Pi, M.-L. Yang, Z. Peng, and X.-Q. Sheng, "Solving problems with over one billion unknowns by the MLFMA," *Antennas and Propaga. IEEE Trans. on*, vol. 60, no. 5, pp. 2571-2574, 2012.

[7] J. Guan, S. Yan, and J.-M. Jin, "An OpenMP-CUDA implementation of multilevel fast multipole algorithm for electromagnetic simulation on multi-GPU computing systems," *Antennas and Propaga., IEEE Trans. on*, vol. 61, no. 7, pp. 3607-3616, 2013.

[8] S. M. Rao, D. R. Wilton, and A. W. Glisson, "Electromagnetic scattering by surfaces of arbitrary shape," *IEEE Trans. Antennas Propag.*, vol. AP-30, no. 3, pp. 409-418, May 1982.

[9] Q. M. Nguyen, V. Dang, O. Kilic, and E. El-Araby, "Parallelizing fast multipole method for large-scale electromagnetic problems using GPU clusters," *Antennas and Wireless Propagation Letters, IEEE*, vol. 12, pp. 868-871, 2013.

[10] V. Dang, Q. Nguyen, and O. Kilic, "Fast multipole method for large-scale electromagnetic scattering problems on GPU cluster and FPGA-accelerated platforms," *Applied Computational Electromagnetics Society Journal*, vol. 28, no. 12, 2013.

Effect of Lorentz Force on Motion of Electrolyte in Magnesium Electrolysis Cell

Cheng-Lin Liu [1], Ze Sun [1], Gui-Min Lu [1], Xing-Fu Song [1], and Jian-Guo Yu [1,2]

[1] National Engineering Research Center for Integrated Utilization of Salt Lake Resource
East China University of Science and Technology, Shanghai, 200237, China
liuchenglin@ecust.edu.cn, zsun@ecust.edu.cn, gmlu@ecust.edu.cn, xfsong@ecust.edu.cn

[2] State Key Laboratory of Chemical Engineering
East China University of Science and Technology, Shanghai, 200237, China
jgyu@ecust.edu.cn

Abstract *–* Magnesium production process is highly energy intensive. Electrolysis process provides an effective route to reduce the energy consumption. In this paper, a three-dimensional electro-magneto-hydro-dynamics coupling model of a 120 kA magnesium electrolysis cell using finite element method is presented. In this model, the electric field, magnetic field, and flow field are included. This paper concerns the effects of the Lorentz force on the motion of the electrolyte in the cell. The model predicts that the magnitude of Lorentz force is at its maximum near the region between the anode and cathode. The direction of the Lorentz force is beneficial to the motion of the electrolyte in the magnesium electrolysis cell.

Index Terms *–* Electro-magneto-hydrodynamics, Lorentz force, magnesium electrolysis cell.

I. INTRODUCTION

Magnesium has found a variety of applications due to a number of advantages including low mass density and high specific strength. Like the Hall-Herault process of aluminum production [1], the electrolysis process for magnesium is one of the most energy intensive industrial processes [2]. Over the years, lots of research efforts have been made on the investigations of the flow field, thermoelectric field, electro-hydrodynamic field, thermoelectromechanical model and magneto-hydrodynamic model by using the commercial software packages [3-7].

Over the years, much attention has been paid on aluminum reduction cell. Little effort, however, has been made on the magnesium electrolysis cell. Shilova and Shcherbinin investigated the distribution of the electromagnetic field with the effects of bus bar and electrode in the magnesium electrolysis cell [8]. The research indicated that the magnetic field will help to improve the circulation and convection of the electrolyte.

Recent years some researches on the multi-physical fields including the electric field, magnetic field and flow fields in magnesium electrolysis cells have been reported [9, 10]. In summary, most of the reported studies of magnesium electrolysis cell only considered the mathematical model based on one physical fields. But little progress has been made on the effect of the Lorentz force on the motion of the electrolyte by using a 3D full cell coupling model of electro-magneto-hydrodynamics fields.

This paper presents an Electro-magneto-hydrodynamics model for the magnesium electrolysis cells to investigate the distributions of electric filed, magnetic field and flow field simultaneously. Moreover, the main objective of the article is to show the Lorentz force distribution throughout the cell and its effects on the motion of the electrolyte.

II. DESCRIPTION OF NUMERICAL SIMULATION

A. Structure of magnesium electrolysis cell

In the present article, a 3D full cell model of 120 kA commercial magnesium electrolysis cell with a set top entry of graphite anodes and a set side entry of steel cathodes typically consists of the molten electrolyte of $MgCl_2$, massive refractory lining, thermal insulating materials, asbestos board, steel shell, capping, and partition wall. The structural parameters used in this work have been reported elsewhere and only a brief description will be given here [11].

B. Governing equations

In the electrolysis process, DC current is fed from the anodes, and flow out from the cathodes after passing through the electrolyte. The study consider the electric field, magnetic field and flow field as the main physical fields in the model. To ensure the feasibility of the model, the following hypotheses are made:

(a) The model only focuses on the resistance voltage without considering the voltage for the decomposition of magnesium chloride, overvoltage, and contact voltage drop in the cell.

(b) Anodes are assumed to share all of the current in cells equally.

(c) All the magnetic line are in the air region.

The problem of electromagnetic analysis is solving Maxwell's equations subject to certain boundary conditions. The Ohm's law is used to predict current distribution as follows:

$$J = \sigma(E + v \times B). \tag{1}$$

The Lorentz force of the electrolyte is:

$$F = J \times B. \tag{2}$$

Magnetic induction is used for magnetic flux density calculation as follow:

$$\frac{\partial B}{\partial t} = \nabla \times (v \times B) + \frac{1}{\sigma\mu}\nabla^2 B. \tag{3}$$

$\nabla \times (v \times B)$ is negligible in comparison with other terms, and reduces to the following:

$$\frac{\partial B}{\partial t} = \frac{1}{\sigma\mu}\nabla^2 B. \tag{4}$$

On the supposition that the molecular viscous stress tensor can be neglected in comparison with the turbulent stress tensor, the momentum equation is given as follows:

$$\nabla \times v = 0, \tag{5}$$

$$\frac{\partial(\rho v)}{\partial t} + (\rho v \cdot \nabla)v = -\nabla p + v\nabla^2 v + \rho g + F. \tag{6}$$

The k-ε model is used for computing velocity profile of electrolyte using the Lorentz forces (F) as source term. The k-ε model is a class of turbulent model, called the two-equation model, where the isotropic eddy viscosity is characterized by the turbulent kinetic energy (k) and its dissipation rate (ε), and the equations can be modified to satisfy no-slip boundary conditions at the walls.

B. Boundary conditions

The Neumann boundary condition is adopted at each top of anodes, with the normalized current density of inward current flow equals to the current intensity divided by the working area:

$$-n \cdot J = J_n. \tag{7}$$

A voltage potential of zero is set at end of cathodes:

$$V = 0. \tag{8}$$

The magnetic vector potential at all of the exterior surfaces of air region is zero:

$$n \times A = 0. \tag{9}$$

The solutions to these equations are carried out by a finite element software of COMSOL. An optimum number of elements were chosen when two consecutive grid refine elements yield an error less than 1% on both magnetic field and Lorentz force calculations.

III. RESULTS AND DISCUSSION

A. Electromagnetic model validation

It is necessary to validate the accuracy of the mathematical model developed in this article, before using them for numerical experiments. The predicted results are validated by an electromagnetic coil. The coils are shown in Fig. 1 and its structure parameters are shown in Table 1.

Table 1: Technical specifications of the original vice coil of 24009 and 24010

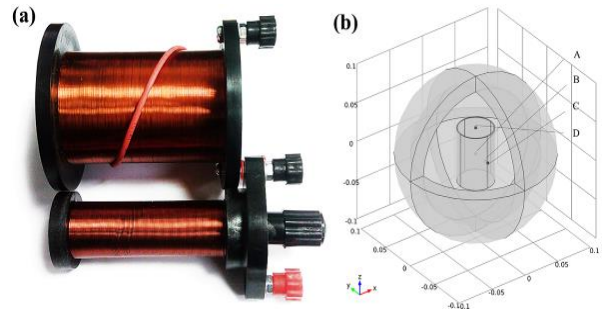| Coil Type	Internal Diameter (mm)	External Diameter (mm)	Length (mm)	Wire Diameter （mm）	Number of Coils
24009	42.9	45.1	66.5	0.22	1150
24010	27.0	29.2	51.7	0.22	370

Fig. 1. Electromagnetic coil: (a) experiment, and (b) mathematical model.

The comparison of the magnetic flux density between the experiments and simulations in the position A, B, C, and D of original vice coils are listed in Table 2. Almost all the relative error in the four test points are less than 10%, which shows that the mathematical model can predict the electromagnetic field accurately.

Table 2: Magnetic flux density of experiments and simulations in the different positions of original vice coils

Coil Type	Voltage/ V	A/Gs			B/Gs			C/Gs			D/Gs		
		Exp.	Simul.	Relative Error	Exp.	Simul.	Relative Error	Exp.	Simul.	Relative Error	Exp.	Simul.	Relative Error
24009	2.57	47.7	43.6	-8.6%	49.1	43.8	-10.8%	2.5	2.3	-8.0%	9.9	10.8	9.1%
	4.25	77.5	72.2	-6.8%	78.6	72.4	-7.9%	3.5	3.8	8.6%	15.8	17.9	7.8%
	5.92	105.4	100.5	-4.6%	102.6	100.9	-1.7%	5.1	5.3	3.9%	21.9	24.8	8.3%
24010	2.52	70.1	63.2	-9.8%	71.4	63.1	-11.6%	2.1	2	-4.7%	25.2	24.5	-2.8%
	4.19	109.8	104.8	-4.6%	116.2	104.9	-9.7%	3.8	3.4	-10.5%	39.6	40.8	3.0%
	5.84	148.5	146.1	-1.6%	154	146.2	-5.1%	5.3	4.7	-11.3%	54.4	56.9	4.6%

B. Distribution of electromagnetic field

In the magnesium electrolysis cell, current fed from up of the anodes, and flow out from the profile, which form a coil with a quarter of a turn. The magnetic flux density is plotted in Fig. 2. The max of the magnetic flux density is about 270 Gauss (27 mTesla) approximately at the position between the anode and cathode, where is the centre of the "electromagnetic coil". The magnetic flux density of the magnesium electrolysis cell and aluminum reduction cell are at the same order of magnitude [12].
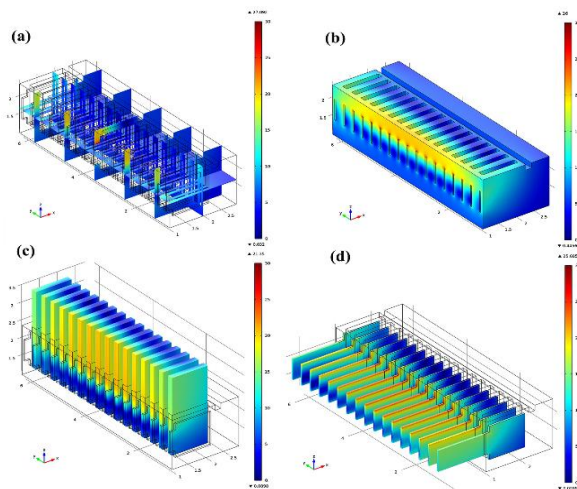
Fig. 2. Magnetic flux density in the magnesium electrolysis cell: (a) slide of electrolyte, (b) electrolyte, (c) anodes, and (d) cathode.

As shown in the Fig. 3, the magnetic flux density in the magnesium electrolysis cell are mainly focused on region between the anode and cathode. The vectors distribution of the magnetic flux density is like a one-fourth of electromagnetic coil.
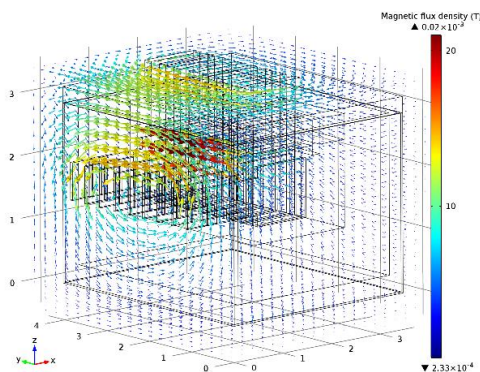

Fig. 3. Vectors of magnetic flux density in the magnesium electrolysis cell.

C. Lorentz force

Lorentz force is an important motive force of the motion of electrolyte in the magnesium electrolysis

processing. In Fig. 4, the typical contour and vector plots of the Lorentz force are plotted in the cell. These results show a high Lorentz force appears between the electrodes in the electrolyte, which is because the higher current density and higher electrolyte velocity. The maximum Lorentz force reaching 93.1 N m^{-3} at corner. And the Lorentz force become lesser toward the collection from electrolysis compartment. This will result in a significant velocity gradient in electrolyte, and may have an influence on the overall flow pattern of the electrolyte in the cell. These results show a high Lorentz force between the electrodes, which is because of the higher current in the region.
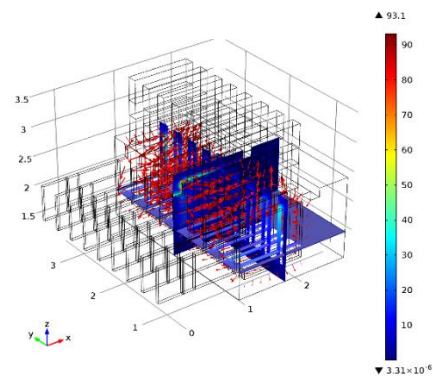

Fig. 4. Lorenz force vectors distribution in the 120 kA magnesium electrolysis cell.

Figure 5 shows the velocity vector and contour plots of the velocity magnitude in the electrolyte of the cell under the effect of the electromagnetic force. The calculated maximum magnitude of the velocity are 0.13 m s^{-1}, at the regions between the electrodes. Velocity decreases toward the center of the cell as the magnetic flux density decreases. The direction of the flow patterns is clearly a function of Lorentz force. Because the Lorentz force is the only volume force in the electrolyte, the motion of the electrolyte in the magnesium electrolysis cell follows the direction of the Lorentz force.
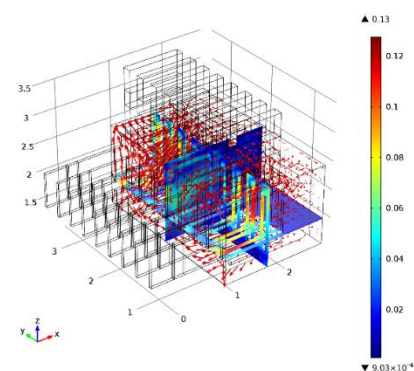

Fig. 5. Velocity vectors distribution in the 120 kA magnesium electrolysis cell.

Figure 6 shows the velocity vector and contour plots of velocity magnitude in the center of the cell (y = 2.05 m). The results show a large vortice with clockwise rotation in the cell. It is note that, the directions of Lorentz forces will help improve the circulation of electrolyte and increase the electrolysis efficiency.
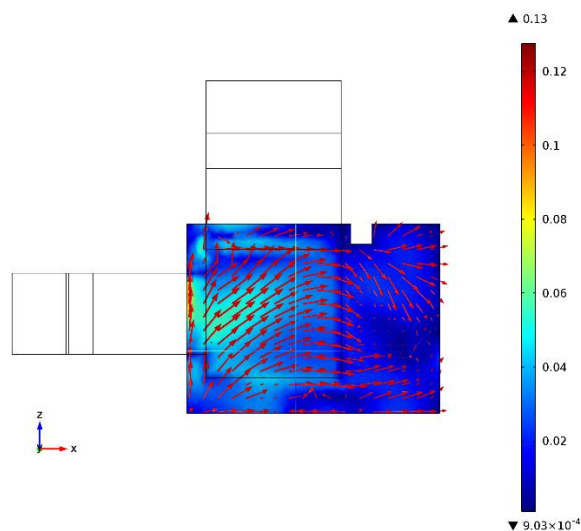

Fig. 6. Velocity vectors and contour plots of velocity magnitude in the center of the cell.

IV. CONCLUSION

A three-dimensional mathematical model of 120 kA magnesium electrolysis cell has been developed using the finite elements method. The electromagnetic field was computed, and its accuracy was validated by an electromagnetic coil. The Lorentz force acting in electrolyte was calculated based on a DC current passing through the cell and the induced magnetic field. On the basis of the electromagnetic field and flow field, the Lorentz force phase was coupled into the electromagnetic field and flow field. The model predicts that the magnitude of Lorentz force is at its maximum near the region between the anode and cathode. The direction of the Lorentz force is beneficial to the motion of the electrolyte in the magnesium electrolysis cell.

ACKNOWLEDGMENT

REFERENCES

[1] A. L. Perron, L. I. Kiss, and S. Poncsák, "Mathematical model to evaluate the ohmic resistance caused by the presence of a large number of bubbles in Hall-Héroult cells," *J. Appl. Electrochem.*, vol. 37, pp. 303-310, 2007.

[2] H. Eklund, P. B. Engseth, B. Langseth, T. Mellerud, and O. Wallevik, *An Improved Process for the Production of Magnesium*. John Wiley & Sons, Inc. 2014.

[3] K. Zhang, Y. Feng, P. Schwarz, Z. Wang, and M. Cooksey, "Computational fluid dynamics (CFD) modeling of bubble dynamics in the aluminum smelting process," *Ind. Eng. Chem. Res.*, vol. 52, pp. 11378-11390, 2013.

[4] Z. Zhang, X. Lu, T. Wang, Y. Yan, and S. Chen, "Synthesis and electrolysis of $K_3NaMgCl_6$," *Ind. Eng. Chem. Res.*, vol. 54, pp. 1433-1438, 2015.

[5] A. Rezvanpour, E. W. C. Lim, and C. Wang, "Computational and experimental studies of electrohydrodynamic atomization for pharmaceutical particle fabrication," *AIChE J.*, vol. 58, pp. 3329-3340, 2012.

[6] M. Dupuis, "Thermo-electric design of a 740 kA cell, is there a size limit?," *Aluminium: Inter. J. Ind., Res. App.*, vol. 81, pp. 324-327, 2005.

[7] D. S. Severo, A. F. Schneider, E. C. Pinto, B. Gusberti, and V. Potocnik, "Modeling magneto-hydrodynamics of aluminum electrolysis cells with ANSYS and CFX," *Light Met.*, pp. 475-480, 2005.

[8] E. I. Shilova and E. V. Shcherbinin, "Control methods MHD convection in the magnesium electrolysis cells," *Magnetohydrodyn.*, vol. 34, pp. 147-152, 1998.

[9] Z. Sun, H. Zhang, P. Li, B. Li, G. Lu, and J. Yu, "Modeling and simulation of the flow field in the electrolysis of magnesium," *JOM*, vol. 61, pp. 29-33, 2009.

[10] Z. Sun, Z. Yun, G. Lu, P. Li, J. Wang, and J. Yu, "Novel method based on electric field simulation and optimization for designing an energy-saving magnesium electrolysis cell," *Ind. Eng. Chem. Res.*, vol. 50, pp. 6161-6173, 2011.

[11] C. Liu, Z. Sun, G. Lu, X. Song, and J. Yu, "Scale-up design of a 300 kA magnesium electrolysis cell based on thermo-electric mathematical models," *Can. J. Chem. Eng.*, vol. 92, pp. 1197-1206, 2014.

[12] S. Das, G. Brooks, and Y. Morsi, "Theoretical investigation of the inclined sidewall design on magnetohydrodynamic (MHD) forces in an aluminum electrolytic cell," *Metall. Mater. Trans. B.*, vol. 42, pp. 243-253. 2011.

Metamaterial-Inspired Split Ring Monopole Antenna for WLAN Applications

S. Imaculate Rosaline and S. Raghavan

Department of Electronics and Communication Engineering
National Institute of Technology, Tiruchirappalli, India
imaculaterosaline@gmail.com, raghavan@nitt.edu

Abstract — This paper describes the design of a compact dual band monopole antenna for WLAN (2.4/5.2/5.8 GHz) applications. The antenna is printed on a $22.5 \times 24 \times 0.8$ mm^3 FR-4 substrate with a partial ground plane and is fed by a microstrip line. The proposed structure consists of a simple hexagonal ring with a split arm along its center. The split in the arm in turn creates a quarter wavelength resonance in the higher frequency range. It also induces magnetic resonance which accounts for band notch between the WLAN lower (2.4 GHz) and upper bands (5.2/5.8 GHz). The extraction of negative permeability of the split ring structure is also discussed. A prototype of the proposed structure is fabricated and the measured results comply greatly with the simulated results. The antenna has consistent radiation pattern and stable gain over all the working region.

Index Terms — Hexagonal monopole antenna, negative permeability, notch frequency, split ring, WLAN.

I. INTRODUCTION

Wireless Local Area Network (WLAN) is a significant component of the wireless computer network which interconnects two or more devices. It is based on IEEE 802.11 standard and operates in the 2.45 (2.4-2.48) GHz, 5.2 (5.15-5.35) GHz and 5.8 (5.75-5.825) GHz frequencies. Design of single antenna capable of operating at all these specified frequencies have attracted many researchers in the recent past. Besides obtaining multiple frequencies, the antenna also demands compactness, cost effectiveness and flexibility to be integrated with other microwave integrated devices. Printed monopole antennas seem to be a good choice to meet these aforementioned challenges. Multi branched radiators [1, 2], slotted monopoles [3-5], meander monopoles [6], fractal shapes [7] are few among them to obtain dual band operation in the WLAN 2.5/5.2/5.8 GHz range. However, these antennas suffer from either complicated geometry [1, 5] or larger dimensions. Reactive slots in the radiating patch [3] have compact dimensions, yet it resulted in poor impedance matching at the lower resonant band. Recently, electromagnetic (EM) metamaterials inspired split ring elements and its complementary are also used as radiating structures for achieving compact and dual band antennas in the WLAN range. Their role in antenna design becomes attractive because of their ability to achieve miniaturization [8, 9], multiband resonances [10] gain and bandwidth enhancement [11]. Split ring monopole antenna proposed in [12] has impedance matching problem in the lower WLAN band, whereas the dual band antennas with CSRRs [13] and triangular split ring resonators (SRRs) [14] has larger dimensions. In general, the overall dimension of these antennas are large compared with the proposed one as shown in Table 1 below. Also, unlike these antenna analysis, this paper emphasizes on the role of metamaterial property (negative permeability) in antenna design, which many papers have failed to prove. As a result, the antenna designer can enjoy the privilege of tuning the operating frequency to the desired range.

In this paper, a simple and compact hexagonal split ring radiating element is proposed for WLAN applications. The split in the ring element is capable of creating band separation (notch band) between the operating bands due to its induced magnetic resonance. The proposed geometry is very simple with good resonant and radiation characteristics, making it a good choice for commercial use.

Table 1: Comparison of the existing antennas with the proposed antenna

Ref.	Dimensions, L x W (mm^2)	Metamaterial Property Verification
[12]	20 x 32	Not verified
[13]	34 x 30	Not verified
[14]	40 x 35	Not verified
Proposed antenna	24 x 22.5	Verified

II. PROPOSED ANTENNA DESIGN

The evolution of the proposed split ring radiating antenna is shown in Fig. 1. Configuration A shows a hexagonal ring monopole fed by a 50 Ω microstrip line and a partial ground plane. The monopole considered in our antenna design is hexagonal in shape whose resonant frequency will be similar to that of a circular monopole

[15], hence the resonant frequency is given as $fr \approx \frac{1.8412*c}{4\pi S\sqrt{\varepsilon_r}}$. Here, c is the velocity of light, S is the side length of the hexagon and ε_r is the dielectric constant of the substrate. Hence, for a side length of 9 mm, the antenna resonates at 2.4 GHz. Now, in configuration B, a vertical arm is introduced at the center which in turn is connected to the feed directly. This vertical arm, opens up the higher order resonance. Finally, in configuration C, a split is introduced at the center of the vertical arm, which in turn opens up the quarter wave resonance corresponding to length L1 and also, induces a narrow magnetic resonance due to its capacitive effect and creates a sharp notch corresponding to the split width, yielding two resonant bands centered at 2.4 GHz and at 6 GHz. A detailed layout of the proposed antenna is shown in Fig. 2 along with its side view and its dimensions are listed in Table 2. Photograph of the proposed structure is shown in Fig. 3.
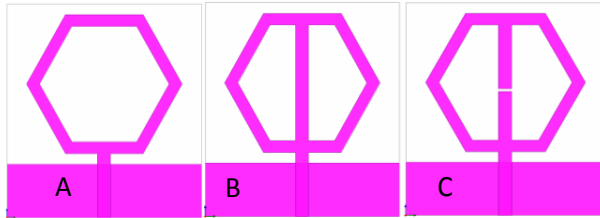

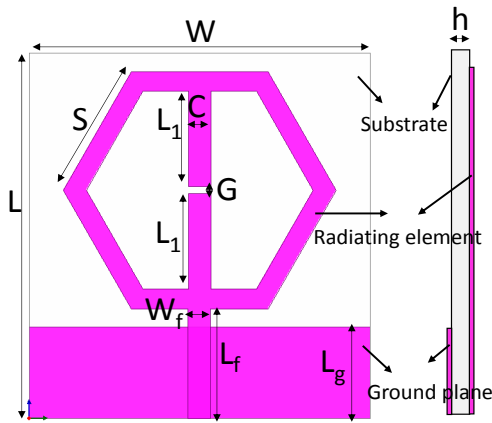Fig. 1. Evolution of the proposed antenna.

Fig. 2. Geometry of the proposed antenna: (a) top view and (b) side view.

Table 2: Dimensions of the proposed antenna

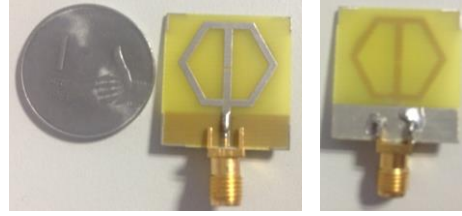| Parameter	Dimension (mm)	Parameter	Dimension (mm)
L	24	W	22.5
S	9	C	1.5
G	0.3	L_1	6.3
L_f	7.2	W_f	1.5
L_g	6	h	0.8

Fig. 3. Photograph of the fabricated dual band antenna (top view and bottom view).

III. SIMULATION RESULTS

Simulations are performed using the Ansoft High Frequency Structure Simulator (HFSS) V.15.0 commercial software package. Figure 4 shows the simulated return loss characteristics of three configurations shown in Fig. 1. Configuration A shows resonance around 2.4 GHz. When the vertical arm is introduced (configuration B), a higher order resonance is noted. The width of the vertical arm plays an important role in determining the higher order resonance. Finally, in configuration C, a split is introduced at the center of the vertical arm to induce magnetic resonance. Now, the higher order resonance is opened from 4 GHz to 7.5 GHz, covering the upper WLAN frequencies (5.15–5.35) and (5.75–5.825). Figure 5 shows the parametric study on the return loss characteristics of configuration C for various vertical arm's width C, ranging from 1.5 mm to 6 mm in steps of 1.5 mm. It is inferred that, as the width C increases, the notch frequency is shifted towards the lower frequencies, opening the upper WLAN band. The lower frequency limit of this band (5.2/5.8 GHz) is determined by the dimension L1 × C. For L_1 = 6.3 mm and C = 1.5 mm, the lower frequency limit is the quarter wave resonance of length (6.3 mm + 1.5 mm). Thus, C = 1.5 mm is chosen to be optimum for our design, which corresponds to the notch around 4 GHz. It is also inferred that, for C = 6 mm, the WiMAX band (3.5 GHz) is also covered.
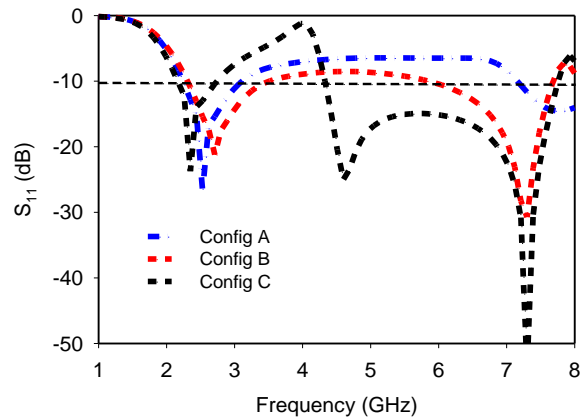

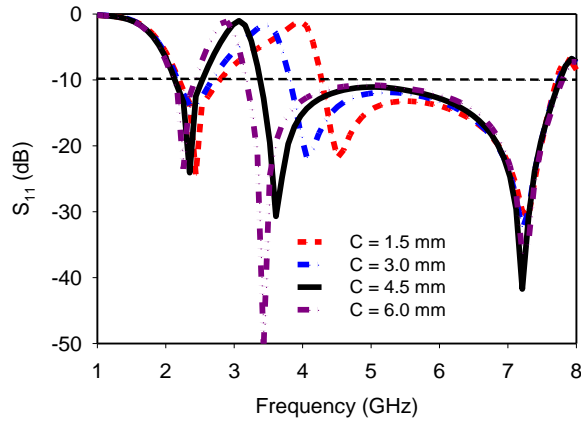Fig. 4. Simulated return loss characteristics of the three configurations A, B, and C.

Fig. 5. Simulated return loss characteristics of configuration C for various width G.

Similarly, the split gap G plays an important role in determining the notch frequency. Figure 6 shows the parametric study on the return loss characteristics of configuration C for various split width G ranging from 0.3 mm to 0.9 mm in steps of 0.2 mm. It is observed that, as the split gap G increases, the notch frequency also increases correspondingly. Hence, G = 0.3 mm is chosen for our design, for a notch to occur around 4 GHz.
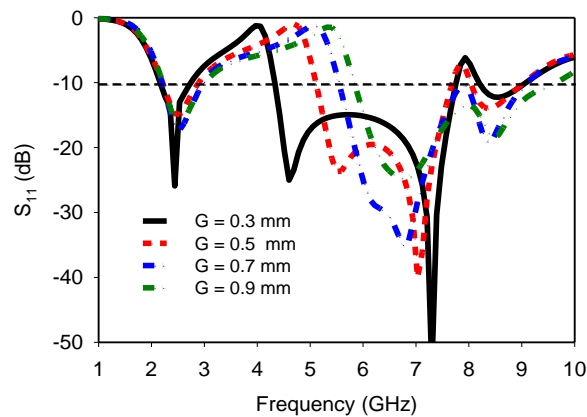

Fig. 6. Simulated return loss characteristics of configuration C for various width C.

IV. SPLIT RING ANALYSIS

The radiating element is itself a split ring structure which is analyzed using the classic waveguide theory approach. The transmission and reflection coefficients are noted and from which the effective material parameters, permeability and permittivity are extracted. Figure 7 shows the real parts of extracted effective permeability values plotted along with the return loss characteristics of the proposed structure. It is inferred that the permeability is negative around 4 GHz. This

negative permeability region has in turn led to the notch frequency, which can be clearly understood by the dashed grey region. Due to the negative permeability, no transmission is practical in this region, thus the S_{11} curve exhibits notch band over this region.
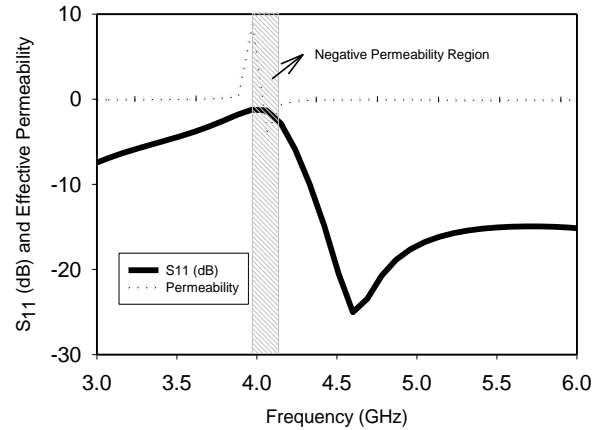

Fig. 7. Comparison of extracted real parts of effective permeability and S_{11} (dB).

V. MEASUREMENT RESULTS

The return loss characteristics are measured using a vector network analyzer. Figure 8 shows the simulated and measured return loss results. The measured data shows dual band resonance centered at 2.4 GHz (2.0 – 2.7 GHz) and at 4.4 GHz and 7.12 GHz (4.12 – 7.66 GHz). The measured data greatly agree with the simulated results. The radiation pattern of the proposed antenna is measured in an anechoic chamber, which is shown in Fig. 9. A consistent omnidirectional pattern is observed in the H plane and a bidirectional pattern is observed in the E plane over all the operating region (2.4, 5.5 GHz).
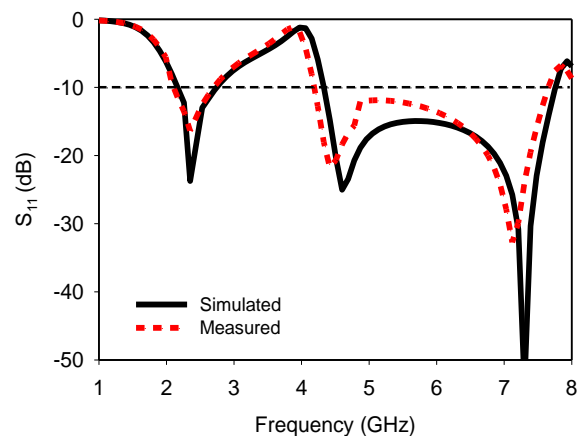

Fig. 8. Simulated and measured return loss characteristics of the proposed antenna.
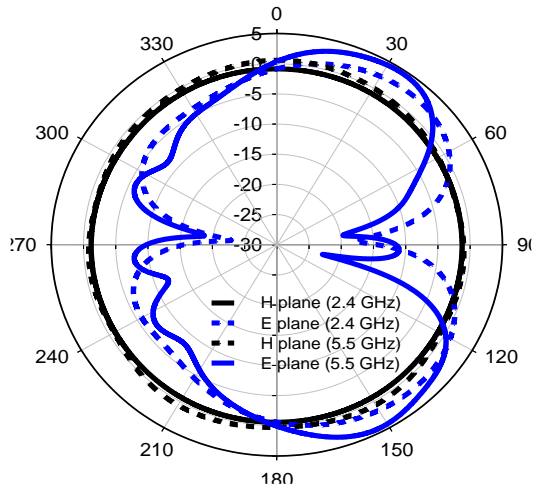
Fig. 9. Measured H plane and E plane pattern of the proposed antenna at 2.4 GHz and 5.5 GHz.

VI. CONCLUSION

A dual band monopole antenna suitable for WLAN 2.4/5.5 GHz applications is presented in this paper. The antenna makes use of a metamaterial inspired split ring structure for achieving the dual band resonance. The antenna geometry is very simple and also compact making mass production easy. The radiation pattern and gain are consistent over all the operating bands making the proposed antenna a good choice for wireless applications.

REFERENCES

[1] Y. Xu, Y.-C. Jiao, and Y.-C. Luan, "Compact CPW-fed printed monopole antenna with triple band characteristics for WLAN/WiMAX applications," *Electron Lett*., vol. 48, pp. 1519-1520, 2012.

[2] X. L. Sun, L. Liu, S. W. Cheung, and T. I. Yuk, "Dual-band antenna with compact radiator for 2.4/5.2/5.8 GHz WLAN applications," *IEEE Trans. Ant. P*, vol. 60, pp. 5924-5931, 2012.

[3] W.-C. Liua, C.-M. Wua, and N.-C. Chu, "A compact low profile dual-band antenna for WLAN and WAVE applications," *AEU Int. J. Electron. C*, vol. 66, pp. 467-471, 2012.

[4] C.-Y. Huang and E.-Z. Yu, "A slot monopole antenna for dual band WLAN applications," *IEEE Antennas Wireless Propag. Lett.*, vol. 10, pp. 500-502, 2011.

[5] X.-Q. Zhang, Y.-C. Jiao, and W.-H. Wang, "Compact wide tri-band slot antenna for WLAN/WiMAX applications," *Electron. Lett*., vol. 48, pp. 64-65, 2012.

[6] H.-Y. Chien, C.-Y.-D. Sim, and C.-H. Lee, "Dual band meander monopole antenna for WLAN operation in laptop computer," *IEEE Antennas Wireless Propag. Lett*., vol. 12, pp. 694-697, 2013.

[7] R. Ghatak, R. K. Mishra, and D. R. Poddar, "Perturbed Sierpinski carpet antenna with CPW feed for IEEE 802.11 a/b WLAN application," *IEEE Antennas Wireless Propag. Lett*., vol. 7, pp. 742-744, 2008.

[8] D. Laila, R. Sujith, V. A. Shameena, C. M. Nijas, V. P. Sarin, and P. Mohanan, "Complementary split ring resonator-based microstrip antenna for compact wireless applications," *Microw. Opt. Technol. Lett*., vol. 55, pp. 814-816, 2013.

[9] R. O. Ouedraogo, E. J. Rothwell, A. R. Diaz, K. Fuchi, and A. Temme, "Miniaturization of patch antennas using a metamaterial-inspired technique," *IEEE Trans. Ant. P*, vol. 60, pp. 2175-2182, 2012.

[10] Y. H. Xie, C. Zhu, L. Li, and C. H. Liang, "A novel dual-band metamaterial antenna based on complementary split ring resonators," *Microw. Opt. Technol. Lett*., vol. 54, pp. 1007-1009, 2012.

[11] L.-W. Li, Y.-N. Li, T. S. Yeo, J. R. Mosig, and O. J. F. Martin, "A broadband and high-gain metamaterial microstrip antenna," *Appl. Phys. Lett*., vol. 96, pp. 164101-164103, 2010.

[12] S. C. Basaran and Y. E. Erdemli, "A dual band split ring monopole antenna for WLAN applications," *Microw. Opt. Technol. Lett*., vol. 51, pp. 2685-2688, 2009.

[13] S. C. Basaran, U. Olgun, and K. Sertel, "Multiband monopole antenna with complementary split ring resonators for WLAN and WiMAX applications," *Electron. Lett*., vol. 49, pp. 636-638, 2013.

[14] K. Yang, H. Wang, Z. Lei, Y. Xie, and H. Lai, "CPW-fed slot antenna with triangular SRR terminated feed line for WLAN/WiMAX applications," *Electron. Lett*., vol. 47, pp. 685-686, 2011.

[15] C. A. Balanis, *Antenna Theory: Analysis and Design*. John Wiley & Sons, 2005.

The Equivalent Circuit Extraction and Application for Arbitrary Shape Graphene Sheet

Ying S. Cao [1], Li Jun Jiang [1], and Albert E. Ruehli [2]

[1] Department of Electrical and Electronic Engineering
The University of Hong Kong, Pokfulam Road, Hong Kong
caoying@eee.hku.hk, jianglj@hku.hk

[2] UMRI/MST EMC Laboratory
Missouri University of Science and Technology, Rolla, MO 65409, USA
albert.ruehli@gmail.com

Abstract ─ In this work, for the first time the electromagnetic features of graphene are characterized by a circuit model derived instead of fitted from the electric field integral equation (EFIE). The atomically thick graphene is equivalently replaced by an impedance surface. When it is magnetized, the impedance surface is anisotropic with a tensor conductivity. Based on EFIE, the graphene's circuit model can be derived by the partial element equivalency circuit (PEEC) concept. The anisotropic resistivity is modeled using a serial resistor with current control voltage sources (CCVSs). From the derived circuit model, electromagnetic properties of graphene can be conveniently analyzed. This work also provides a new characterization method for dispersive and anisotropic materials.

Index Terms ─ Graphene, magnetized, non-magnetized, PEEC.

I. INTRODUCTION

Graphene is an atomically thin nanomaterial with promising application potentials. However, its electromagnetic modeling is usually totally numerical [1] or physically empirical. This fact motivated us, in this paper, to derive an equivalent circuit model based on EFIE for both non-magnetized and magnetized graphene. It employs the PEEC [2-7] process to convert the electromagnetic interactions on the graphene surface into resistive, inductive and capacitive effects on an isotropic or anisotropic impedance surface. By solving the derived circuit model, graphene's electromagnetic properties can be fully predicted conveniently and efficiently.

In this paper, a novel circuit model based on the electric field integral equation (EFIE) is proposed to solve the dispersivity of non-magnetized and magnetized graphene. For the non-magnetized graphene, the conductivity is composed of intraband and interband contributions. The resistive part in the equivalent circuit model is modeled as a resistor, an inductor and they are in series with Zinter which accounts for the interband contribution of the surface conductivity of graphene. In the equivalent circuit model for magnetized graphene, the diagonal elements of the surface conductivity tensor intrinsically correspond to the resistance of each inductive branch, which is the same as the unbiased scalar conductivity of graphene. For the off-diagonal elements of the conductivity tensor, a new equivalent circuit model is developed to model the resistive characteristics by utilizing current-controlled voltage sources (CCVSs).

The advantages of the proposed new method are: (i) The model is derived based on EM wave equations. It is not empirical or curve fitted. Hence, it is more reliable and general. Based on our search, this is the first derived model for graphene. (ii) The new method is much more efficient than the volumetric based graphene modeling process. (iii) Compared with the numerical process [8], the derived circuit model of graphene provides a convenient bridge to integrate graphene EM parasitic effects with lumped circuit designs.

II. EQUIVALENT CIRCUIT MODEL DERIVATION FOR GRAPHENE

The 2D atomically thin graphene can be considered as an impedance surface with the dispersive conductivity that is isotropic or anisotropic along tangential directions [9]. Based on the electric field integral equation, we have:

$$\mathbf{E}^{inc}(\mathbf{r},t) = \frac{\mathbf{J}(\mathbf{r},t)}{\sigma} + \mu \int_{v'} G(\mathbf{r},\mathbf{r}') \frac{\partial \mathbf{J}(\mathbf{r}',t)}{\partial t} dv' + \frac{\nabla}{\varepsilon} \int_{v'} G(\mathbf{r},\mathbf{r}') q(\mathbf{r}',t) dv', \tag{1}$$

where $G(\mathbf{r},\mathbf{r}')$ is the full wave Green's function. For one current filament on the graphene sheet, using the partial equivalence element concept, Eq. (1) becomes Kirchhoff's

Voltage Law:

$$V = RI + Lp\frac{dI}{dt} + Q \cdot Pp,$$ (2)

where Lp is the partial inductance, and Pp is the partial coefficient of potential. The Lp between cell α and β and Pp between cell i and j are:

$$Lp_{\alpha\beta} = \frac{\mu}{a_\alpha a_\beta}\int_{v_\alpha}\int_{v_\beta}G(\mathbf{r}_\alpha,\mathbf{r}_\beta)dv_\alpha dv_\beta,$$ (3)

$$Pp_{ij} = \frac{1}{\varepsilon S_i S_j}\int_{S_i}\int_{S_j}G(\mathbf{r}_i,\mathbf{r}_j)dS_i dS_j.$$ (4)

The 2-dimensional meshing scheme can be represented as Fig. 1 [2].

Discretization

x-directional Inductive Cells   y-directional Inductive Cells   Capacitive Cells
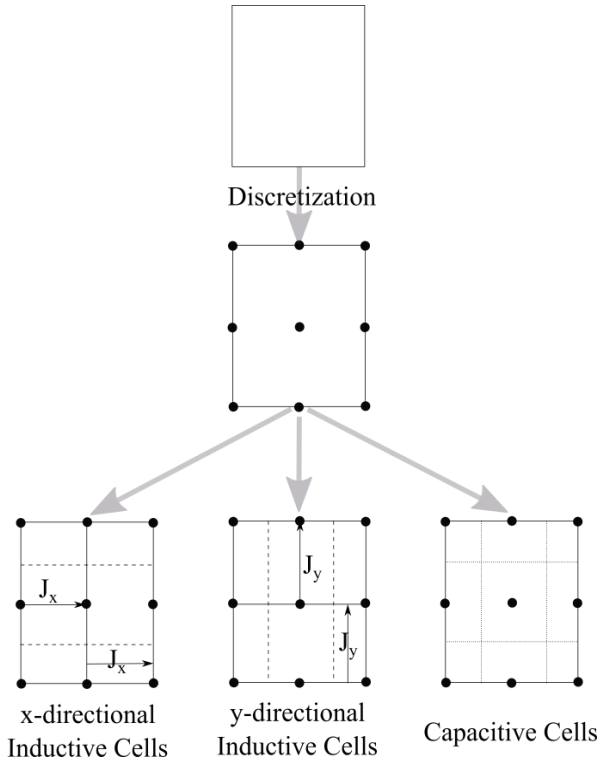
Fig. 1. This is the 2D discretization of thin conductive plate. Dark circles indicate nodes, dashed lines separate inductive cells, and dotted lines separate capacitive cells.

A. Non-magnetized graphene

The graphene dispersive conductivity σ is a summation of both intraband and interband contributions [10]. For a one freestanding rectangular patch with the length l and width w, its surface resistance can be derived from Kubo's formula:

$$R_{surf} = R_r + j\omega L_r + Z_{inter},$$ (5)

where R_r is from the real part of σ, L_r is from the imaginary part of σ, and Z_{inter} is from the inter band part.

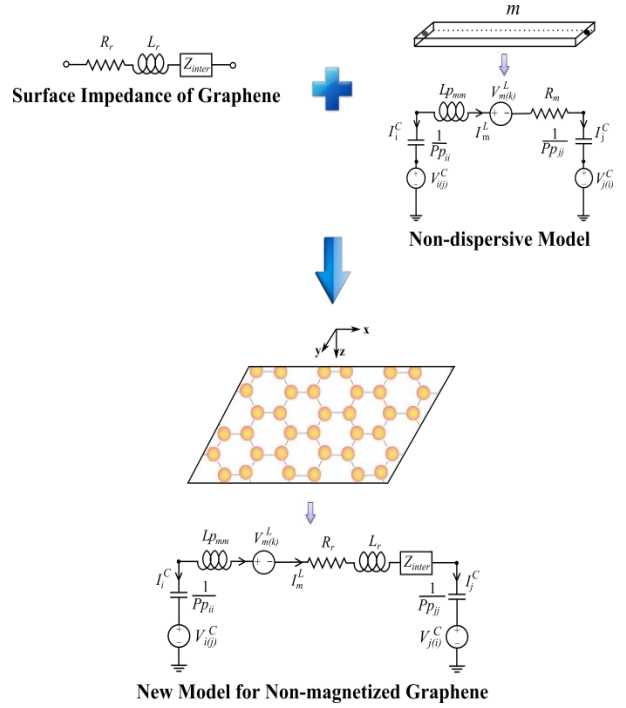Combining other parts of the cell model, the new non-magnetized graphene unit model is illustrated in Fig. 2.

Surface Impedance of Graphene

Non-dispersive Model

New Model for Non-magnetized Graphene

Fig. 2. One cell's model for a non-magnetized graphene patch. The left bottom model is the traditional equivalent model, where Lp_{mm}, Pp_{ii} and Pp_{jj} are self-inductance and self-coefficients of potential, respectively. $V_{m(k)}^L$ is the voltage control voltage source (VCVS) corresponding to mutual inductive couplings. $V_{i(j)}^C$ and $V_{j(i)}^C$ are the voltage control voltage sources (VCVSs) due to mutual capacitive coupling between two capacitors i and j.

B. Magnetized graphene

With the bias by a static magnetic field, the surface conductivity of graphene becomes an anisotropic tensor, which complicates the problem. Hence, the numerical methods have to settle the dispersive and anisotropic properties of graphene simultaneously.

For the magnetized graphene, its surface conductivity becomes an anisotropic and dispersive tensor $\bar{\bar{\sigma}}$ [9]. Hence, the electric field has contributions from orthogonal current components. For example, the x-direction electric field is a function of J_x and J_y. Hence, for off-diagonal elements of the conductivity tensor, a current control voltage source (CCVS) can be used to represent each of these orthogonal contributions. CCVSs are in series with the intrinsic resistances that are derived from diagonal terms of $\bar{\bar{\sigma}}$. In Fig. 3, the new equivalent circuit model for the magnetized graphene is sketched.

$$\mathbf{E}^{tot} = \mathbf{E}^{inc} - \frac{\partial \mathbf{A}}{\partial t} - \nabla\phi = \overset{=-1}{\sigma}\mathbf{J} = \overset{=}{\rho}\mathbf{J}$$

$$= (\rho_{xx}J_x + \rho_{xy}J_y)\hat{x} + (\rho_{yx}J_x + \rho_{yy}J_y)\hat{y}$$
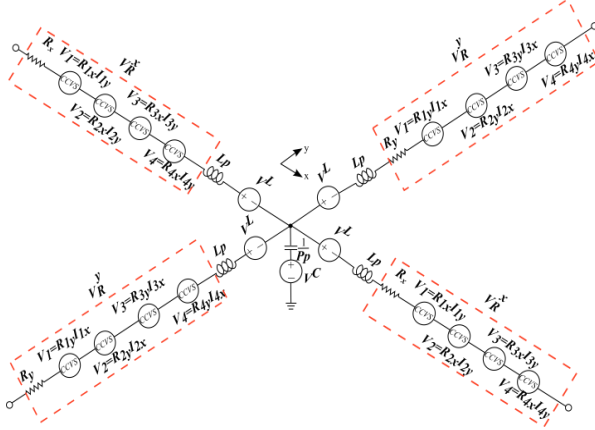

Fig. 3. A schematic diagram of the complete equivalent circuit for magnetized anisotropic conductivity graphene. This circuit model is for four nearby cells which share a common node, and two for x–directional cells and the other two for y–directional cells. The scripts for each cell are omitted for simplicity.

III. NUMERICAL RESULTS

A. Non-magnetized graphene

The $5 \times 0.5 \ \mu m^2$ graphene patch is illuminated by a plane wave linearly polarized along the patch length. The direction of propagation is normal to the surface of graphene. The absorption cross section of the graphene patch is shown in Fig. 4 for different relaxation times. By comparing with results from [10] (represented by circles), it is seen that perfect agreements are achieved including the positions of important resonant frequencies.
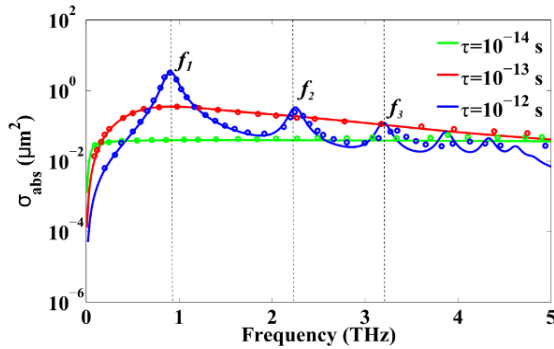

Fig. 4. Absorption cross section (in logarithmic scale) of a non-magnetized graphene patch as a function of frequency, for different relaxation time: 10^{-14} s (green line), 10^{-13} s (red line), 10^{-12} s (blue line). The results are compared with [5] plotted using circles. f_1, f_2 and f_3 are resonant frequencies.

B. Magnetized graphene

To validate the accuracy of the proposed algorithm for the magnetized graphene, a $10 \times 2 \ \mu m^2$ graphene patch is studied first. Relaxation time $t = 1.3 \times 10^{-13}$ s. The magnetic bias $B_0 = 0.25$ T. The graphene patch is biased by a z–directional static magnetic field and the same excitation plane wave in III.A is used. The absorption cross section and extinction cross section calculated by new method and discontinuous Galerkin method [9] are compared in Fig. 5. The definitions of the absorption cross section and extinction cross section can be found in [6].
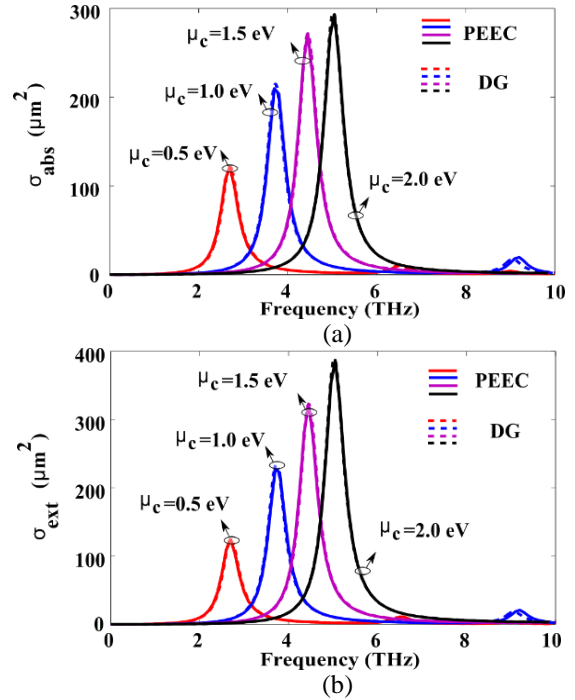

Fig. 5. (a) σ_{abs} (absorption cross section) and (b) σ_{ext} (extinction cross section) of the magnetized graphene patch.

IV. CONCLUSION

In this paper, a novel equivalent circuit model is derived for the general graphene sheet based on EM integral equations. It provides a new bridge between EM parasitic effects and lumped circuit designs for researches on graphene and other dispersive anisotropic media.

REFERENCES

[1] V. Nayyeri, M. Soleimani, and M. Ramahi, "Wideband modeling of graphene using the finite-difference time-domain metod," *IEEE Trans. Antenna Propag.*, vol. 6, no. 12, pp. 6107-6114, Dec. 2013.

[2] A. E. Ruehli, "Equivalent circuit models for three dimensional multiconductor systems," *IEEE Trans. Microw. Theory Tech.*, vol. MTT-22, no. 3, pp. 216-221, Mar. 1974.

[3] A. E. Ruehli, "Inductance calculations in a complex integrated circuit environment," *IBM J. Res. Develop.*, vol. 16, no. 5, pp. 470-481, Sep. 1972.

[4] A. E. Ruehli and P. A. Brennan, "Efficient capacitance calculations for three-dimensional multiconductor systems," *IEEE Trans. Microw. Theory Tech.*, vol. 21, no. 2, pp. 76-82, Feb. 1973.

[5] Y. S. Cao, L. Jiang, and A. E. Ruehli, "Distributive radiation and transfer characterization based on the PEEC method," *IEEE Trans. Electromag. Compat.*, vol. 57, no. 4, pp. 734-742, Aug. 2015.

[6] Y. S. Cao, L. Jiang, and A. E. Ruehli, "An equivalent circuit model for graphene-based terahertz antenna using the PEEC method," *IEEE Trans. Antenna Propag.*, vol. 64, no. 4, pp. 1385-1393, Apr. 2016.

[7] Y. S. Cao, L. Jiang, and A. E. Ruehli, "The derived equivalent circuit model for magnetized anisotropic grapheme," submitted to *IEEE Trans. Antenna Propag.*

[8] O. V. Shapoval, J. S. G.-Diaz, J. P.-Carrier, J. R. Mosig, and A. I. Nosich, "Integral equation analysis of plane wave scattering by coplanar graphenestrip gratings in the THz range," *IEEE Trans. Terahertz Sci. Techn.*, vol. 6, no. 3, pp. 666-674, Sept. 2013.

[9] P. Li and L. J. Jiang, "Modeling of magnetized graphene from microwave to THz range by DGTD with a scalar RBC and an ADE," *IEEE Trans. Antennas Propag.*, vol. 63, no. 10, Oct. 2015.

[10] I. Llatser, C. Kremers, D. N. Chigrin, J. M. Jornet, M. C. Lemme, A. Cabellos-Aparicio, and E. Alarcon, "Radiation characteristics of tunable graphene as in the terahertz band," Radioengineering, vol. 21, no. 4, pp. 946-953, 2012.

A Subwavelength Perfect Absorbing Metamaterial Patch Array Coupled with a Molecular Resonance

Michael F. Finch and Brian A. Lail

Department of Electrical and Computer Engineering
Florida Institute of Technology, 150 West University Blvd., Melbourne, FL, 32901, USA
mfinch2009@my.fit.edu, blail@fit.edu

Abstract ─ A perfectly absorbing metamaterial (PAMM) coupled with vibrational modes has varied applications ranging from surface-enhanced vibrational spectroscopy to biological sensing. This endeavor considers a subwavelength PAMM sensor design and analysis using a commercially available finite element method (FEM) solver and analytically with temporal coupled mode theory (TCMT). A carbon double oxygen bond (C=O) at 52 THz or 1733 cm^{-1} that resides in poly(methyl methacrylate), PMMA, will be used as a stand-in analyte. Normal mode splitting that results from the resonant coupling between the PAMM and analyte's molecular resonance is investigated and analyzed.

Index Terms ─ Electromagnetic Induced Absorption (EIA), Electromagnetic Induced Transparency (EIT), metamaterial, perfect absorbing, PMMA, resonant coupling, superscattering.

I. INTRODUCTION

Metamaterial are engineered materials that are designed with periodic or aperiodic elements known as meta-atoms [1]. A perfect absorbing metamaterial (PAMM), as the name implies, is designed to "perfectly" absorb incident fields, and have been described with a metamaterial impedance matched to free space [2-5]. A commercially available finite element method solver, Ansys high frequency structural simulator (HFSS), is employed to design and simulate a gold patch array metasurface with ground plane to form a subwavelength optical resonant cavity.

Resonant coupling results in avoidance crossing dispersion relationships, or normal mode splitting, that is described with temporal coupled mode theory (TCMT) [2, 6-9]. Coupled resonant models have been used a classical analogy treatment of quantum phenomena, namely Fano resonance with bright/dark mode interactions [10-12], and electromagnetically induced transparency or absorption (EIT or EIA) [2, 13-17]. In a Fano resonance, EIT, or EIA treatment a molecular resonance is modeled as a dark mode. TCMT has also

been used to describe light-matter interactions in polaritonic systems [18].

In the case of this work, poly(methyl methacrylate) (PMMA), a thermal plastic for very-large-scale integration (VLSI) and material for plastic fibers, has a carbon double oxygen (C=O) molecular resonance at 52 THz/1733 cm^{-1} with optical material properties measured using ellipsometry [19]. PMMA's C=O infrared (IR) active molecular bond will be used as an analyte stand-in to demonstrate the mode splitting. Metamaterial coupled to molecular resonances has varied applications including, but not limited to, biosensing and surface enhanced-vibrational spectroscopy [4, 9, 10, 13, 20, 21].

II. PAMM SENSOR

The PAMM unit cell under consideration is a gold patch stood-off above from a gold ground plane by a spacer of amorphous silicon as shown in Fig. 1. For the preliminary analysis of the PAMM, a thin overlay of material with the refractive index that of dispersionless PMMA is introduced to reduce the red shifting induced on the PAMM's resonance [4]. The PAMM spacer thickness range is on the order of 20 to 200 nm which results in a subwavelength nanoresonator cavity.
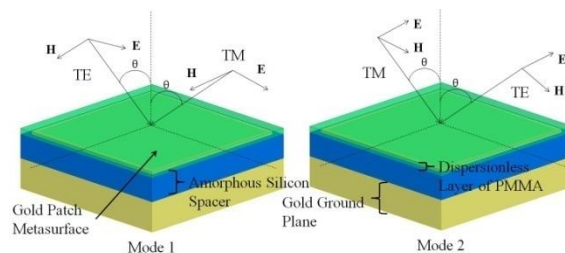

Fig. 1. PAMM unit cell modeled in Ansys HFSS.

A. Theory

Derived from circuit theory or mass spring relationships [2, 6, 21], it can be seen from the TCMT that the PAMM can be modeled as a single input uncoupled system (SI-US) as seen in Fig. 2. The TCMT

equations for normalized resonator energy "a" at resonator frequency ω_0 can be written as:

$$\frac{da}{dt} = j\omega_0 a - (\gamma_0 + \gamma_e)a + \alpha S^+, \qquad (1.a)$$

$$S^- = cS^+ + da, \qquad (1.b)$$

where [2, 6],

$$|\alpha| = \sqrt{2\gamma_e}, \qquad (2.a)$$

$$d = \sqrt{2\gamma_e}, \qquad (2.b)$$

$$c = -\left|\frac{\alpha}{d}\right| = -1, \qquad (2.c)$$

and γ_0 and γ_e represent internal and external losses, respectively. The internal loss can be thought of as Ohmic loss within the PAMM or circuit resonator. External loss results from excitation that does not couple into the PAMM or resonator, but reflects or scatters out of the structure. The excitation and reflected field is denoted as S^+ and S^- respectively.
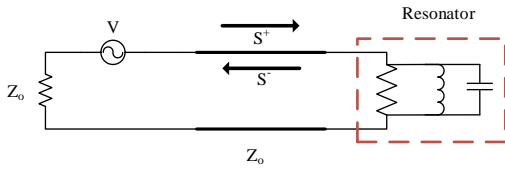

Fig. 2. Circuit visualization for a SI-US where the resonator circuit models the PAMM.

Given the single input description of the PAMM and from (1) and (2) it can be seen that the only scattering parameter is the reflection coefficient:

$$S^- = \left[\frac{j(\gamma_e - \gamma_0) + (\omega - \omega_0)}{j(\gamma_e + \gamma_0) - (\omega - \omega_0)}\right]S^+, \qquad (3)$$

and at resonance (3) becomes:

$$S_{11}(\omega_0) = \frac{\gamma_e - \gamma_0}{\gamma_e + \gamma_0}. \qquad (4)$$

Due to the ground plane in the PAMM the transmission spectrum can be taken as negligible; therefore,

$$A(\omega) = 1 - |S_{11}(\omega)|^2, \qquad (5)$$

where $A(\omega)$ is the absorbed power in the PAMM. It can be seen from (4) when $\gamma_e = \gamma_0$ then $|S_{11}(\omega_0)|$ tends to zero, or $A(\omega)$ becomes unity. The condition $\gamma e = \gamma_0$ can be described via impedance matching and results in the "perfect absorption" condition in the metamaterial known as criticallycoupled [2, 3, 6]. The condition $\gamma_e < \gamma_0$ and $\gamma_e > \gamma_0$ is referred to as undercoupled and overcoupled respectively. The "coupling" in critically-(CC), under-(UC), and overcoupled (OC) is in reference to the impinging field "coupling" into the PAMM and not related to the molecular resonance.

B. PAMM numerical results

While maintaining the PAMM at approximately at 52 THz (molecular resonant frequency), the thickness of the A-Si spacers was varied. A case of OC, CC, and UC can be seen in Fig. 3 for spacer thickness of 150, 90, and 50 nm respectively. From (4), it can be seen that the CC case is the only case where approximately the perfect absorption condition is met. Figure 4 is the spectral and angular resolution for the absorption for the CC case for both incident modes seen in Fig. 1. It can be seen that the design patch metamaterial provides a polarization insensitive to transverse electrical (TE), transverse magnetic (TM), or either mode orientation while providing a wide-field-of-view [3]. With the use of the TCMT equations, the damping rates, external (γ_e) and internal (γ_0), were determined by a parametric fit and the absorption is plotted in Fig. 3 as seen as asterisks (*), and linear fit as seen in Fig. 5.
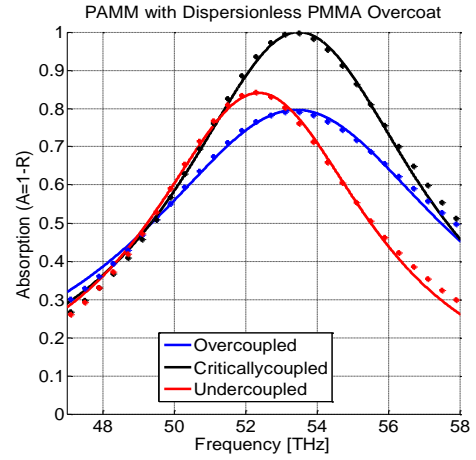

Fig. 3. Numerical (*) and mathematical model (solid line) results are shown for case of over-(OC), critically-(CC), and undercoupled (UC) cases.
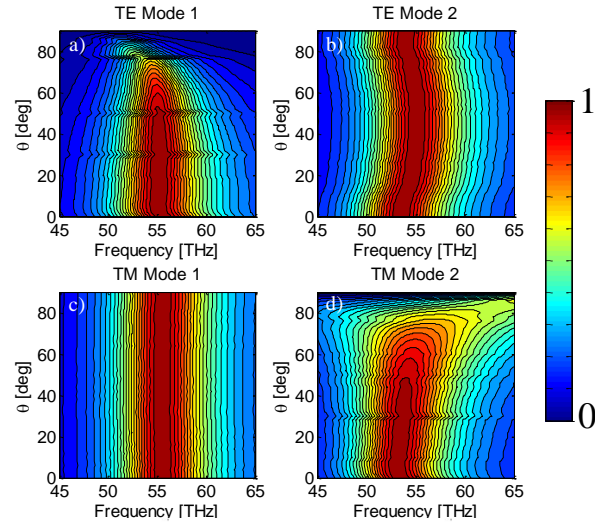

Fig. 4. Absorption spectra and angular-resolved for the criticallycoupled cases as seen in Fig. 1. Both TE (a, b) and TM (c, d) polarizations are shown for both mode orientations.
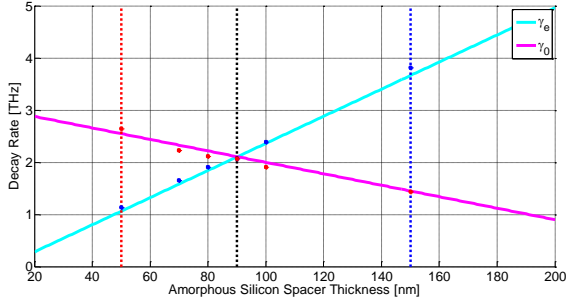
Fig. 5. Linear relationship between external (γ_e) and internal (γ_0) loss as a function of amorphous space thickness. CC case is seen at around 90 nm Si spacer thickness.

III. PAMM-MOLECULAR COUPLING

A. TCMT description

Similar to a EIT/EIA or Fano resonate description, the molecular resonance is a weak coupling to the incident field and is a spectrally narrow resonance, therefore; it can be model as a dark resonance mode [2, 10, 12, 13] as seen in the circuit realization in Fig. 6. However, the PAMM provides a spectrally broad resonance and couples very strongly with the incident field, and thus can be thought as a bright mode as seen in Fig. 6.
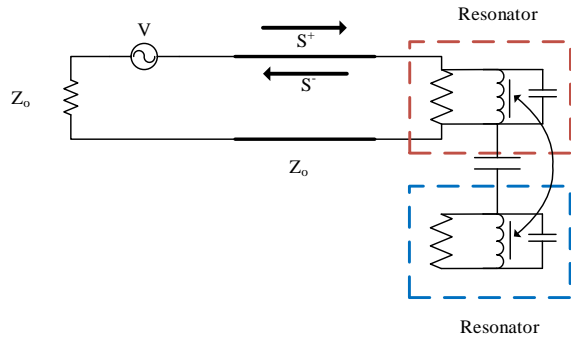

Fig. 6. Circuit visualization for a single input couple resonator system.

Using TCMT, an addition equation describing the molecule resonance, "a_2", with complex frequency $\omega_2 + j\gamma_2$. The interaction or mode coupling between the PAMM and molecule resonance is describe with the coupling strength "V". The coupling strength is a result of the interactions of the near field from the PAMM on to the PMMA molecular bond. The TCMT system of equations that results are as follow [2, 18]:

$$\frac{da_1}{dt} = j\omega_1 a_1 - (\gamma_1 + \gamma_e)a_1 + jVa_2 + \alpha S_1^+, \quad (6.a)$$

$$\frac{da_2}{dt} = j\omega_2 a_2 - \gamma_2 a_2 + jVa_1, \quad (6.b)$$

$$S^- = cS^+ + da_1. \quad (6.c)$$

B. PAMM molecular resonance coupled numerical results

With the introduction of the PMMA phonon resonance, it can be observed that for the CC and UC cases in mode splitting or EIT [12, 14] while OC case results in superscattering or EIA [15, 16] as seen in Figs. 7 and 8. Figure 8 (b) shows the avoidance crossing dispersion relation [8] where the variation of the only the A-Si spacer thickness resulting changing the PAMM resonances or absorption maximum spectrally.

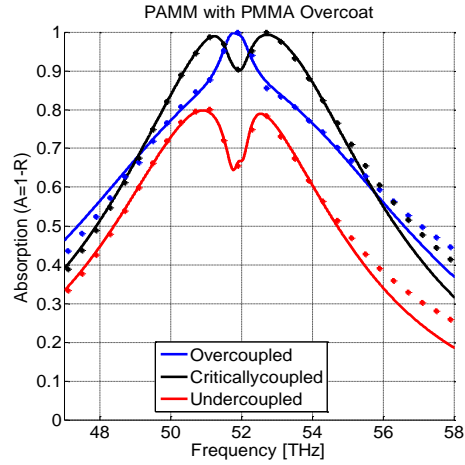Fig. 7. Numerical (*) and mathematical model (solid line) results are shown for resonant coupling between PAMM and molecular resonance.
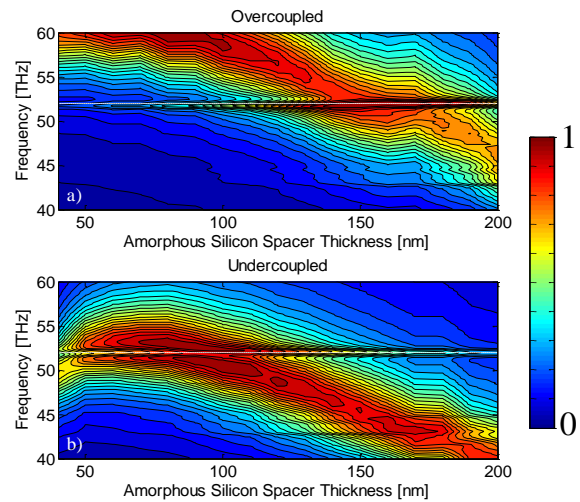

Fig. 8. Absorption spectra for PAMM resonantly coupled to a molecular resonance at 52 THz (white dashed line) for OC (a) and UC (b) cases for varies A-Si spacer thickness.

IV. CONCLUSION

With the use of finite element method and TCMT, a perfect absorbing patch metamaterial resonant coupled

to a C=O doubled bond at 52 THz of PMMA was investigated. EIT and EIA responses and anti-crossing dispersion are evident when resonated couple between a PAMM and molecule resonance is presented.

REFERENCES

[1] A. V. K., A. Boltasseva, and V. M. Shalaev, "Planar photonics with metasurfaces," *Science*, vol. 339, no. 6125, pp. 1232009-1-1232009-6, 2013.

[2] R. Adato, A. Artar, S. Erramilli, and H. Altug, "Engineered absorption enhancement and induced transparency in coupled molecular and plasmonic resonator systems," *Nano Letters*, vol. 13, no. 6, pp. 2584-2591, May 2013.

[3] C. Wu, B. Neuner, and G. Shvets, "Large-area, wide-angle, spectrally selective plasmonic absorber," *Physical Review B*, vol. 84, no. 7, pp. 075105:1-7, August 2011.

[4] K. Chen, R. Adato, and H. Altug, "Dual-band perfect absorber for multispectral plasmon-enhanced infrared spectroscopy," *ACS Nano*, vol. 6, no. 9, pp. 7998-8006, August 2012.

[5] N. I. Landy, S. Sajuyigbe, J. J. Mock, D. R. Smith, and W. J. Padilla, "Perfect metamaterial absorber," *Physical Review Letters*, vol. 100, no. 20, pp. 207402-207408, May 2008.

[6] H. A. Haus, *Waves and Fields in Optoelectronics*. Englewood Cliffs, NJ: Prentice-Hall, Inc., 1984.

[7] H. A. Haus and W. Huang, "Coupled-mode theory," *Proceeding of the IEEE*, vol. 79, no. 10, pp. 1505-1518, 1991.

[8] L. Novotny, "Strong coupling, energy splitting, and level crossings: A classical perspective," *American Journal of Physics*, vol. 78, no. 11, pp. 1199-1202, 2010.

[9] S. Savasta, et al., "Nanopolaritons: Vacuum Rabi splitting with a single quantum dot in the center of a dimer nanoantenna," *ACS Nano*, vol. 4, no. 11, pp. 6369-6376, 2010.

[10] C. Wu, et al., "Fano-resonant asymmetric metamaterials for ultrasensitive spectroscopy and identification of molecular monolayers," *Nature Materials*, vol. 11, pp. 69-75, January 2012.

[11] B. Luk'yanchuk, et al., "The Fano resonance in plasmonic nanostructures and metamaterials," *Nature Materials*, vol. 9, pp. 707-715, 2010.

[12] A. B. Khanikaev, C. Wu, and G. Shvets, "Fano-resonant metamaterials and their applications," *Nanophotonics*, vol. 2, no. 4, pp. 247-264, 2013.

[13] R. Adato, S. Aksu, and H. Altug, "Engineering mid-infrared nanoantennas for surface enhanced infrared absorption spectroscopy," *Materials Today*, vol. 00, no. 00, pp. 1-11, March 2015.

[14] C. L. Garrido Alzar, M. A. G. Martinez, and P. Nussenzveig, "Classical analog of electro-megnetically induced transparency," *American Journal of Physics*, vol. 70, no. 1, pp. 37-41, 2002.

[15] X. Zhang, et al., "Electromagnetically induced absorption in a three-resonator metasurface system," *Scientific Reports*, vol. 5, no. 10737, pp. 1-9, May 2015.

[16] W. Wan, W. Zheng, Y. Chen, and Z. Liu, "From Fano-like interference to superscattering with a single metallic nanodisk," *Nanoscale*, vol. 6, no. 15, pp. 9093-9102, May 2014.

[17] A. Lovera, B. Gallinet, P. Nordlander, and O. J. F. Martin, "Mechanisms of Fano resonances in coupled plasmonic systems," *ACS Nano*, vol. 7, no. 5, pp. 4527-4536, April 2013.

[18] S. Zanotto, et al., "Perfect energy-feeding into strongly coupled systems and interferometric control of polariton absorption," *Nature Physics*, vol. 10, no. 11, pp. 830-834, August 2014.

[19] J. Ginn, et al., "Characterizing infrared frequency selective surfaces on dispersive media," *ACES Journal*, vol. 22, no. 1, pp. 184-188, March 2007.

[20] T. Chen, S. Li, and H. Sun, "Metamaterials application in sensing," *Sensors*, vol. 12, no. 3, pp. 2742-2765, February 2012.

[21] D. J. Shelton, et al., "Strong coupling between nanoscale metamaterials and phonos," *Nano Letters*, vol. 11, no. 5, pp. 2101-2108, April 2011.

Unmanned Aerial Vehicle Platform Stabilization for Remote Radar Life Sensing

Robert H. Nakata, Brian Haruna, Scott K. Clemens, Daren Martin, Charles Jaquiro, and Victor M. Lubecke

Department of Electrical Engineering
University of Hawaii at Manoa, Honolulu, HI 96816, USA
rnakata6@hawaii.edu

Abstract ─ Unmanned Aerial Vehicle (UAV) platforms are increasingly ubiquitous and an ideal platform for rapid deployment to conduct remote sensing. However, for radar sensors that measure the phase of the signal of interest, the platform must be stabilized to avoid signal distortion. Measurement of respiratory motion with a continuous wave Doppler radar sensor is vulnerable to platform motion and requires a stable platform and post-detection motion compensation signal processing. We have investigated feedback stabilization techniques via simulation and empirical measurements using a bench top test fixture to remove the motion noise, where we observed a 86% reduction in motion, resulting in a SNR improvement of 29 dB after motion compensation.

Index Terms ─ Motion compensation, radar, remote sensing, UAV.

I. INTRODUCTION

Unmanned Aerial Vehicles (UAVs) have the potential for post-disaster search and rescue missions where triage can be conducted on victims using an on-board radar sensor to detect respiratory motion [1].

Vital signs measurements using a stationary Continuous Wave (CW) Doppler radar sensor have been previously demonstrated [2]. Since the signal corresponds to the phase modulation resulting from the range variation between the radar and the subject, any sensor platform motion will induce an undesired phase component to the respiration signal. Other papers describe motion cancellation techniques for vital signs sensing when the subject motion interferes with the measurement [3]. For our scenario, the assumption is that the subject is stationary as is likely the case for a post-disaster scenario where victims are prone on the ground.

Our hypothesis is that the combination of mechanically stabilizing the platform in conjunction with baseband signal processing will improve the SNR of the target signal and improve the probability of detection. The concept is similar to camera stabilization systems that apply vibration reduction and pixel shifting to de-blur images resulting from camera shake.

In this paper, we describe the concept of operations (CONOPS) for a UAV post-disaster search platform, platform motion compensation architecture and experimental results for ultrasonic sensor driven motion compensation. A programmable bench top test platform was constructed to mimic the unwanted UAV platform motion. A motion compensation sub-platform was mounted on the base platform and was programmed to compensate for the base platform motion based on inputs from an ultrasonic sensor that measured the unwanted motion in real-time. A 10 GHz radar sensor was mounted on the bench top test platform and pointed at a respiration phantom to assess the performance with and without the motion compensation.

In the following sections, we describe the CONOPS, the motion analysis, simulation, experiment, ultrasonic ranging sensor and experimental results.

II. CONCEPT OF OPERATIONS

The operational concept is for the system to have a search mode, a platform stabilization and signal acquisition mode as shown in Fig. 1. In the search mode, the UAV navigates to the area of interest using GPS waypoint coordinates or could navigate autonomously. An onboard camera with image recognition could be used to identify potential victims (targets) [4, 5]. In the stabilization mode, a suite of sensors, including GPS, IMU, LIDAR and/or ultrasonic range sensors, are used to adjust the UAV Electronic Speed Controllers (ESC) to maintain a steady altitude and fixed Yaw, Pitch and Roll (YPR) attitude. In the platform stabilization mode, the UAV hovers above the subject and uses the stabilization techniques described in this paper to improve the target SNR.
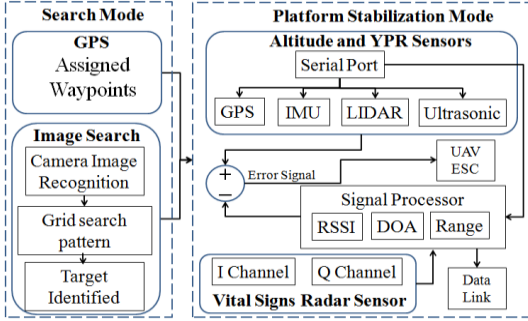
Fig. 1. System block diagram for the search and stabilization modes for the UAV radar sensor platform. Conventional sensors (GPS, IMU, barometric sensors) that determine altitude and yaw, pitch, roll (YPR) are supplemented with range sensors (ultrasonic, LIDAR) for motion compensation. The SNR of the on-board vital signs radar sensor pointed downward is improved by the platform stabilization and post detection baseband signal processing.

III. MOTION ANALYSIS, SIMULATION AND EXPERIMENT

A. Motion analysis

The time domain representations for the respiration signal of interest (modeled as a sinusoid for simplicity) and platform motion components are:

$$\text{respiration signal: } x_1(t) = A\sin(\omega_1 t), \quad (1)$$
$$\text{UAV motion: } x_2(t) = \sin(\omega_2 t). \quad (2)$$

The motion compensation signal $x_3(t)$ is derived from the secondary sensors on board the UAV including the inertial measurement unit IMU(t), ultrasonic sensor U(t) and lidar sensor L(t) with the composite signal represented as:

$$x_3(t) = A*IMU(t) + B*U(t+\tau_u) + C*L(t+\tau_L), \quad (3)$$

where A, B, C are scaling factors applied to each sensor and τ_i is the sensor signal delay for each sensor i.

The IMU sensor signal is:

$$IMU(t) = a(t) + m(t) + g(t), \quad (4)$$

where a(t), m(t) and g(t) are the accelerometer, magnetometer and gyroscope signals, respectively.

It should be noted that IMU sensors are subject to position errors due to the double integration operation required to derive position from acceleration. However, the IMU can be used to determine yaw, pitch and roll to determine the platform attitude to account for the offset pointing angle of the ultrasonic and/or LIDAR sensor that are used to determine the range to ground.

The ultrasonic sensor signal is:

$$U(t) = u(t + \varphi_u + \tau_u) = x_u(t), \quad (5)$$

the lidar sensor signal is:

$$L(t) = l(t + \varphi_L + \tau_L) = x_L(t), \quad (6)$$

and φ_i is the measured sensor phase. Note that we explicitly include the sensor phase noise and delays (φ_u and τ_u) as these parameters contribute to the phase compensation error.

As described in the introduction, the range sensor outputs, $x_u(t)$ and $x_L(t)$, can be used to adjust the platform position by providing proportional inputs to the ESC that adjusts the thrust of each motor. Additionally, the same range sensor outputs can be used to extract the platform motion via baseband signal processing.

The motion compensated radar signal $x_1'(t)$ in-phase (I) and quadrature component (Q) include the subtracted sensor phase components as shown below:

$$I = kA\cos[(\omega_1 t) + \varphi_r u(t - \varphi_u - \tau_u) + l(t - \varphi_L - \tau_L)], \quad (7)$$
$$Q = kA\sin[(\omega_1 t) + \varphi_r u(t - \varphi_u - \tau_u) + l(t - \varphi_L - \tau_L)]. \quad (8)$$

B. Motion compensation simulation

A Matlab Simulink program was written to simulate the effect of the UAV platform motion on the respiration signal of interest. An example with sinusoidal respiration and platform motion is shown in Fig. 2.
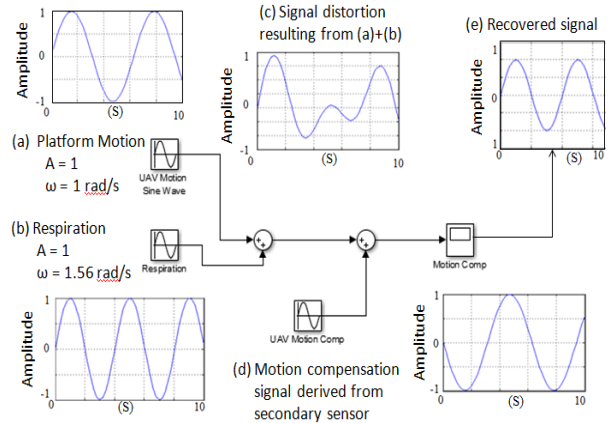

Fig. 2. Motion compensation simulation using Simulink: (a) UAV sinusoidal platform motion at 1 rad/s, (b) respiration sinusoid signal of interest at 1.56 rad/s, (c) motion distortion of signal of interest, (d) motion compensation error signal, and (e) recovered signal.

C. Experiment configuration

The motion compensation experiment block diagram is shown in Fig. 3. The target signal of interest is created with the Mover 1 linear actuator representing a respiration phantom. The UAV platform motion is injected with Mover 2 using a programmable linear actuator from Galil Motion Systems. The motion compensation corrective motion is created with Mover 3. Mover 3 was implemented using a slide potentiometer linear actuator mounted on wheels to allow independent motion from the Mover 2 base platform. An Arduino controller was programmed to control the position of Mover 3 based on the ultrasonic sensor range value. If optimally implemented, Mover 3 will cancel the undesirable Mover 2 motion. The test platform hardware configuration is shown in Fig. 4.
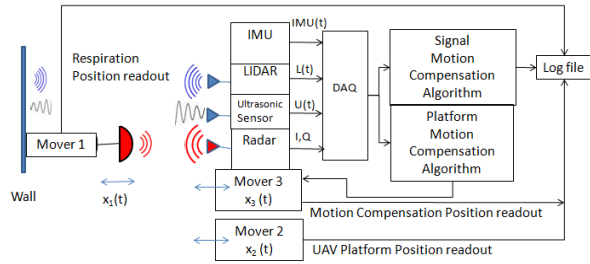
Fig. 3. Motion compensation experiment configuration block diagram. The respiration signal of interest (Mover 1) is detected by the radar that is mounted on the motion compensator (Mover 3) that is mounted on the base platform representing the unwanted motion from a UAV (Mover 2).
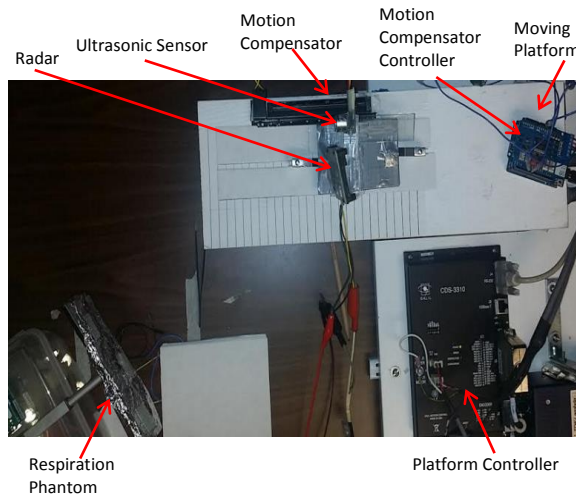

Fig. 4. Top view of motion compensation test bench components. Motion compensator (top center) is mounted on white plate mounted on platform motion actuator. Ultrasonic sensor is pointed at flat plate to the left (not shown) representing the ground. Radar is pointed at respiration phantom (lower left).

D. Ultrasonic sensor for feedback motion compensation

We focused on the ultrasonic sensor for this paper. Ultrasonic sensors operate by emitting high frequency pulses that are then reflected by a target. After reflecting off the target, the echo is then received by the sensor and the time difference is measured. With the time difference, the distance of the object can be calculated using the speed of sound. The IMU and lidar sensors will be tested in the future when the sensor fusion algorithm is developed and refined.

E. Motion compensation algorithm

A motion controller from Galil and mover were operated using code designed in Galil Tools. The controller made the mover perform a sinusoidal motion emulating the undesired platform motion. Additional

code was written for an Arduino Uno controller that controlled the motion compensation mover. The Arduino code reads the output from the ultrasonic sensor and moves the motion compensation actuator relative to the difference between the received distance value and a reference distance value.

A PID (Proportional, Integral, Derivative) stage was also implemented in the feedback algorithm as shown in Fig. 5. The PID parameters allows for tuning of the feedback response to optimize the motion compensation. The optimal PID values were empirically derived as P = 68, I = 7.6 and D = 73. We also modeled the system transfer function using the Matlab System ID Toolbox to reduce the number of PID empirical permutations.
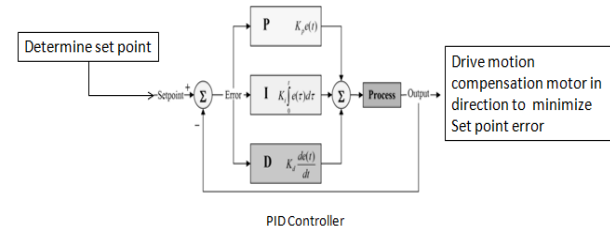

Fig. 5. PID controller in the motion compensation feedback loop.

IV. EXPERIMENT RESULTS

With the motion compensation enabled, the platform motion was reduced from 4.6 cm to 1.1 cm peak-to-peak for a 76% reduction in unwanted motion as shown in Fig. 6. This result was based on a constant gain feedback signal. After tuning the PID loop, we were able to reduce the compensated peak-to-peak motion to 0.6 cm for an 86% reduction in unwanted motion.

We also attempted to improve the compensation response with a position dependent gain factor, where the gain was proportional to the error voltage, resulting in faster convergence of the motion compensation. The result was not as significant as the PID result and was abandoned.
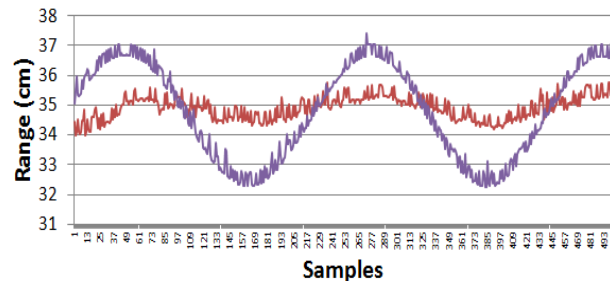

Fig. 6. Platform motion with 4.6 cm peak-to-peak sinusoidal waveform and 1.1 cm peak-to-peak compensated motion for a 76% reduction in unwanted platform motion amplitude using the constant gain feedback loop.

To assess the motion compensation system performance, we plotted the spectrum of the respiration signal alone (Fig. 7 (a)), the platform motion spectrum (Fig. 7 (b)), and the compensated simultaneous respiration and platform motion spectrum (Fig. 7 (c)). The spectra were obtained by performing a FFT in MatLab from the baseband radar signal. As shown in Fig. 7 (b), the respiration signal is masked by the platform motion noise. In Fig. 7 (c), the respiration signal is recovered after motion compensation is enabled.

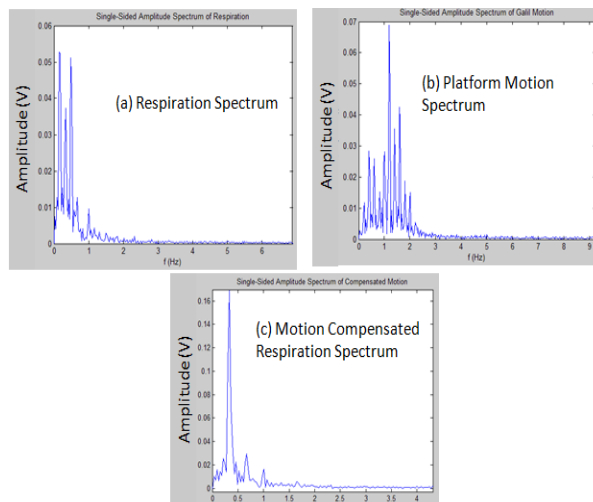The SNR for each case shown in Fig. 7 is summarized in Table 1.

Fig. 7. Baseband radar spectra for: (a) respiration signal at 0.4 Hz, (b) platform motion at 1.1 Hz, and (c) both respiration and platform motion with motion compensation enabled.

Table 1: SNR with and without motion compensation

	Voltage (V)		Power (V^2)		SNR (dB)
	Respiration	Motion	Respiration	Motion	
No compensation	0.028	0.07	0.0008	0.0049	-8.0
Motion compensation	0.17	0.015	0.0289	0.0002	21.1
SNR improvement					29.1

V. CONCLUSION

We demonstrated that unwanted platform motion can be compensated for, thereby improving the SNR of a Doppler radar signal. Analysis and simulation of secondary sensors to derive motion compensation signals in a feedback control system and empirical measurements with an ultrasonic sensor were conducted to compensate for unwanted platform motion. The experimental result was a 29 dB improvement in SNR. Future work will investigate other sensor types and multiple sensors with sensor fusion to further increase the system performance.

REFERENCES

[1] V. Ferrara, "Technical survey about available technologies for detecting buried people under rubble or avalanches," *WIT Transactions on The Built Environment*, vol. 150, pp. 91-101, May 2015.

[2] L. Ren and A. E. Fathy, "Noncontact heartbeat detection using UWB impulse Doppler radar," in *IEEE Radio Science Meeting*, Vancouver, BC, Canada, pp. 221, July 2015.

[3] C. Li and J. Lin, "Random body movement cancellation in Doppler radar vital sign detection," *IEEE Trans. on Microw. Theory Techn.*, vol. 56, no. 12, pp. 3143-3152, Nov. 2008.

[4] R. Sabatini, C. Bartel, A. Kaharkar, and T. Shaid, "Design and integration of vision based sensors for unmanned aerial vehicles navigation and guidance," *Proceedings of the SPIE, Optical Sensing and Detection II*, vol. 8439, Apr. 16-19, 2012.

[5] L. Xinhua and Y. Cao, "Research on the application of the vision-based autonomous navigation to the landing of the UAV," *Proceedings of 5th International Symposium on Instrumentation and Control Technology*, vol. 5253, pp. 385-388, Oct. 24-27, 2003.

Radar Noise Floor Method for Occupancy Detection

Pooja Nuti [1], Ehsan Yavari [2], and Olga Boric-Lubecke [1]

[1] Department of Electrical Engineering
University of Hawaii at Manoa, Honolulu, HI 96822, USA
poojan@hawaii.edu, olgabl@hawaii.edu

[2] Adnoviv LLC
Adnoviv LLC, Honolulu, HI 96822, USA
ehsan@adnoviv.edu

Abstract ─ A 2.4 GHz continuous wave Doppler radar sensor is utilized to carry out occupancy detection through detection of human presence over an empty room based on time domain root-mean-square (RMS) values. An existing system-on-chip with custom-made baseband board is employed for developing the radio.

Index Terms ─ Doppler radar, noise level, occupancy detection, Root-Mean-Square (RMS).

I. INTRODUCTION

Previously, occupancy detection has been conducted through detection of large motions or heat detection. Occupancy sensing technology is now moving away from such methods towards vital sign detection. In an effort to design such an occupancy sensor, noise level in conjunction with root-mean-square (RMS) is utilized [4].

Additionally, by rapid increase in global energy use, majority coming from fossil fuels, energy efficiency and conservation are becoming increasingly important. Studies show occupancy sensors can save up to 50% of that energy use [1]. Passive infrared and ultrasonic sensors are the two most common occupancy sensors in the market, however, they suffer from high rates of false alarms due to inconsistent ability to distinguish occupancy [2]. The feasibility of Doppler radar as an occupancy sensor is investigated in [3]. In [4] we investigated the effects of motion on the noise floor of a room and the potential to use that as a measure to discern an occupied room vs. an unoccupied one. In this paper we present experimental results with a human occupant, confirming that this technique can be used to discern human presence.

II. EXPERIMENTAL SETUP

A custom radar with single antenna is used for the measurements (Fig. 1). Radar transmits 2.4 GHz signal.

The radiated signal will be reflected back and received by same antenna. The received signal after down conversion and a conditioning circuit is digitized by the onboard ADC. The digitized signal is sent to a computer via usb port.
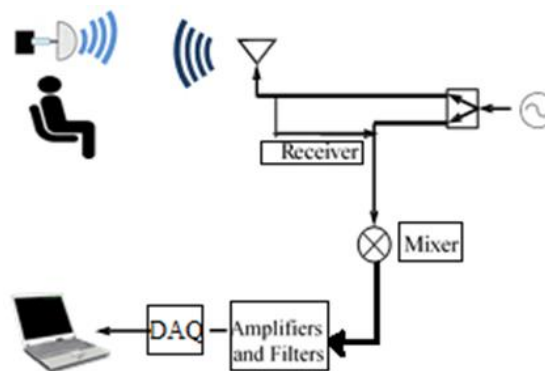
Fig. 1. Block diagram of system.

National Electrical Manufacturers Association (NEMA) standards are adhered to in order to obtain valid data during collection. NEMA requires that occupancy sensors be tested in an indoor area. The indoor area should be split into uniform cells in a grid pattern. By standards, these cells should be 3ft by 3ft in area. The testing environment should be controlled, such that temperature and humidity remain constant. In order to meet these requirements, a room with dimensions 3.5 m by 4.5 m with no windows was utilized. Additionally, the room was broken into 27 cells where the mechanical target/human subject used for vital-sign modeling could be moved through. See Fig. 2 for cell layout in room. Blue tapes in Fig. 3 mark the mechanical target locations throughout the room. The occupancy sensor along with passive infrared/ultrasonic hybrid sensor were wall-mounted. As per standard, the direction of motion

produced by the mechanical target was perpendicular to the sensor face.

Radar is used to detect presence in the room by detecting small periodic motions such as respiration in each individual cell. A precision single-axis linear stage is from Galil motion control (CDS-3310) with a pulse-width modulation (PWM) driver is employed for generating such periodic motions simulating human respiration. This mechanical target was moved throughout 27 full cells and radar return from the target at each of the 27 cells was recorded. The duration of recordings in each of the 27 cells was 90 seconds. A similar test with a human subject in place of the mechanical target was also performed under same conditions. Additionally, data collections were taken consisting of radar reflected signal from the same empty room with no mechanical target for estimating noise level in our measurements.
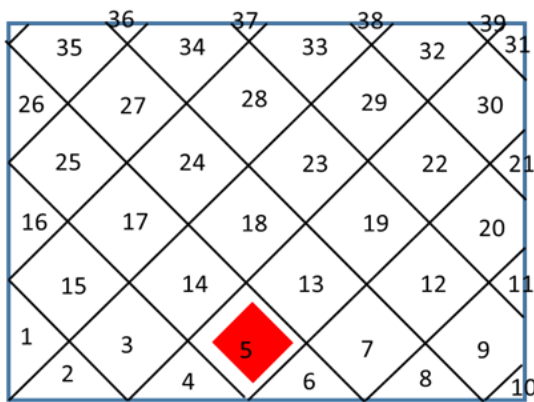

Fig. 2. Room layout consisting of 27 cells used for data collection. Note Cell 5 location in front of sensor face.

Fig. 3. Test setup [4].

III. EXPERIEMENTAL RESULTS

Analysis of data included plotting time-domain profiles to qualitatively observe the radar return from different cells, and noise return. Additionally, RMS was applied to all data collects to further quantify signal level for comparison purposes.

A. Results of tests with mechanical target

In the time domain, radar data from tests using a mechanical target can look similar to data collected from the radar with an empty room. This results in difficulty distinguishing between noise and radar signals. The similarities between radar signal from a mechanical target and noise are observed in comparing amplitudes of Figs. 4 and 5 to that of Fig. 6. These figures depict the reflected signal from a mechanical target at different locations in the testing room. After studying the raw radar data from each of the 27 cells tested, we found Cell 5 to have the strongest signal due to the closeness (0.5m) and perpendicularity to the radar antenna. Cell 21 is farther away from the radar field of view and has the weakest signal. The RMS value of Cell 5 is .2432 units and the RMS value of Cell 21 is .2538 units, comparatively the RMS of one set of empty room data (noise) is 0.2439 units. These values illustrate the similarities in return.

In order to account for the amplitude fluctuations observed in the signal the root mean square (RMS) of the time-domain data from various cells were taken. RMS is also used because it can account for both non-periodic variability associated with signal return from radar and noise, and also be compared easily. The built-in RMS function provided in Matlab was used to take the RMS of data matrices. The used RMS in these tests are defined in [4]. RMS was taken of the time domain data of radar from mechanical target in all 27 cells/locations in the room and compared with the RMS values of multiple noise recordings. The result of plotting both noise and radar data from mechanical target is shown in Fig. 6.
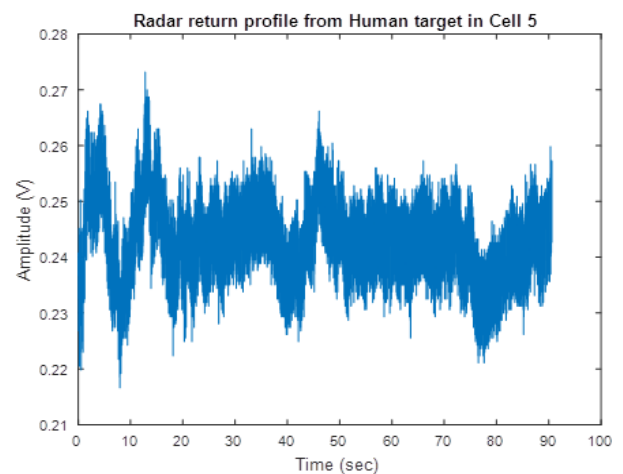

Fig. 4. Radar data from mechanical target in region of strong radar return.
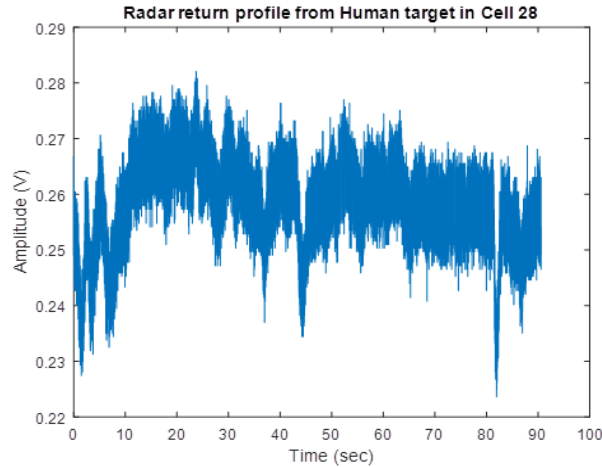
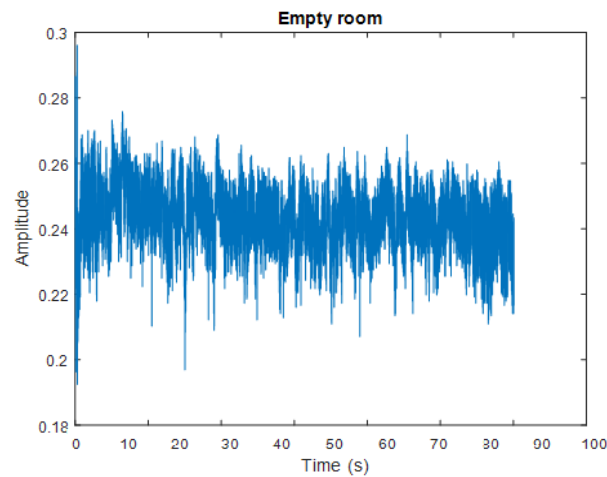Fig. 5. Radar data from mechanical target in region of weak radar return.

Fig. 6. Radar signal from empty room no mechanical target.

Figure 7 illustrates that despite the raw noise data and radar data with mechanical target seeming qualitatively similar; there is a distinction between their time domain RMS values. The mean of the RMS values of mechanical target return is 0.2523 units while the mean of RMS values of noise (empty room return) is 0.2430 units. These values yield an average difference of 0.0093 units. This difference is low due to the low frequency utilized in data collection (0.2 Hz). Radar baseband filters may have contributed to the low difference between radar signals with the mechanical target and noise data. The difference between mechanical target return and noise is consistent. 92.6% radar signal from the mechanical target could be distinguished from noise, as two RMS values

(corresponding to data collects from two unique cells) from mechanical target are within noise region and therefore undistinguishable.
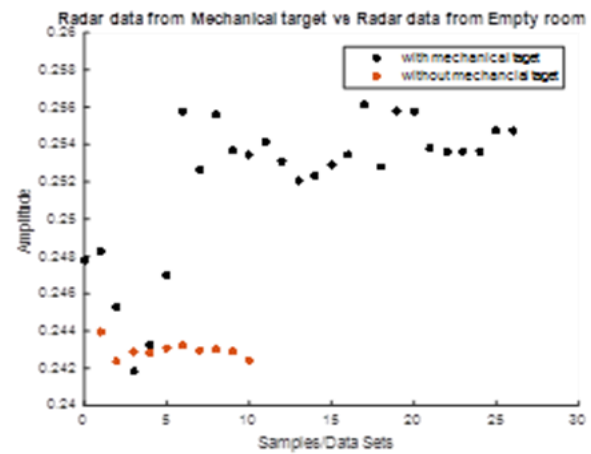

Fig. 7. Time domain RMS with and without mechanical target [4].

B. Results of tests with human subject

There is a qualitative difference between RMS amplitudes of return from a mechanical target in Fig. 7 and a human target (Fig. 10). Using a mechanical target resulted in a more distinct signal level from motion due to the more similar and concentrated amplitudes. In Fig. 10, the RMS amplitudes from the human target are less concentrated than that of the mechanical target. This may be due to the less-controlled frequency of periodic motion observed in human respiration in comparison to the constant frequency that the mechanical target operates at. The result of radar return from human respiration matches the result of radar return from a mechanical target. In these tests, a human was placed in each of the same 27 cells used to test the mechanical target. Similar to the results of the mechanical target, strongest return from the target was yielded from Cell 5 which is located right in front of radar antenna (see Fig. 8). The RMS amplitude of return from this cell was .2495 V. The ability of radar to detect vital signs is observed through the sinusoidal waveform (Fig. 8). This sinusoidal motion corresponds to human respiration. As the human target is moved into cells farther from the line of sight of antenna, radar return degrades. In Fig. 9, Cell 22 yields the lowest RMS amplitude of radar return (.2436 V). This lower amplitude is due to the distance between the radar antenna and test subject, and the location of cell 22 is not in direct path of antenna but rather on a side of the room. Despite the lower amplitude, the presence of a human target is still detected over the mean noise level of .2430 V.
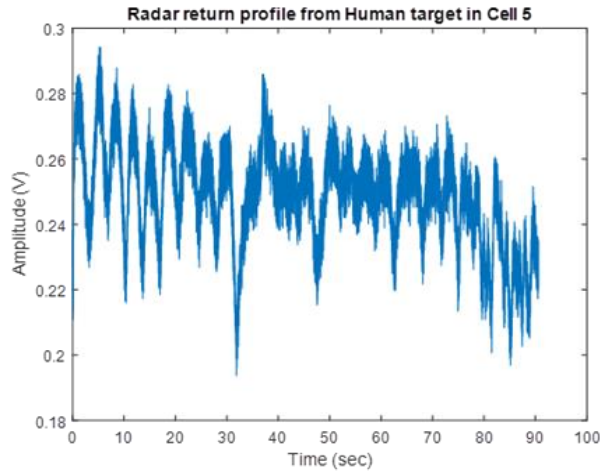
Fig. 8. Human respiration observed through sinusoidal waveform in time-domain.
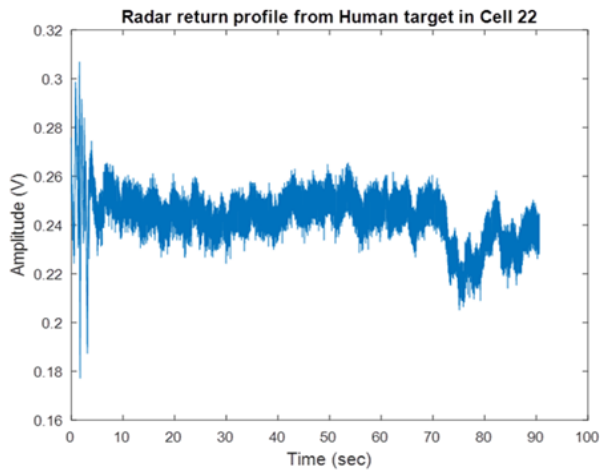

Fig. 9. Radar return from human target at distance away from antenna line of sight. Note similar profile to noise.
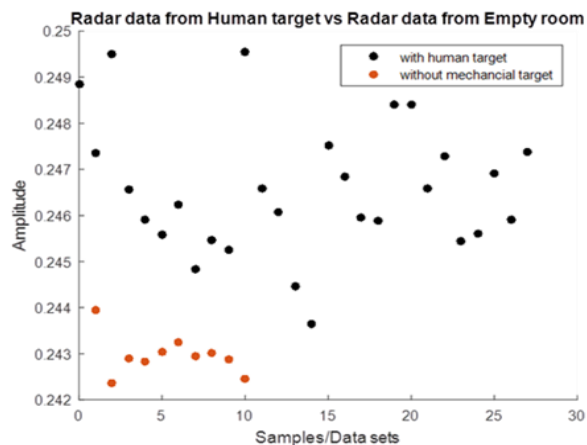
Fig. 10. Return from human target at generally higher signal level than noise level.

IV. CONCLUSION

A Doppler radar occupancy sensor was used for detecting presence. Noise floor of a room is used as a measure to detect occupancy. Experiments were performed to distinguish empty room radar return versus radar return from a mechanical target simulating respiration signal and human target with a resting respiration rate. Tests with a mechanical target agree with tests with a human target. In both cases, stronger human presence is observed in cells closest to the radar. Additionally, both cases result in 93% accuracy in detection of human presence over noise. The gap between RMS values can be utilized to distinguish movement from noise. Future work would focus on testing radar at different frequencies and quantifying the noise floor for the system. Additionally, future work could include modeling performance of radar and finding limits to which this method of occupancy detection is valid by attempting human presence detection over different environments.

REFERENCES

[1] X. Guo, D. K. Tiller, G. P. Henze, and C. E. Waters, "The performance of occupancy-based lighting control systems: A review," *Lighting Research and Technology*, vol. 42, no. 4, pp. 415-431, 2010.

[2] M. Kumar, D. Garge, and R. Zachery, "Multi-sensor fusion strategy to obtain 3-D occupancy profile," *Proceedings of the 31st Annual Conference of the IEEE Industrial Electronics Society (IECON)*, pp. 2083-2088, 2005.

[3] E. Yavari, H. Jou, V. Lubecke, and O. Boric-Lubecke, "Doppler radar sensor for occupancy monitoring," in *Proc. 2013 IEEE Radio Wireless Symp.*, pp. 316-318, 2013.

[4] E. Yavari, P. Nuti, and O. Boric-Lubecke, "Occupancy detection using radar noise floor," in *Proc. ICWITS/ACES 2016 Conference*, pp. 152-154, 2016.

UAV-Radar System for Vital Sign Monitoring

Ashikur Rahman, Yuta Ishii, and Victor Lubecke

Department of Electrical Engineering
University of Hawaii at Manoa
Honolulu, USA

Abstract ─ Feasibility and fabrication of components of a life-sign radar system on unmanned aerial vehicle (UAV) has been studied. A signal conditioning circuit has been simulated, fabricated, and tested for data preparation and acquisition. Application such as, vital-sign detection, using UAV in real-time requires wireless transmission of baseband data to the monitoring station. Methods have been devised to achieve this goal. A quadrature Doppler radar has been assembled using two single channel x-band MDU1020 radars. The united radars can avoid null point distortion in physiological monitoring. Examinations were performed with both mechanical targets.

Index Terms ─ Amplifier, demodulation, Doppler radar, filter, microwave, phase, phase noise.

I. INTRODUCTION

Short distance Doppler radars have been thoroughly studied and reported in RF based noninvasive measurements. Short distance radars can detect small motions (mm or submillimeter, based on frequency of operation) and hence can be used in vital sign measurements based on respiratory effort. In vital sign measurements continuous wave (CW) Doppler radar uses RF echo reflected from human torso. The phase of echo is proportional to the variable displacement across the body surface corresponding to the motion of heart and lungs. Due to the high sensitivity of short distance radar any motion artifact will distort the received RF signal bounced of any moving surface. Some potential application of vital sign measurements from a mobile platform can be very useful. For example, a UAV carrying a vital sign radar can be used as first responder in a battle field or a search and rescue robot during natural catastrophe. Such a system requires a portable vital-sign radar system, a data acquisition and processing system as shown in Fig. 1.

Single channel vital sign radar has major limitation since a null case may arise in numerous positions in front of the radar which gives inaccurate measurement [1]; hence, quadrature radar is preferred. Some manufacturers produce cheap single channel motion detector radars for specific purposes, i.e., MDU1020 motion detector by

Microsemi. These radars are application specific and subject to null point distortions depending on the nominal distance of the target and the radar. However, two of these radars can be placed in a way to avoid the single channel radar's constraint providing a cost effective means of vital sign detection [2]. Commercial motion detection units (MDU) or vital sign radars have reasonably good power output; 13-dBm, for instance, in a MDU 2400 module mass-produced by Microwave Solutions Ltd.
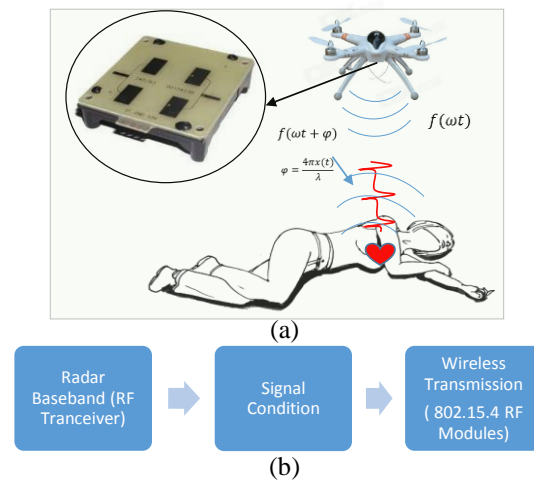
Fig. 1. An illustration of life sign detection using shot distance radars reveals: (a) UAV mounted radar (MDU-1020) for vital sign detection, and (b) shows a block diagram of the system components [3].

However, the baseband signal output is only a small portion of 13-dBm, due to the fact that vital sign related motion is very small compared to the radar's full range of exposure ability. This imposes the fact that the radar signal needs to be enhanced prior to sending the output to a communication link. A feasible solution is to apply ZigBee based on the IEEE 802.15.4 protocol. ZigBee device's analog sampling is limited to the range of 0 to 1.2 volts. Testing several MDU's over a range of a meter, the MDU's voltage output swing was found to be -20 mV to 20 mV for respiration which may easily be buried in

the noise during wireless transmission [3]. Moreover, since the ZigBee device cannot sample negative voltages, the radar output needs to be shifted, amplified and filtered. Integrating these functions requires a custom design of signal conditioning circuitry.

II. SYSTEM ARCHITECTURE

Radar measurement from a mobile platform is modulated by the motion artifact introduced by the mobile platform. The combined motion may show phase distortion, and saturation problems in received backscattered signal. Additionally, the occurrence of null points in the received radar signal makes the extraction of life signs challenging [4], [5]. These issues stem from inconstant traveling distance seen by backscattered RF between the radar antenna and the target [5]. A motion compensation technique for mobile CW Doppler vital signs radar has been demonstrated using high precision cameras [5]. Another work included development of a new, field applicable, low-cost methodology for motion artifact compensation using only a small RF tag [5]. Some other techniques related to vital sign measurements were reported in [6]-[8]. However, these works either focused on body motion cancellation, whereas, our work focuses on vital sign detection from a UAV. Additional sensors can characterize UAV platform motion to filter out the noise introduced by vibration and drift. The output of these sensors also requires some conditioning (amplification and filtering) as shown in the detailed system architecture in Fig. 2.
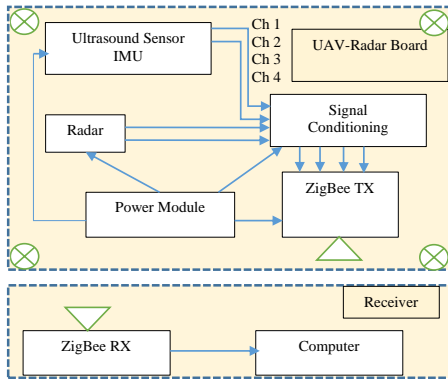

Fig. 2. A detail system diagram of UAV radar system for vital sign monitoring.

Application of ultrasound sensor can generate and approximate spatial map (in a confined space, or ground elevation for open space) of a moving object within the range of speed and sensitivity. This can help characterize platform motion, hence, useful data for radar signal processing in noise cancellation.

To test the concept by starting from the scratch, we needed low cost radars, able to perform demodulation of signal with greater quality. On the other hand, recorded signal needs to be cleaned up for further processing and filtering unwanted components. So, this worked focused on finding an engineering solution to a low cost radar also creation of custom signal conditioning circuit.

III. LOW COST RADAR SENSOR

We proposed a low cost solution to an IQ radar system. The price is well under of exisitng solutions. A typical qudrature radar MDU-4200 from Microwave Solution Ltd. Costs about $120, whereas, a single-channel MDU 1020 can be found as low as $7. A theoritical explanation presented below justifies the composition of two single channel radars for making an alternate solution to qudrature radar.

A. Radar theory of vital sign detection

A typical coherent continuous wave vital sign Doppler radar system sends an RF signal towards human torso, the echo is phase modulated due to the positional variation of moving body parts. The echo is mixed and down converted to retrieve the target's displacement. The output of a single channel receiver is given by [1]:

$$B(t) \approx A_B \cos\left(\theta + \frac{4\pi x(t)}{\lambda} + \frac{4\pi y(t)}{\lambda} + \Delta\emptyset\left(t - \frac{2d_0}{c}\right)\right), (1)$$

where θ is constant phase shift and $\Delta\emptyset$ is residual phase noise. $A_B, \lambda, x(t), y(t)$, and d_0 are baseband amplitude, wavelength, chest movement, heart movement and nominal distance between the radar and the target. Now, if somehow $\frac{\pi}{2}$ phase change is introduced, (1) will result in:

$$B(t) \approx A_B \sin\left(\theta + \frac{4\pi x(t)}{\lambda} + \frac{4\pi y(t)}{\lambda} + \Delta\emptyset\left(t - \frac{2d_0}{c}\right)\right). (2)$$

Our work proves that physical offset between two single-channel radars can achieve proper phase shift for quadrature radar channels.

B. Null and optimum demodulation

The constant phase shift in (1) is related to nominal distance and can be expressed as:

$$\theta = \frac{4\pi d_0}{\lambda} + \theta_0; \ d_0 = \frac{\lambda(\theta - \theta_0)}{4\pi}; \ d_0 = \frac{\lambda\left(\frac{k\pi}{2} - \theta_0\right)}{4\pi}. \quad (3)$$

If θ is denoted as the k multiple of $\frac{\pi}{2}$, the baseband output $B(t)$ will be either null or optimum for an integer value of k when small signal approximation is applicable [1]. From (1) and (2) we see that null and optimum baseband output occurs for the nominal distances as:

$$d_{NULL} = \frac{\lambda(m\pi - \theta_0)}{4\pi}; \ d_{OPT} = \frac{\lambda\left(m\pi + \frac{\pi}{2} - \theta_0\right)}{4\pi}, \quad (4)$$

where m is an integer. (3) reveals that the adjacent null and optimum separation is:

$$|d_{NULL} - d_{OPT}| = \frac{\lambda}{8}. \quad (5)$$

This examination proves the fact that if two single channel radars are oriented in similar way with a separation of $\frac{\lambda}{8}$ in the plane of nominal distance between

the target and the radar transceivers, one of the radars will be in optimum position given that the other is in null position and vice versa [2]. One fundamental assumption is that the difference of the residual phase noise of the radars is negligible.

C. Fabrication and testing

Two MDU1020 radars were used for the experiment. MDU1020 has the operating frequency is 10.525 GHz. The theoretical distance between consecutive null and optimum points is 3.5629 mm; $\frac{\lambda}{8}$ in other words [2]. Now creating the physical separation of this length in traveling plane of the wave should ensure fail-safe design. In case the nominal distance between one of the radars is in null position, the other will be in optimum position. The two radars were physically separated approximately 3.5 mm, as shown in Fig. 3. A mechanical mover was programmed to move in 0.5 Hz sinusoidal motion. The nominal distance between the radar pair and target was varied in the range of [0.5 m (0.5+.0035) m]. Linear demodulation has been performed using the output (voltage proportional to displacement) of the two radars as shown in Fig. 4.
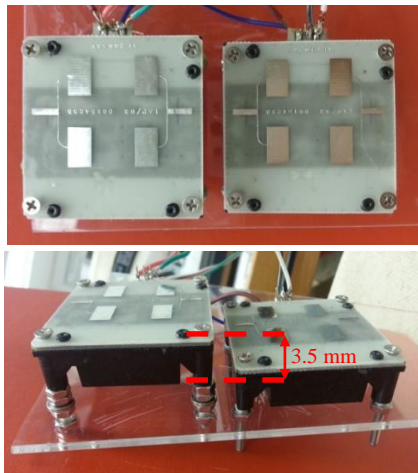
Fig. 3. Quadrature radar assembly is illustrated. A physical offset of approximately 3.5 mm has been kept between the two MDU1020 single channel k band radars [2].
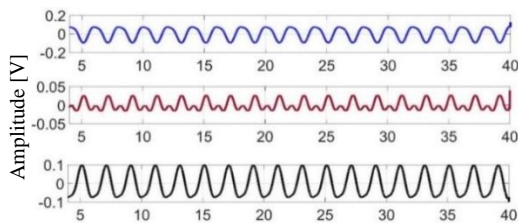
Fig. 4. Shows the two radar output and combined output when the nominal distance is gradually changed. It is evident that one radar suffers distortion, but the other radar helps demodulating the correct signal [2].

IV. SIGNAL PROCESSING MODULE

A. Design requirment

The design requirements of the signal conditioning circuit are amplification, filtering, light-weight, DC offsetting capability for level shifting.

B. Simulation

General purpose uA741 operational amplifiers were selected for design and fabrication of the quick prototype. The circuit has two stages as shown in Fig. 5. The first stage is a voltage follower while the second stage is a summing inverting amplifier. Passive components have been used for filtering, i.e., variable resistors along with capacitors were used to provide DC offset, tunable band and gain. Figure 6 shows some tuning simulation.

(a)   (b)
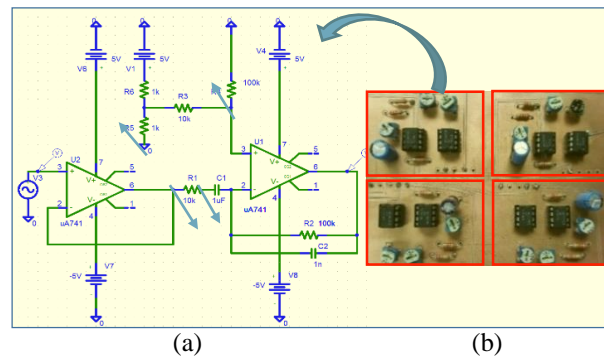
Fig. 5. Software simulation and layout is shown: (a) depicts the circuit diagram of the single channel amplifier, and in (b) fabricated circuit containing four amplifier blocks [3].

(a)

(b)
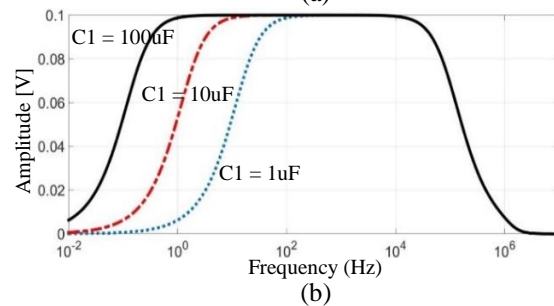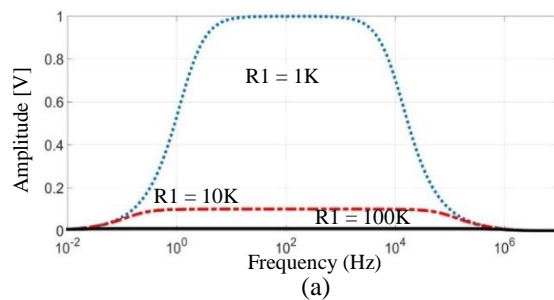
Fig. 6. Bandwidth and gain variation by tuning resistance in (a), and capacitance in (b). R1 and C1 in Fig. 5 (a) were varied.

C. Fabrication and testing

PCB layout was made using a LPKF mid-range PCB milling machine. A double sided FR-4 board was used to print the layout. A four-channel signal conditioning circuit was designed, so the circuit is capable of conditioning four channels.

V. RESULTS

Experiments were to check the performance of signal conditioning, wireless data acquisition. Target's motion and a simple platform motions were simulated using mechanical moving stages, as demonstrated in Fig. 7. The radar sensor and signal conditinging circuitry was mounted on a linear stages. Figure 8 illustrates the presense of noise in radar signal due to motion artifact simulated using linear stage. The radar sensor data was recorded via ZigBee communication link.
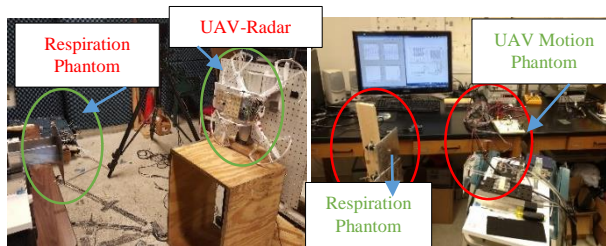

Fig. 7. (a) Shows experimental assembly, a DJI phantom quadcopter carrying the circuitry, and (b) shows experiments, simulating platform motion and target motion with two different linear stage.
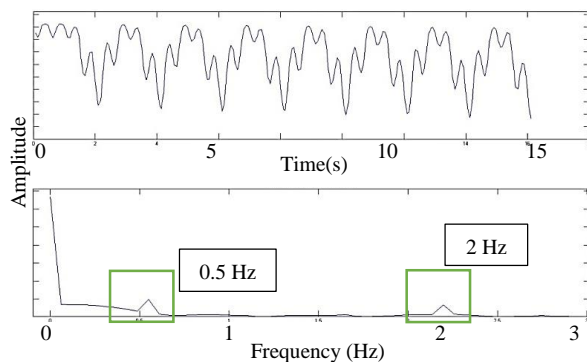

Fig. 8. Shows the time and frequency domain plots of radar data that contains two motion signature, platform motion was 2 Hz and target motion was 0.5 Hz simple periodic motion. Both platform motion and target's motion are in composite signal.

VI. CEM APPLICATION

This work implemented off-the-shelf radar having vertical and horizontal polarization of 36 degree, and 72 degree respectively in patch antenna. For motion artifact compensation low-IF RF tags can be used. A

computational EM (CEM) simulation will help making a robust system.

VII. CONCLUSION

UAV-Life sign Radar system has been studied and system components were built to provide a low cost fabrication of the platform. A low noise lightweight four-channel amplifier with DC offsetting and frequency tuning capability was fabricated. A low cost solution for quick prototyping of distortion-less radars system has been proposed with practical implementation. Overall, our work opens up the potentials for using UAV for life sign monitoring.

REFERENCES

[1] A. Droitcour, V. Lubecke, J. Lin, and O. Boric-Lubecke, "A microwave radio for Doppler radar sensing of vital signs," in *IEEE MTT-S Int. Microw. Symp. Dig.*, Phoenix, pp. 175-178, 2001.

[2] A. Rahman, E. Yavari, and V. Lubecke, "Single channel radar fusion for qudrature life-sign Doppler radar," in *IEEE ACES*, Honolulu, pp. 1-2, 2016.

[3] A. Rahman, A. Nakahara, R. Nakata, and V. Lubecke, "Signal conditioning for UAV-radar in vital sign monitoring," in *IEEE ACES*, Honolulu, pp. 1-2, 2016.

[4] A. Rahman, E. Yavari, X. Gao, V. Lubecke, and O. Boric-Lubecke, "Signal processing techniques for vital sign monitoring using mobile short range Doppler radar," in *IEEE Radio Wireless Symp.*, San Diego, CA, pp. 1-3, 2015.

[5] A. Rahman, E. Yavari, V. Lubecke, and O.-B. Lubecke, "A low-IF tag-based motion compensation technique for mobile doppler radar life sign monitoring," *IEEE Transactions on Microwave Theory and Techniques*, vol. 63, no. 10, pp. 3034-3041, Oct. 2015.

[6] L. Ren and A. E. Fathy, "Noncontact heartbeat detection using UWB impulse Doppler radar," in *IEEE Radio Science Meeting*, Vancouver, BC, Canada, pp. 221, July 2015.

[7] V. Ferrara, "Technical survey about available technologies for detecting buried people under rubble or avalanches," *WIT Transactions on the Built Environment*, vol. 150, pp. 91-101, May 2015.

[8] C. Li and J. Lin, "Random body movement cancellation in Doppler radar vital sign detection," *IEEE Trans. on Microw. Theory Techn.*, vol. 56, no. 12, pp. 3143-3152, Nov. 2008.

Efficient Electromagnetic Simulation Including Thin Structures by Using Multi-GPU HIE-FDTD Method

Yuta Inoue and Hideki Asai

Research Institute of Electronics
Shizuoka University, Hamamatsu, 432-8561, Japan
inoue.yuta@shizuoka.ac.jp, asai.hideki@shizuoka.ac.jp

Abstract ─ This paper describes an efficient simulation method to solve the large scale electromagnetic problems with thin unit cells in the finite-difference time-domain (FDTD) simulation. The proposed method is based on the hybrid implicit-explicit and multi-GPU techniques, which can choose a larger time step size than that in the conventional one, and using the multiple graphic processing units (GPUs). In the proposed method, the computational time is significantly reduced.

Index Terms ─ Electromagnetic simulation, FDTD method, multi-GPU, time domain analysis.

I. INTRODUCTION

The finite-difference time-domain (FDTD) [1] method is one of the numerical simulation techniques for solving the electromagnetic problems. The FDTD method is a conditionally stable method. Therefore, the maximum time step size is constrained by the Courant-Friedrichs-Lewy (CFL) condition. If the time step size is not satisfied with the CFL condition, the FDTD method becomes unstable. For the analysis of the large scale electromagnetic problems with thin structures such as printed circuit boards, the time step size must be small and it can make the FDTD simulation a huge time consuming task. Thus, the efficient electromagnetic simulation technique is strongly demanded for the efficient designs. In order to overcome the CFL condition problem, the several unconditionally stable methods have been proposed for an arbitrary time step size [1], [2]. However, these methods are unsuitable for the parallel implementation because several overheads degrade the efficiency of parallel computing.

In order to alleviate the CFL condition problem, the hybrid implicit-explicit (HIE)-FDTD method has been proposed and studied for the fast electromagnetic simulation with thin unit cells in the FDTD simulation [3]-[7]. The HIE-FDTD method can adopt a larger time step size than that for the conventional FDTD method. The implicit technique is employed partially, and the computational domain can be easily divided for the parallel computing. Therefore, the message passing interface (MPI)-based parallel-distributed HIE-FDTD method [6] and the general purpose computing on graphic processing unit (GPGPU) based massively parallel HIE-FDTD method [7] have been proposed for the efficient simulation. However, the parallel distributed HIE-FDTD method is required to be faster since CPU is slower than the graphic processing unit (GPU). On the other hand, the memory size of the GPU boards is not sufficient for the large scale problems. Hence, the combination of the MPI based method and the GPGPU based method is demanded for solving the large scale problems. As a result, the proposed method can solve the large scale problems and can reduce the elapsed time drastically from MPI-based and single GPU-based method.

In this paper, the multi-GPU based HIE-FDTD method with MPI and CUDA is proposed for the efficient electromagnetic simulation of the object with thin structures. First, the HIE-FDTD method is reviewed briefly. Next, the proposed method is described. Finally, the efficiency of the proposed method is evaluated through several FDTD simulations.

II. HIE-FDTD METHOD

The HIE-FDTD method [4] has been proposed for the efficient 3-D electromagnetic simulation of the given object with thin unit cells in the FDTD-based computational domain. Here, it is assumed that the given object has the fine scale dimension in the z direction such as printed circuit boards. In such a case, the updating formulas of the HIE-FDTD method consist of the two explicit equations which do not contain the derivatives with respect to z direction and four implicit equations which include the derivatives with respect to z direction. The updating formulas are described in [6].

The updating procedures of the HIE-FDTD method are partially different from the conventional FDTD method. First, E_z and H_z are explicitly updated. Next, E_x and E_y are updated by numerical solution method of simultaneous linear equations such as LU decomposition

method. After updating E_x and E_y, H_x and H_y are explicitly updated.

The CFL condition of the HIE-FDTD method is alleviated than that of the conventional FDTD method, which is given by:

$$\Delta t_{FDTD} \leq \frac{1}{c\sqrt{\Delta x_{\min}^{-2} + \Delta y_{\min}^{-2} + \Delta z_{\min}^{-2}}}, \qquad (1)$$

where Δt_{FDTD} is maximum time step size for the conventional FDTD method, c is the speed of light, $\Delta x_{\min}$, $\Delta y_{\min}$ and $\Delta z_{\min}$ are minimum cell sizes along the x, y, and z direction in the computational domain. The CFL condition of the HIE-FDTD method is determined as follows:

$$\Delta t_{HIE} \leq \frac{1}{c\sqrt{\Delta x_{\min}^{-2} + \Delta y_{\min}^{-2}}}, \qquad (2)$$

where Δt_{HIE} is maximum time step size for the HIE-FDTD method. From (2), the Δz is removed from the CFL condition of the HIE-FDTD method. Therefore, the time step size can choose the larger time step size than that for the conventional FDTD method. In the case of thin cell along with z direction, the HIE-FDTD method can efficiently simulate.

III. MULTI-GPU HIE-FDTD METHOD

The multi-GPU HIE-FDTD method is combination of the parallel distributed HIE-FDTD method and the GPGPU HIE-FDTD method. In the proposed method, the arithmetic operations are performed by a GPU instead of a CPU. In order to employ multi GPUs, the proposed method uses the three types of domain decomposition techniques. One is the domain decomposition technique for allocating a GPU to a subdomain. The others are the domain decomposition techniques for the GPU computing. In this section, the domain decomposition techniques and updating procedure are described. Here, the MPI library is employed for network communication between the neighboring subdomains and the CUDA is done for the GPU computing.

A. Domain decomposition

First, the domain decomposition technique for allocating a GPU to a subdomain is described. In the proposed method, the whole 3D spatial domain is divided into the several subdomains along the x and y directions. The number of subdomains is same as the number of GPUs. Note that the boundary cells of subdomain are overlapped with the neighboring subdomains. The overlapping boundary cells are employed to communicate magnetic components between the neighboring subdomains. Furthermore, the dummy cells are added to each subdomain in order to correct the total number of cells of x-y plane to a multiple of 64. The dummy cells are employed for the GPU

computing. In the updating process, the electromagnetic components at the dummy cells are not updated.

Next, the domain decomposition techniques for GPU computing are shown. The partitioned subdomains for GPU computing are illustrated in Fig. 1. Here, NX, NY, and NZ are the numbers of cells for the x, y, and z directions, respectively. In the proposed method, the domain decomposition techniques are different in the explicit updating procedure and the implicit updating procedure. Figure 1 (a) shows the divided subdomain for explicit updating procedure. The subdomain is partitioned by the blocks which are composed of the 64 threads. A cell is allocated to a thread and is updated by the thread. The thread is smallest element of the process. Therefore, in the explicit updating procedure, the subdomain is divided into (NX×NY×NZ)/64 blocks. On the other hand, the domain decomposition technique for implicit updating procedure is illustrated in Fig. 1 (b). In the implicit updating procedure, x-y plane of the sub-domain is partitioned by the blocks. Thus, each block is assigned to 64×NZ cells. Each thread in the block is allocated to the NZ cells. These cells are updated by using the LU decomposition method [8].
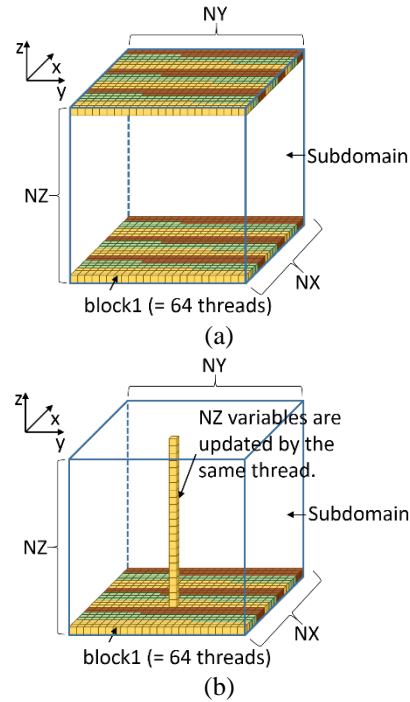

Fig. 1. The domain decomposition and block structure for GPU computing. (a) Domain decomposition for explicit updating procedure, and (b) domain decomposition for implicit updating procedure.

B. Updating procedure

In the proposed method, two types of updating procedures are employed. One is the updating procedure which invokes the blocking data communication

function. The other is the updating procedure which calls the nonblocking data communication function. In the case of invoking the blocking data communication function, each magnetic component at boundary cells is communicated after updating each magnetic component. On the other hand, in the case of calling the nonblacking data communication function, the updating process is divided into two parts. One is the boundary part. The other is except the boundary part. Figure 2 shows the pseudo-code of updating procedure for calling the nonblocking data communication function. From Fig. 2, the data communication and the computation is over lapped by the nonblocking data communication function. Therefore, it is efficiently performed than invoking the blocking data communication function.

```
Transient()
{
    COMMUNICATE Hx and Hy of boundary part
between neighboring subdomains
    WHILE Current Time < Ending Time
        COMPUTE Hz of the boundary part
        COMMUNICATE Hz of boundary part between
        neighboring subdomains
        COMPUTE Hz except boundary part
        COMPUTE Ez except boundary part
        WAIT for completion of Hx and Hy of boundary
        part communicate
        COMPUTE Ez of boundary part
        COMPUTE Ex and Ey except boundary part
        WAIT for completion of Hz of boundary part
        communicate
        COMPUTE Ex and Ey of boundary part
        COMPUTE Hx and Hy of boundary part
        COMMUNICATE Hx and Hy of boundary part
        between neighboring subdomains
    ENDWHILE
}
```

Fig. 2. Pseudo-code of the proposed method.

IV. NUMERICAL RESULTS

First, in order to estimate the accuracy of the proposed method, the simulation has been performed for multi conductor transmission lines illustrated in Fig. 3. Each transmission line is terminated with the resistor (100 Ω). The voltage source is appended to the near end of the trace2. A pulse excitation with 0.5×10^{-9} sec rise/fall time, a width 4×10^{-9} sec, a period 1×10^{-8} sec, and an amplitude 3.3V was used. Mur's first order absorbing boundary condition is used for the absorbing boundary condition. The computational domain consists of $46\times40\times50$ cells and discretized by nonuniform meshes. The minimum cell sizes are Δx=0.2mm, Δy=1 mm, Δz=0.01 mm, respectively. The time step size is 3.33×10^{-14} sec in the FDTD method and 6.53×10^{-13} sec

in the HIE-FDTD method. All of the simulations are performed by PC cluster, which is composed of two PCs. These PCs are connected by the gigabit Ethernet. Each PC has an Intel Xeon E5-2650 2GHz and four GPU boards, which are Tesla C2075. Tesla C2075 is one of the GPU boards for the high performance computing. In this simulation, the Intel Xeon E5-2650 was used for the FDTD method and the HIE-FDTD method and eight GPU boards were used for the proposed method. Open MPI is employed for MPI library. Figure 4 shows waveform results of the far end of the trace2 and the trace3. From Fig. 4, the waveform results show good agreement between these methods.

Next, in order to verify the efficiency of the proposed method, the large scale problem has been performed. The number of cells is $1270\times1270\times102$ cells. The minimum cell sizes are Δx=Δy=1 mm, Δz=0.01 mm. Mur's first order absorbing boundary condition is adopted. The time step size for the FDTD method is 3.33×10^{-14} sec. That for the HIE-FDTD method is 1.89×10^{-12} sec. Table 1 shows the simulation results, which are the elapsed time and the speed up ratio, by the FDTD method, the multi-GPU FDTD method, the HIE-FDTD method and the proposed method. The proposed method is performed with single precision floating point and double precision floating point. The peak performance of the Tesla C2075 by using the single precision floating point is two times faster than the peak performance by using the double precision floating point. In the bracket, the communication time is described. From Table 1, the proposed method is about 4000 times faster than the FDTD method by using the 8 GPUs with single precision floating point.
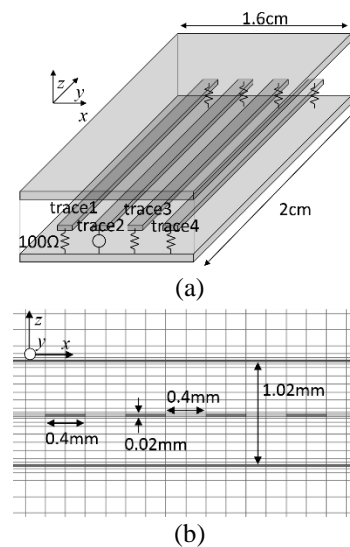

Fig. 3. Example printed circuit board: (a) overhead view of the example circuit, and (b) cross section view of the example circuit.
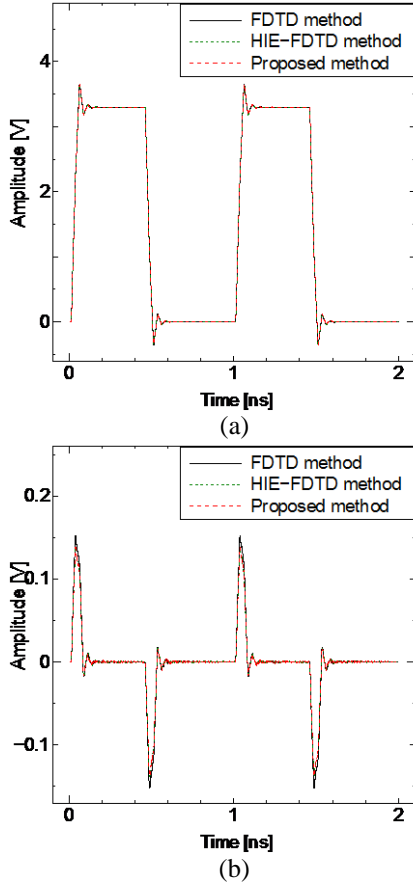
(a)

(b)

Fig. 4. Waveform results: (a) far end of the trace2, and (b) far end of the trace3.

Table 1: Elapsed time and speed up ratio

	Elapsed Time (sec)	Speed up Ratio (vs FDTD Method)
FDTD method	499395.0	1.00
Multi-GPU FDTD method	6189.88	80.68
HIE-FDTD method	15110.43	33.05
Proposed method (double precision and blocking)	287.39 [221.23]	1737.69
Proposed method (single precision and blocking)	145.62 [117.33]	3429.44
Proposed method (double precision and nonblocking)	244.86	2039.51
Proposed method (single precision and nonblocking)	126.64	3943.42

V. CONCLUSION

The multi-GPU HIE-FDTD method proposed for the efficient simulation of the large-scale electromagnetic problems including thin structures. In the case of suitable given objects for the HIE-FDTD method, it has been verified that the proposed method is more than about 4000 times faster than the conventional FDTD method in the case of using 8 GPUs with single precision floating point.

ACKNOWLEDGMENT

REFERENCES

[1] A. Taflove and S. C. Hagness, *Computational Electrodynamics: The Finite-Difference Time-Domain Method*. Artech House, Inc., Norwood, 2005.

[2] Y. Yang, R. S. Chen, and E. K. N. Yung, "The unconditionally stable Crank-Nicolson FDTD method for three-dimensional Maxwell's equations," *Microw. Opt. Tech. Lett.*, vol. 48, no. 8, pp. 1619-1622, May 2006.

[3] J. Chen and J. Wang, "A three-dimensional semi-implicit FDTD scheme for calculation of shielding effectiveness of enclosure with thin slots," *IEEE Trans. Electromagn. Compat.*, vol. 49, no. 2, pp. 354-360, May 2007.

[4] M. Unno and H. Asai, "HIE-FDTD method for hybrid system with lumped elements and conductive media," *IEEE Microw. Wireless Compon. Lett.*, vol. 21, no. 9, pp. 453-455, Sep. 2011.

[5] H. Muraoka, Y. Inoue, T. Sekine, and H. Asai, "A hybrid implicit-explicit and conformal (HIE/C) FDTD method for efficient electromagnetic simulation of nonorthogonally aligned thin structures," *IEEE Trans. Electromagn. Compat.*, vol. 57, no. 3, pp. 505-512, June 2015.

[6] Y. Inoue and H. Asai, "Fast fullwave simulation based on parallel distributed HIE-FDTD method," *IEEE APMC 2012*, Kaohsiung, Taiwan, pp. 1253-1255, Dec. 2012.

[7] M. Unno, S. Aono, and H. Asai, "GPU-based massively parallel 3-D HIE-FDTD method for high-speed electromagnetic field simulation," *IEEE Trans. Electromagn. Compat.*, vol. 54, no. 4, pp. 912-921, Aug. 2012.

On The Beam Forming Characteristics of Linear Array Using Nature Inspired Computing Techniques

Suraya Mubeen [1*], A. M. Prasad [2], and A. Jhansi Rani [3]

[1] Department of ECE, JNTU Kakinada, Kakinada-533001, Andhra Pradesh, India
Suraya418@gmail.com

[2] Department of ECE, JNTU Kakinada, Kakinada-533001, Andhra Pradesh, India
a_malli65@yahoo.com

[3] Department of ECE, VR. Siddhartha College of Engineering, Vijayawada-520001, Andhra Pradesh, India
jhansi9rani@gmail.com

Abstract ─ Beamforming is a serious problem in wireless communication. Many deterministic and numerical techniques are proposed to achieve beamforming. However, the application of evolutionary computing techniques produced better results over many existing conventional methods. In this paper, one such attempt of applying a novel nature-inspired technique known as Firefly algorithm (FFA) to demonstrate beamforming in linear arrays and compared with particle swarm optimization (PSO). The desired objectives of the synthesis process are defined as sidelobe level (SLL) suppression and null positioning. The optimal set of amplitude distribution for the elements in the linear array is obtained using the technique in order to achieve the desired objectives. The results are evaluated in terms of radiation pattern plots.

Index Terms ─ Antenna array, Firefly algorithm, particle swarm optimization.

I. INTRODUCTION

Beamforming, on one hand it involves in observing deep nulls in the direction of the undesired signal and, on the other hand, it involves in positioning the main beam in the direction of the desired signal. Antenna array which have perfect control of radiation characteristics like sidelobe level (SLL), control on beam width (BW) along with beam steering (BS) capabilities have emerged as the best candidates for beamforming applications [1]. Linear arrays are the simplest form of array antenna geometry. All the elements of the array are oriented along a straight line defined in terms of array length and number of elements. These linear arrays are the best candidates for beamforming applications with efficient sidelobe level suppression [2-4] and null control [5] characteristics. These arrays are capable of interference suppression through beam-forming technique. Nulls are located in the direction of arrival (DOA) of the interference signal while the main beam is steered to the DOA of the desired signal in order achieve the above said characteristics. Many conventional techniques are proposed to solve the problem of beamforming. Unfortunately, these are time-consuming as well as provide poor performance. In order to overcome these hurdles, in the recent past several evolutionary techniques are proposed [3-6]. These techniques are quite efficient and often express the supremacy over traditional techniques.

In this paper, novel nature-inspired metaheuristic evolutionary computing algorithm known as Firefly algorithm (FFA) [6, 7] is employed to achieve the specified two objectives of SLL suppression and null position in linear arrays. The rest of the paper is organized as follows. Sections 2 and 3 are dedicated to description of the problems statement and formulation of cost function. Implementation of the algorithm for the proposed problem is given in Section 3. The case wise presentation of results is given in Section 4 which is followed by overall conclusion in Section 5.

II. PROBLEM STATEMENT

The problem statement can be considered as to determine the amplitude distribution of linear array which produces the desired radiation pattern. The geometry of the broadside linear array of N equi-spaced isotropic elements with symmetric excitation positioned along Z-axis is shown in Fig. 1.

Mathematically, the array factor can be stated as:

$$E(\theta) = \sum_{n=1}^{N} I_n \exp\left[j(n-1)(kd\cos\theta + \beta) \right]. \qquad (1)$$

Here, x_n is Amplitude of excitation of the nth element of the array, k is wave number, d is spacing between the elements ($\lambda/2$), $\beta = kd\cos\theta_d$ and θ_d = Scan angle and

N = Number of elements in the array. Progressive phase β is zero as pattern maxima is directed towards $\theta_d = 90^0$ for broadside linear array. However, in scanned array pattern maxima is oriented at an angle θ_d.
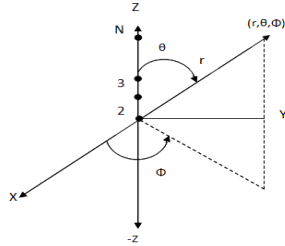

Fig.1. Geometry of broadside linear array.

III. COST FUNCTION FORMULATION

The two objectives considered in this work corresponding to linear array synthesis are SLL suppression and null positioning. The cost functions corresponding to these two objectives are given as below.

For SLL optimization:

$f1= \max[E_{\Theta=FN \text{ to } \pi/2}]$

$$cf1 \quad = |f1\text{-}40| \quad \text{if } f1<40$$
$$= 0 \quad \text{otherwise,} \quad (2)$$

where cf1 is cost function and FN is first null.

For null positioning:

$E_{nmax}= \max[E_{(\Theta=n1)}, E_{(\Theta=n2)}, E_{(\Theta=n3)}]$

$$cf2 \quad = |E_{nmax}\text{-}60| \quad \text{if } E_{nmax}<60$$
$$= 0 \quad \text{otherwise,} \quad (3)$$

where n1, n2 and n3 are desired null positions in degrees.

IV. OPTIMIZATION ALGORITHMS

A. Firefly algorithm

Genesis

FFA is proposed by Yang [6]. Like several other swarm intelligence based algorithms, FFA is also inspired by the natural behaviour of fireflies and the phenomenon of bioluminescent communication. Firefly species produce short and rhythmic flashes. These flashes of light play a vital role in bioluminescent communication. Mostly the flash pattern is unique for every species in terms of frequency, colour and time for which the flash of light is generated. These patterned flashes are used by the fireflies to attract other fireflies for mating. According to inverse square law the intensity of light at a certain distance r from the light source. Which means that the intensity of the light I goes on decreasing as the distance r increase in terms of $I \alpha 1/r^2$. In addition to this another phenomenon known as absorption, the light becomes weaker as it travels along the distance. Due to these factors when combined make most fireflies visible at a limited distance, normally to a few hundred meters at night, which is quite enough for fireflies to communicate with each other. The following

are the rules that are used to describe the structure of the FFA:
1) All fireflies are unisex so that one firefly will be attracted to other fireflies regardless of their sex.
2) Attractiveness depends on the brightness of the flashed light. Accordingly, firefly with brighter flash is supposed to have good attractiveness than the firefly with less brightness.
3) Brightness is defined by the objective function or the landscape of the species.

Algorithm Construction

The construction of the algorithm considers several base rules which mimic the actual behaviour of the fireflies. It is interesting to note that these fireflies (FF) are unisex and are capable of drawing the attention of any companion FF. The degree of attraction is referred as individual's brightness. The brightness or light intensity of a firefly is influenced by the landscape of fitness/cost function. The basic steps of the FA can be summarized as the pseudo code in Fig. 2.
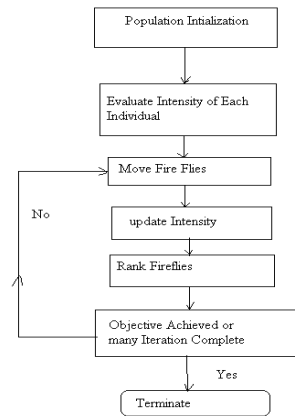
Fig. 2. Flowchart of Firefly algorithm.

Implementation of FFA for LA Synthesis

The adoption of the algorithm for the array synthesis problem starts with population initialization. Every individual FF in the population is considered as a respective array and the amplitude distribution for each array in the population is randomly generated. If P individual FF are considered, then the population is a vector of size 1xK. However, each individual FF is again a vector of size equal to the number of elements in each array. For example, the ith FF is given as follows:

$$x_i = [x_1, x_2, x_3 \ldots x_K], \quad (4)$$

Similarly, the population matrix is given as:

$$X = \begin{bmatrix} x_1 \\ x_2 \\ . \\ . \\ . \\ x_P \end{bmatrix} = \begin{bmatrix} x_{11} \, x_{12} \ldots x_{1K} \\ x_{21} \, x_{22} \ldots x_{2K} \\ . \quad . \quad . \quad . \quad . \quad . \\ . \quad . \quad . \quad . \quad . \quad . \\ . \quad . \quad . \quad . \quad . \quad . \\ x_{P1} \, x_{P2} \ldots x_{PK} \end{bmatrix}. \quad (5)$$

After the population initialization the immediate step if FF evaluation, which means to evaluate the fitness of the FF and the corresponding amplitude distribution:

$$I_i = ObjFunc\ (x_i).\qquad(6)$$

The calculation of the attractiveness of an FF is given by:

$$\beta(r) = \beta_o * \exp\left(-\gamma r_{ij}^2\right).\qquad(7)$$

Here, r is the distance between any two fireflies, β_o is the initial attractiveness at r=0 and γ is the absorption coefficient which controls the decrease of the light intensity. The distance between any two fireflies i and j at x_i and x_j respectively can be defined as a Cartesian distance 'r_{ij}' using the following equation:

$$r_{ij} = \left|x_i - x_j\right| = \sqrt{\sum_{k=1}^{n}\left(x_{ij} - x_{jk}\right)^2}.\qquad(8)$$

The movement of a firefly 'i', which is attracted by a more attractive (i.e., brighter) firefly 'j' is given by the following equation:

$$x_i = x_i + \beta_o * \exp\left(-\gamma r_{ij}^2\right) * \left(x_j - x_i\right) + \alpha * \left(rand - 1/2\right).\qquad(9)$$

B. Particle swarm optimization

PSO is developed by Kennedy and Eberhart and applied to several engineering problems. PSO belongs to the class of population based optimization technique which is inspired by social behaviour of bird flocking or fish schooling. The population in PSO constitutes particles. Each particle refers to the position in the search space. Every particle is subjected to move around the search space in search of the target and updates in every iteration. The displacement is governed by both velocity and position of the particle. The velocity depends on personal best position and the global best position of the particle. Personal best refers to the best position of the particle from the displacement history of the particle, while the global best refers to the best position among the flock in the current iteration. This is mathematically modelled using the following equation:

$$v_{k+1}^i = wv_k^i + \frac{\beta * rand(p^i - x_k^i)}{\Delta t} + \gamma * \frac{rand(p_k^g - x_k^i)}{\Delta t},\qquad(10)$$

$$x_{k+1}^i = x_k^i + \Delta t * v_{k+1}^i$$

v_k^i = velocity of ith particle in the kth iteration, p_k^i = personal best of ith particle in the kth iteration, p_k^g = global best in the kth iteration, w = inertial weight, β and α are cognitive and communal coefficients. In this case both are equal to 0.5.

V. RESULTS AND DISCUSSION

The simulation based experimentation is carried out for two different objectives. As discussed earlier, one of the objective is to obtain very low SLL than uniform distribution under scanned and un-scanned beam conditions. Un-scanned beam refers to position of main

beam at 0^0 while scanned beam refers to main beam positioned at 30^0. In both the cases, the SLL is maintained well below the level observed with uniform distribution in which all the elements in the array are uniformly excited. Suppressed SLL to -40 dB is achieved by applying FFA to obtain non-uniform amplitude distribution. The results pertaining to this discussion are mentioned in Fig. 3. The pattern with dashed line refers to uniform distribution where the SLL is -13 dB while the solid line and dotted line refers to un-scanned and scanned radiation patterns with non-uniform distribution using FFA respectively.
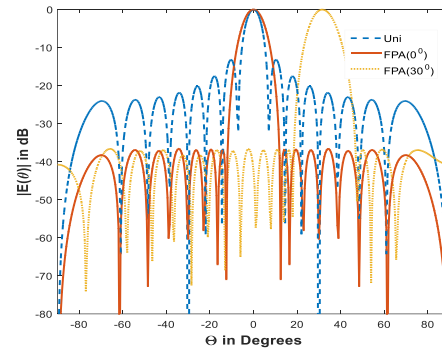
Fig. 3. SLL suppression in unscanned and scanned patterns.

As a second objective the problem of null positioning is considered for both unscanned and scanned beams. As a frst case in this objective, multiple nulls with fixed main position at 0^0 using both PSO and FFA is considered. The DOA of the interference signals are -25^0, 40^0 and 60^0. This is demonstrated in terms of radiation pattern in Fig. 4 (a) and the corresponding convergence plots in Fig. 4 (b). The amplitude distribution obtained using the FFA and PSO is given in the corresponding column of Table 1. As a second case beamsteering along with three nulls is considered. This is shown in the Fig. 5 (a), where the main beam is scanned to an angle of 30^0 which is considered as the DOA of the desired signal. The respective convergence plot is as shown in Fig. 5 (b). The corresponding amplitude distribution using PSO and FFA are as given in Table 1.

(a)

(b)
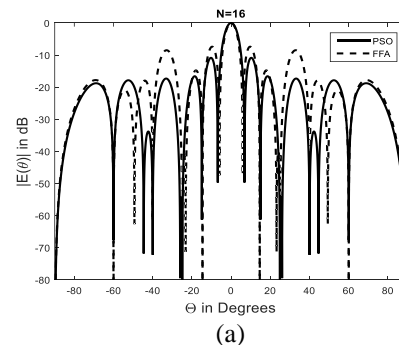
Fig. 4. (a) Radiation pattern with nulls at -25^0, 40^0 and 60^0 while main beam positioned at 0^0, and (b) convergence plot for PSO and FFA.

(a)

(b)

Fig. 5. (a) Radiation pattern with nulls at -25^0, 40^0 and 60^0 while main beam positioned at 25^0, and (b) convergence plot for PSO and FFA.

Table 1: Amplitude distribution obtained using FFA for different cases
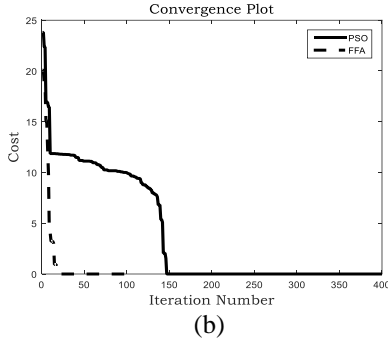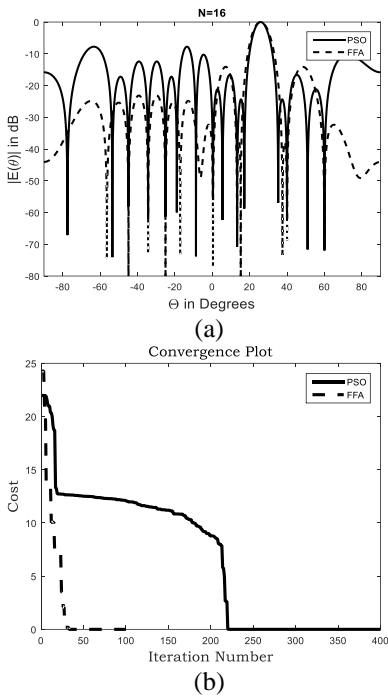
Element Number	Amplitude Distribution					
	Fig. 3		Fig. 4		Fig. 5	
	0^0	30^0	FFA	PSO	FFA	PSO
1&1'	0.76	0.80	0.54	0.76	0.99	0.47
2&2'	0.72	0.76	0.04	0.36	0.86	0.57
3&3'	0.64	0.68	0.11	0.56	0.56	0.14
4&4'	0.53	0.56	0.77	0.75	0.41	0.38
5&5'	0.41	0.44	0.71	0.75	0.48	0.50
6&6'	0.29	0.31	0.46	0.85	0.59	0.13
7&7'	0.18	0.20	0.60	0.55	0.37	0.00
8&8'	0.13	0.15	0.85	0.73	0.38	0.68

VI CONCLUSION

FFA has been successfully applied to array synthesis problems in beamforming applications with multiple objectives. The technique of generating nulls in the desired directions in order to suppress the interference signals is well demonstrated under un-scanned and scanned conditions for beamforming characteristics. FFA has shown its efficiency and simplicity in terms of computation time and complexity. It is evident from the plots that the convergence is achieved earlier with FFA than PSO in all the cases mentioned above. Moreover, the convergence is delayed when the beam is scanned to a certain angle. The technique demonstrated in this paper can easily be extended to any multimodal problems with several constraints.

REFERENCES

[1] S. A. Schelkunoff and H. T. Friis, *Antennas Theory and Practice*. JW & Sons Inc., London, 1952.

[2] K.-K. Yan and Y. Lu, "Sidelobe reduction in array-pattern synthesis using genetic algorithm," *IEEE Transactions on Antennas and Propagation*, vol. 45, no. 7, pp. 1117-1122, 1997.

[3] V. V. S. S. S. Chakravarthy, K. Naveen Babu, S. Suresh, P. Chaya Devi, and P. Mallikarjuna Rao, "Linear array optimization using teaching learning based optimization," *Advances in Intelligent Systems and Computing*, vol. 338, pp. 183-187, 2015.

[4] V. S. S. S. Chakravarthy Vedula, S. R. Chowdary Paladuga, and M. Rao Prithvi, "Synthesis of circular array antenna for sidelobe level and aperture size control using flower pollination algorithm," *International Journal of Antennas and Propagation*, vol. 2015, 2015.

[5] D. Karaboga, K. Guney, and A. Akdagli, "Antenna array pattern nulling by controlling both the amplitude and the phase using modified touring ant colony optimisation algorithm," *Int. Journal of Electronics*, vol. 91, pp. 241-251, 2004.

[6] X.-S. Yang, "Firefly algorithms for multimodal optimization," *in Stochastic Algorithms: Foundations and Applications: 5th International Symposium, SAGA 2009*, Sapporo, Japan, October 26-28, 2009. Proceedings, vol. 5792 of Lecture Notes in Computer Science, pp. 169-178, Springer, Berlin, Germany, 2009.

[7] B. Basu and G. K. Mahanti, "Fire fly and artificial bees colony algorithm for synthesis of scanned and broadside linear array antenna," *Progress In Electromagnetics Research B*, vol. 32, pp. 169-190, 2011.

Bandwidth Enhancement of Dipole Antennas using Parasitic Elements

Garrett Hoch, Atef Elsherbeni, and Payam Nayeri

Department of Electrical Engineering and Computer Science
Colorado School of Mines, Golden, CO 80401, USA
ghoch@mines.edu, aelsherb@mines.edu, pnayeri@mines.edu

Abstract ─ Dipole antennas have a limited bandwidth which restricts their use to narrow bandwidth applications. To improve the functionality of a dipole antenna, additional wire elements can be added to the dipole antenna to increase the impedance matching and bandwidth. A design for a modified dipole antenna is proposed and analyzed using multiple computational electromagnetic software to determine how the additional wire elements affect the input impedance bandwidth and radiated fields. The antenna was then fabricated and tested and compared to simulation results. The modified dipole antenna shows a bandwidth improvement of more than four times, approaching 31%, while maintaining radiation patterns similar to a traditional dipole antenna with slightly higher gain.

Index Terms ─ Antenna, bandwidth, dipole, far field, gain, impedance.

I. INTRODUCTION

Many modern wireless communication systems require antennas that can provide omnidirectional radiation coverage. As such, the classic half-wavelength dipole antenna which provides a uniform omnidirectional coverage is a popular candidate for these systems. The primary drawback for these antennas however is the fact that they operate in a very narrow bandwidth, which limits their application in modern communication systems. In this work, we propose a new configuration for the classic wire dipole antenna, to improve its bandwidth. The presented design is an extension of the configuration proposed in [1]. Four wire elements connected to the wire dipole antenna, as shown in Fig. 1. This essentially creates a quasi-log periodic wire antenna, which minimizes the reactive part of the impedance and improves the matching at the input port of the antenna, resulting in a wider bandwidth. It is shown that by proper tuning of these parasitic wires, the bandwidth of the classic dipole antenna can be increased by more than four times, while maintaining an omnidirectional radiation pattern and slightly increasing the gain.
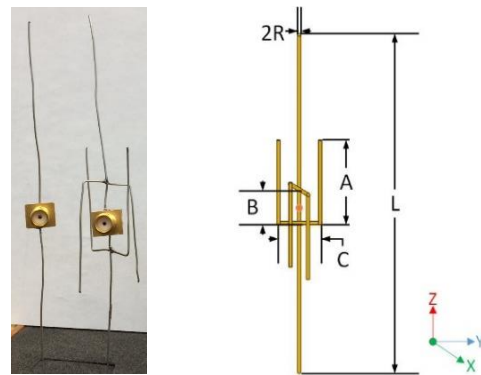

Fig. 1. Dipole antenna models: (left) dipole antenna, (center) modified dipole antenna, and (right) the modified dipole configuration along with design parameters.

II. DESIGN

The initial design for this study is based on the work reported in [1]-[3]. The objective of this study is to further the analysis of the modified dipole antenna configuration shown in Fig. 1. To this end, we compare the performance of the classic wire dipole antenna, with the modified dipole. The analysis is conducted using three computational electromagnetic software: FEKO [4], Ansys HFSS [5], and CEMS [6]. A parametric study for design parameters A, B, and C was conducted using HFSS to understand how the impedance matching of the antenna is effected by the additional wire elements. The parameters R and L were set to 0.2 mm and 128 mm respectively in all used software packages and for the parametric study.

Figures 2-4 show the result of the parametric study. Figure 2 shows the results when varying A with B=10 mm and C = 18 mm. As the length of A increases the real part of the input impedance decreases. The imaginary part of the input impedance curve straightens out as the length of A increases. From the results depicted here, it can be seen that the length of A that corresponds to a real input impedance close to 50 Ohms and an imaginary input impedance around zero, i.e., best matching over a broad frequency range, is for A = 33 mm.
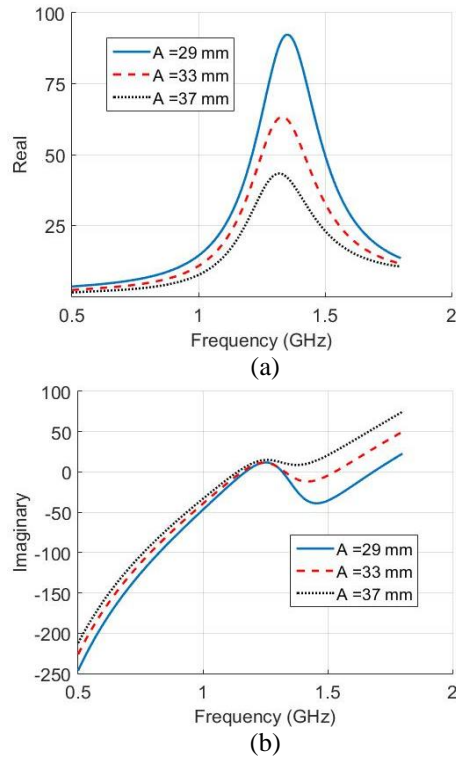
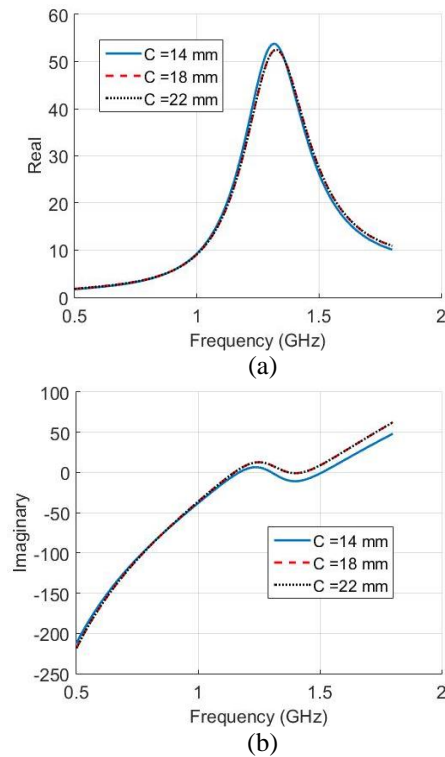Fig. 2. Input impedance versus frequency: (a) real and (b) imaginary.

Fig. 3. Input impedance versus frequency: (a) real and (b) imaginary.

Fig. 4. Input impedance versus frequency: (a) real and (b) imaginary.
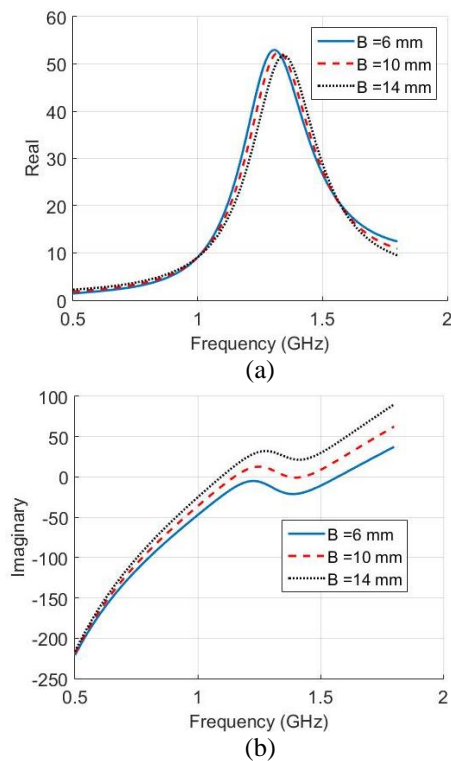
Figures 3 and 4 show the parametric study results for B and C. For the parametric study with B, A = 33 mm and C = 18 mm. For the parametric study with C, A = 33 mm and B = 10 mm. Parameters B and C show similar behaviors as for A, but to less extent. From the study reported here it can be seen that A is the main tuning parameter for impedance matching in this modified dipole configuration. The dimensions of the design parameters were chosen based on the results reported here and are A = 33 mm, B = 10 mm, C = 18 mm, L = 128 mm, and R = 0.2 mm.

III. RESULTS

Using the design dimensions determined in the previous section, the antenna was built and measured as well as simulated using three different computational electromagnetic software, namely: FEKO, Ansys HFSS, and CEMS. For comparison purposes a dipole antenna with the same length as the modified dipole antenna design is also simulated and tested. Figure 5 shows the magnitude of S_{11} versus frequency for the dipole antenna modeled in FEKO and the modified dipole configuration in the three different software packages.

Table 1 shows the original bandwidth and the corresponding improvement for each dipole simulation. The center frequency was determined by finding the -10 dB crossings and calculating the frequency in the middle. For the three modified dipole antennas results in

Fig. 5, the curves are different regions below -10 dB but exhibit a similar behavior. The slight differences can be due to the nature of the three different software packages and how they represent the antenna and the source of excitation. All of the modified dipole models operate at a higher frequency than the traditional dipole antenna and the reflection coefficient curves have two local minimums. The mean improvement in bandwidth is more than four times when compared to the traditional dipole antennas as can be seen in Table 1. Figure 6 shows the measured reflection coefficient for the two models shown in Fig. 1. The curve does not closely resemble what was simulated, but the overall result is similar. The difference is due to the un-professional construction of the antenna, SMA connector, and from the solder joints which was not accounted for in simulations. The bandwidth of the fabricated modified dipole was improved by approximately four times when compared to the fabricated dipole antenna.

To study the effect of the additional wire elements FEKO was used to examine the radiated far-field patterns. The electric field was examined for the modified dipole antenna as well as a dipole antenna operating at the center frequency of the modified dipole antenna.
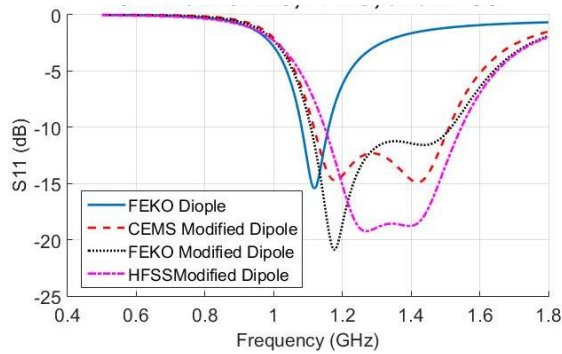
Fig. 5. Reflection coefficient for a regular dipole and the modified dipole using three different software packages.
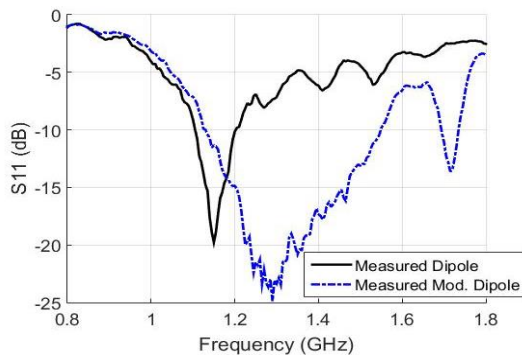
Fig. 6. Measured reflection coefficient for a regular dipole and the modified dipole.

Table 1: Bandwidths of the dipole antennas

Antenna	BW (%)	% BW Improvement
FEKO dipole	7.05	-
CEMS mod. dipole	30.37	23.32
FEKO mod. dipole	31.17	24.12
HFSS mod. dipole	30.35	25.36
Measured dipole	8.24	-
Measured mod. dipole	31.78	23.54

From Figs. 7 and 8 it is observed that the addition of the wire elements does not drastically change the magnitude of E_θ in all three planes across the frequency band. Table 2 shows the gain for both the modified antenna and dipole antenna at three different frequencies within the band.

(a)

(b)

(c)
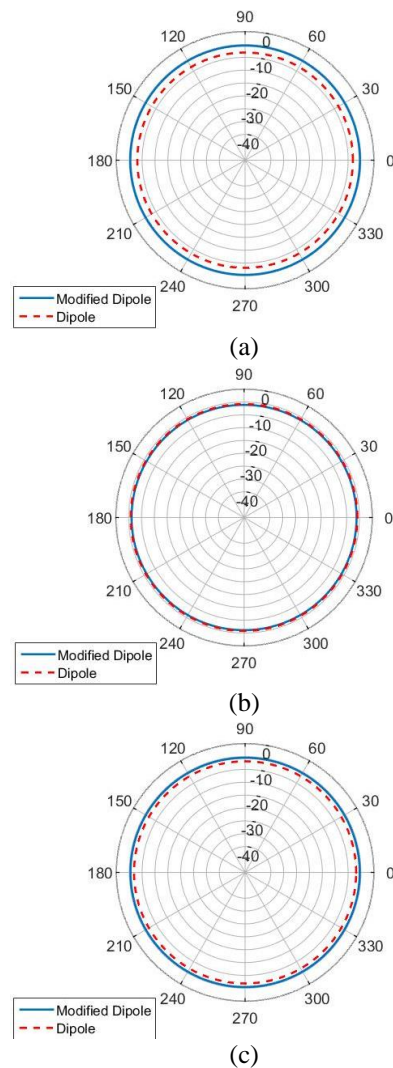
Fig. 7. Electric field in the XY-plane: (a) 1.202 GHz, (b) 1.306 GHz, and (c) 1.410 GHz.

(a)

(b)

(c)
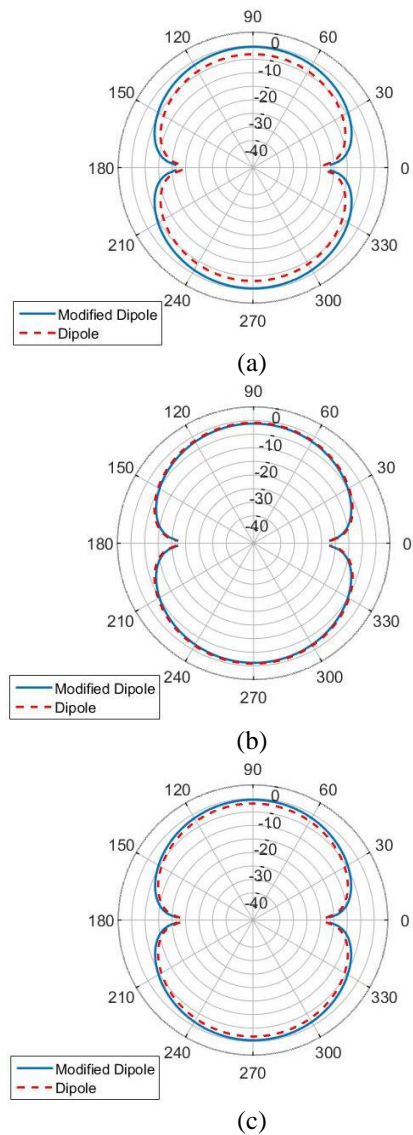
Fig. 8. Electric field in the XZ-plane and YZ-plane: (a) 1.202 GHz, (b) 1.306 GHz, and (c) 1.410 GHz.

Table 2: Realized gain of modified dipole using FEKO

Antenna	Realized Gain (dBi)		
	1.202 GHz	1.31 GHz	1.41 GHz
Modified dipole	2.117	1.92	1.92

Due to the better matching at the input port of the modified dipole, the realized gain of the modified antenna remains around 2 dBi for all three frequencies. The calculated gain of the traditional dipole antenna is 1.94 dBi.

VI. CONCLUSION

A modified dipole antenna configuration is proposed and investigated using three different software packages as well as measured results from built antennas. Additional wire elements are added to a traditional half wave length antenna that enable impedance matching across the band. It is shown numerically and experimentally that this configuration can increase the input impedance bandwidth four times when compared to a traditional dipole antenna. The additional wires do not change the radiation characteristics of the dipole antenna as demonstrated using simulations. The resulting cross polarized electric field component E_ϕ is found to be less than 40 dB of the E_θ component. The gain of the modified antenna is close to 2 dBi over a broad range of frequencies. Based on this study, the proposed modified dipole antenna has the potential to expand the use of dipole antennas in a wide range of applications.

REFERENCES

[1] H.-T. Hsu, J. Rautio, and S.-W. Chang, "Novel planar wideband omni-directional quasi logperiodic antenna," *Asia Pacific Microwave Conference (APMC)*, Suzhou, China, Dec. 2005.

[2] A. Z. Elsherbeni, P. Nayeri, and C.J. Reddy, *Wire Dipole and Monopole Antennas*, in Antenna Analysis and Design Using FEKO Electromagnetic Simulation Software, ACES Series on Computational Electromagnetics and Engineering, SciTech Publishing, an Imprint of IET, Edison, NJ, pp. 36-38, 2014.

[3] G. R. Hoch, P. Nayeri, and A. Z. Elsherbeni, "Bandwidth enhancement of dipole antennas using parasitic elements," *The 31st International Review of Progress in Applied Computational Electromagnetics (ACES 2015)*, 2015.

[4] FEKO, EM Software & Systems Inc., Version 7.0.2. https://www.feko.info/

[5] HFSS, Ansys, Release 16.2.0. http://www.ansys.com

[6] CEMS, EM software package based on A. Elsherbeni and V. Demir, *The Finite-Difference Time-Domain Method for Electromagnetics with Matlab Simulations*, ACES Series on Computational Electromagnetics and Engineering, 2nd Edition, SciTech Publishing, an Imprint from IET, Edison, NJ, 2016.

CUDA-MPI Implementation of Fast Multipole Method on GPU Clusters for Dielectric Objects

Nghia Tran, Tuan Phan, and Ozlem Kilic

Department of Electrical Engineering and Computer Science
The Catholic University of America, Washington, DC, 20064, USA
16tran@cua.edu, 30phan@cua.edu, kilic@cua.edu

Abstract ─This paper investigates the Fast Multipole Method (FMM) for large-scale electromagnetics scattering problems for dielectric objects. The algorithm is implemented on a Graphical Processing Unit (GPU) cluster using CUDA programming and Message Passing Interface (MPI). Its performance is investigated in terms of accuracy, speedup, and scalability. The details of the implementation and the performance achievements are shown and analyzed, demonstrating a scalable parallelization while maintaining a good degree of accuracy.

Index Terms ─ Dielectric objects, Fast Multipole Method (FMM), Graphics Processing Unit (GPU), Message Passing Interface (MPI).

I. INTRODUCTION

Modelling electrically large dielectric objects plays an important role in the research of target identification or the stealth and anti-stealth technology. The excessively heavy requirements of memory and computational resources of such applications pose a challenging problem in the computational electromagnetics community. The past two decades have witnessed many numerical techniques developed to reduce this burden without significant loss of accuracy, including Adaptive Integral Method (AIM) [1], Impedance Matrix Localization (IML) [2], and Fast Multipole Method (FMM) [3]. Among others, FMM is the most suitable technique for large-scale problems in reducing the computational complexity of the conventional technique, Method of Moment (MoM) [4] from $O(N^3)$ to $O(N^{1.5})$, where N denotes the number of unknowns. Some other approaches such as AIM and IML have the complexity of $O(N^{1.5}logN)$ and $O(N^2logN)$, respectively. Many authors have investigated the parallelization of FMM on CPU clusters for solving the dielectric problems [5]. However, to the best of our knowledge, FMM has not been studied for dielectric electromagnetics problems on GPU clusters. Recently our group has applied single-level FMM for perfect electric conductor (PEC) objects [6]-[7], and good speedup factors compared to the CPU implementations

were achieved. However, our previous implementations focused only on PEC objects, which can be limiting for simulating real-life scenarios.

In this paper, we investigate the parallelization of FMM for dielectric electromagnetics structures using the PMCHW formula [8] on a multi-node GPU cluster with CUDA and MPI programing. We demonstrate that the implementation of FMM on GPU clusters outperforms that of the CPU counterpart in terms of speedup and scalability.

The rest of the paper is organized such that Section II provides an overview of FMM for dielectric objects. Section III describes the parallelization of FMM on GPU clusters. Experimental results are discussed in Section IV. Finally, our findings are summarized in the conclusions section.

II. OVERVIEW OF THE FAST MULTIPOLE METHOD FOR DIELECTRIC OBJECTS

In this section, we provide a brief overview to help our discussion on the parallel implementation of FMM for dielectric objects, which is presented in detail in Section III.

FMM solves for the linear equation system created in the form of $ZI = V$, where I represents the unknown currents, V depends on the incident field, and Z is the impedance matrix. The main idea of FMM is the grouping concept as shown in Fig. 1, where N edges in the mesh of a given structure are categorized into M localized groups based on their proximity. According to this approach, two interaction types can be defined: near and far, as depicted in Fig. 1. The Z matrix is split into two components, Z^{near} and Z^{far}, as shown in (1):

$$\sum_{m'}^{M} Z_{mm'} I_{m'} = \sum_{m'}^{M} Z_{mm'}^{near} I_{m'} + \sum_{m'}^{M} Z_{mm'}^{far} I_{m'} = V_m,$$

$$\text{where } Z_{mm'} = \begin{bmatrix} Z_{mm',JJ} & Z_{mm',JM} \\ Z_{mm',MJ} & Z_{mm',MM} \end{bmatrix}, V_m = \begin{bmatrix} E_m \\ H_m \end{bmatrix}, \quad (1)$$

and m and m' are observation and source edges in the mesh, respectively.

The near term comprises of interactions between

spatially close edges, and is computed and stored in a similar manner to MoM [4]. For dielectric objects, PMCHW formula [8] is used in this paper and the four components of Z^{near} is shown in (2)-(4):

$$Z_{mm',JJ}^{near} = \frac{j\omega\mu}{4\pi}\int_S f_{m'}(\boldsymbol{r'})\int_S G_1(\boldsymbol{r},\boldsymbol{r'})\bullet f_m(\boldsymbol{r})dS'dS$$
$$+\frac{1}{4\pi j\omega\varepsilon}\int_S (\nabla\bullet f_m(\boldsymbol{r}))\int_S G_1(\boldsymbol{r},\boldsymbol{r'})\bullet(\nabla\bullet f_{m'}(\boldsymbol{r'}))dS'dS \quad , (2)$$

$$Z_{mm',MM}^{near} = \frac{j\omega\varepsilon\eta^2}{4\pi}\int_S f_{m'}(\boldsymbol{r'})\int_S G_1(\boldsymbol{r},\boldsymbol{r'})\bullet f_m(\boldsymbol{r})dS'dS$$
$$+\frac{\eta^2}{4\pi j\omega\mu}\int_S (\nabla\bullet f_m(\boldsymbol{r}))\int_S G_1(\boldsymbol{r},\boldsymbol{r'})\bullet(\nabla\bullet f_{m'}(\boldsymbol{r'}))dS'dS \quad , (3)$$

$$Z_{mm',MJ}^{near} = \frac{-\eta}{4\pi}\int_S f_{m'}(\boldsymbol{r'})\int_S \nabla G_1(\boldsymbol{r},\boldsymbol{r'})\times f_m(\boldsymbol{r})dS'dS$$
$$= -Z_{mm',JM}^{near} \quad . (4)$$

The interactions between the remaining edges that are spatially far from each other constitute the far term. The advantage of separating the Z matrix into two components is that the Z^{far} matrix does not need to be computed and stored ahead of time. Instead it is factorized into the radiation, T^E/T^{ED}, receive, R^E, and translation functions, T_L. Equations (5)-(11) depict these functions based on PMCHW formula:

$$Z_{mm',JJ}^{far} = \frac{\omega\mu k}{16\pi^2}\int d^2\hat{\boldsymbol{k}} \; \boldsymbol{T}_{r_{im}}^E(\hat{\boldsymbol{k}})T_L(k,\hat{\boldsymbol{k}},\boldsymbol{r}_{ii'})\bullet R_{m'i'}^E(\hat{\boldsymbol{k}}), \quad (5)$$

$$Z_{mm',MM}^{far} = \frac{\omega\varepsilon k\eta^2}{16\pi^2}\int d^2\hat{\boldsymbol{k}} \; \boldsymbol{T}_{r_{im}}^E(\hat{\boldsymbol{k}})T_L(k,\hat{\boldsymbol{k}},\boldsymbol{r}_{ii'})\bullet R_{m'i'}^E(\hat{\boldsymbol{k}}), \quad (6)$$

$$Z_{mm',MJ}^{far} = \frac{k\eta}{16\pi^2}\int d^2\hat{\boldsymbol{k}} \; \boldsymbol{T}_{r_{im}}^{ED}{}'(\hat{\boldsymbol{k}})T_L(k,\hat{\boldsymbol{k}},\boldsymbol{r}_{ii'})\bullet R_{m'i'}^E(\hat{\boldsymbol{k}})$$
$$= -Z_{mm',JM}^{far} \quad , (7)$$

where

$$T_{r_{im}}^E = \int_S (\boldsymbol{I} - \hat{\boldsymbol{k}}\hat{\boldsymbol{k}})\bullet f_m(\boldsymbol{r}_{im})e^{-j\boldsymbol{k}\bullet\boldsymbol{r}_{im}}dS , \quad (8)$$

$$T_{r_{im}}^{ED} = \int_S \hat{\boldsymbol{k}}\times f_m(\boldsymbol{r}_{im})e^{-j\boldsymbol{k}\bullet\boldsymbol{r}_{im}}dS , \quad (9)$$

$$R_{r_{m'i'}}^E = \int_S f_{m'}(\boldsymbol{r}_{m'i'})e^{-j\boldsymbol{k}\bullet\boldsymbol{r}_{m'i'}}dS , \quad (10)$$

$$T_L = \sum_{l=0}^{L}(-j)^l(2l+1)h_l^{(2)}(\boldsymbol{k}\bullet\boldsymbol{r}_{ii'})P_l(\hat{\boldsymbol{k}}\bullet\boldsymbol{r}_{ii'}) . \quad (11)$$
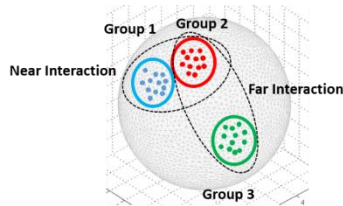

Fig. 1. FMM grouping concept.

In the equations above, the prime syntax denotes the source points, and i refers to the groups in the mesh. The unit vector $\hat{k}$ denotes the K possible field directions in k space, $f(\mathbf{r})$ denotes the basic functions, $h_l^2(r)$ is the spherical Hankel function of the second kind, and $P_l(r)$

is the Legendre polynomial.

III. PARALLELIZATION OF FMM ON GPU CLUSTERS

The platform utilized in our FMM implementation is a GPU cluster, which consists of 13 computing nodes. Each node has a dual 6-core 2.66 GHz Intel Xeon processor, 48 GB RAM along with one NVidia Tesla M2090 GPU running at 1.3 GHz with 6 GB memory. The nodes are interconnected through the infiniBand interconnection. The cluster populates CUDA v7.0 and MVAPICH2 v1.8.1. (a well-known implementation of Message Passing Interface (MPI)). Two parallel programming approaches of CUDA and MPI are combined to provide the use of GPU programming across the cluster.

In this section, we provide an overview of our implementation on GPU. Figure 2 shows the main blocks which consist of pre-processing, processing and post-processing, where processes which utilize GPU are shown in solid green line, and CPU based operations are shown in dashed black line. The geometry mesh data resulting from the pre-processing step is transferred to the GPU memory, and the entire computation is performed on the GPU. The user defined results such as radar cross section, scattered fields are post-processed on the CPU.

The parallelization of the processing step in GPU cluster implementation is performed at two levels: (i) among computing nodes using MPI library, and (ii) within GPU per node using CUDA programming model. The workload of the computational task is equally distributed among the computing nodes, and the inter-node communication is minimized. This is achieved by uniformly distributing the total number of groups, M, among the n computing nodes. We only present the far interactions in this paper, since the near field and V vector calculations implementations can be found in [6]-[7].

A. Far interactions calculations

This task comprises of three calculations: radiation, and receive functions and translation matrices. The first step in the far interaction calculations is the calculation of the radiation, T^E/T^{ED} and receive, R^E, functions for Z^{far} matrix as seen in Fig. 2. It is worth noting that the radiation and the receive functions as well as the translation matrix have to be evaluated at all K directions for the unit sphere integration. The computational workload is distributed across the nodes using the group-based partitioning scheme such that M_{node} groups are allocated for each node. K evaluations for the radiation and receive functions are required for each node. The threads are grouped into blocks such that each block of threads performs N_{group} radiation/receive functions at a given group, then a total number of blocks per node

equal to $M_{node}K$.

The second task for far interactions is the calculation of the translation matrix, T_L. Similar to the calculation of the radiation/receive functions, the translation matrix has to be evaluated at all K directions for the unit sphere integration. The workload for the T_L calculations is also distributed across the nodes following the group-based technique. Each CUDA block is assigned to compute one sparse row of the T_L matrix for a given direction, and each thread computes one element in that row.
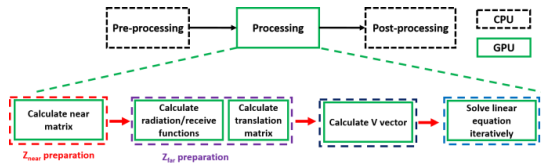

Fig. 2. FMM flow chart implementation.

B. Fast matrix-vector multiplication

The processing step is followed by solving linear equation iteratively. In this paper, the Biconjugate Gradient Stabilized method (BICGSTAB) is employed where each iteration involves the Matrix-Vector Multiplication (MVM). In this part, MVM of the far interactions are focused and the calculation of $Z_{far}I$ comprises of aggregation, translation, and disaggregation. By using the group-based scheme, the inter-node communication is required only at two steps: (i) at the beginning of the MVM to exchange the estimated values for the unknowns among the nodes, and (ii) after the aggregation step and before performing the translation step in order to update all nodes with the aggregated fields. The flow chart of the implementation is shown in Fig. 3.
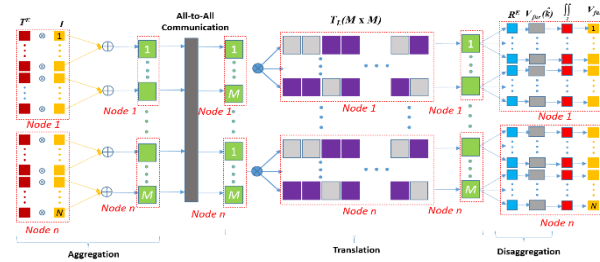

Fig. 3. Far-interaction flow chart implementation.

Each node already calculates its own portion of the radiation/receive functions, and the translation matrix. In the aggregation stage, each unknown is multiplied with its corresponding radiation functions, and is summed in each group. The aggregation stage can be performed independently for all K directions, and thus can be performed concurrently with minimal need for inter-node communication. An all-to-all communication is employed by each node to broadcast the aggregated

fields to all other nodes. Then the received fields at each direction are determined by multiplying the aggregated fields with the translation matrix. Next the received fields are multiplied with the receive functions to obtain the field for each group at a given direction. Finally, an integration over the K directions of the unit sphere is performed to calculate the fields at each observed edge. The far MVM is incorporated with the near MVM to get the full ZI.

IV. EXPERIMENTAL RESULTS

A. Accuracy

To validate the accuracy of FMM implementations on GPU clusters, we calculate the radar cross section (RCS) of a 14λ diameter (254,274 unknowns) sphere with permittivity $\varepsilon = 4 - 0.1j$. It is illuminated by a normally incident 1 GHz x-polarized plane wave. The RCS based on our GPU implementation is compared to the results of the analytical Mie scattering. Figure 4 shows that the GPU and Mie solutions achieve a good agreement.
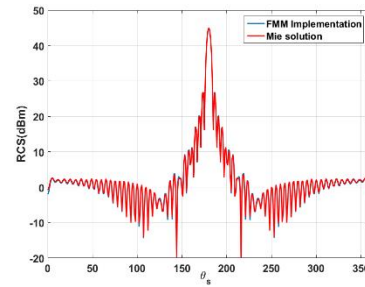

Fig. 4. RCS of a 14λ diameter dielectric sphere

B. Implementation performance on GPU cluster

Our GPU implementations is evaluated using the fixed-workload model in the first experiment. We choose a sphere diameter of 12.4λ, ~ 200K unknowns for the fixed problem size such that it demands the use of at least 8 nodes to satisfy the required memory. Speedup and scalability are used to evaluate the GPU implementations. The speed up is defined as the ratio of time required by multi-node GPU implementation with respect to the 8-node CPU implementation. Scalability is the normalized speedup of multiple nodes in reference to the speedup of 8 nodes. In our analysis, we consider the total executional time and computational time. Figure 5 shows the speedup factors and the measured time of two cases. It is observed that the computation time achieves a speedup factor of 171.5 on 8 nodes, and takes 158.9 seconds. Due to the nature of the fixed workload model, each node carries less workload when the number of computing nodes increase. Therefore the computation time decreases linearly with the increase in nodes (274.5 seconds for computation and 229.9 seconds for total). The slightly smaller speedup factors for the total time as compared to the computation time are due to the inter-node

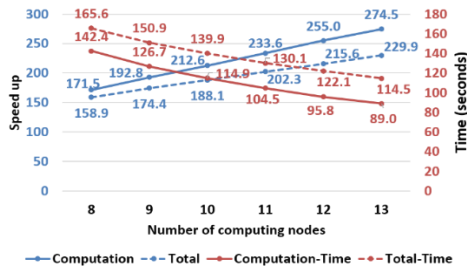communications for transferring the data in our GPU implementations.

Fig. 5. Speedup analysis for the fixed-workload model (vs. 8 nodes CPU implementation, 100 iterations) Computational CPU exec time = 24,421 sec, total CPU exec time = 26,315 sec.

It is observed in Fig. 6 that both computation and total time scale closely to the theoretical linear expectation for the fixed workload problem. This good scalability demonstrates that our implementation has efficiently parallelized the algorithm and reduced the communication overhead.

The second experiment is the fixed-time model. As the problem size is increased, the number of nodes also increases, so that the GPU memory in each node is fully utilized. Our GPU implementation can process a maximum problem size of 254K unknowns with a speedup factor of 169.6 for the computation, and 137 for total execution time as shown in Fig. 7.
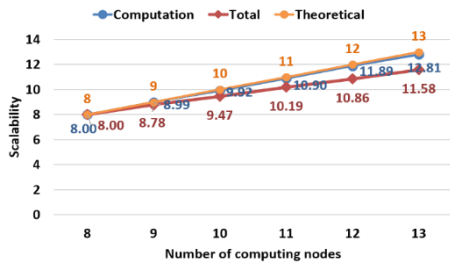

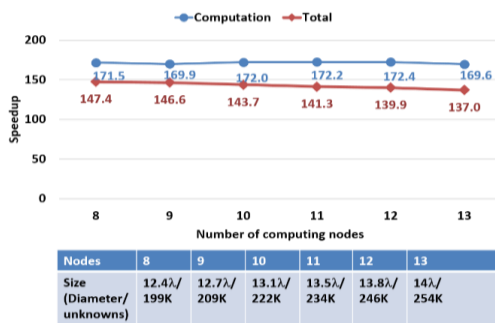Fig. 6. Scalability analysis for the fixed-workload model.

Nodes	8	9	10	11	12	13
Size (Diameter/ unknowns)	12.4λ/ 199K	12.7λ/ 209K	13.1λ/ 222K	13.5λ/ 234K	13.8λ/ 246K	14λ/ 254K

Fig. 7. Speedup analysis when the number of nodes increases along with problem size increases (vs. multi-node CPU, 100 iterations).

VI. CONCLUSION

In this paper, the GPU implementation of FMM for dielectric electromagnetic scattering problems using our 13-node GPU cluster is demonstrated. The maximum problem size is determined by the available on-board GPU memory. For the same degree of accuracy, the GPU implementation outperforms the CPU implementation. Moreover, the GPU implementation has a good scalability as the number of computing nodes increases.

REFERENCES

[1] E. Bleszynski, M. Bleszynski, and T. Jaroszewicz, "AIM: Adaptive integral method for solving large-scale electromagnetic scattering and radiation problems," *Radio Science*, vol. 31, no. 5, pp. 1225-1251, 1996.

[2] F. X. Canning, "The impedance matrix localization (IML) method for moment-method calculations," *IEEE Antennas Propagat. Mag.*, vol. 32, no. 5, pp. 18-30, 1990.

[3] R. Coifman, V. Rokhlin, and S. Wandzura, "The fast multipole method for the wave equation: A pedestrian prescription," *IEEE Antennas Propagat. Mag.*, vol. 35, no. 3, pp. 7-12, June 1993.

[4] S. M. Rao, D. R. Wilton, and A. W. Glisson, "Electromagnetic scattering by surfaces of arbitrary shape," *IEEE Trans. Antennas Propag.*, vol. AP-30, no. 3. pp. 409-418, May 1982.

[5] O. Ergul and L. Gurel, "Efficient parallelization of the multilevel fast multipole algorithm for the solution of large-scale scattering problems," *IEEE Trans. Antennas Propag.*, vol. 56, no. 8, pp. 2335-2345, August 2008.

[6] Q. M. Nguyen, V. Dang, O. Kilic, and E. El-Araby, "Parallelizing fast multipole method for large-scale electromagnetic problems using GPU clusters," *Antennas and Wireless Propagation Letters, IEEE*, vol. 12, pp. 868-871, 2013.

[7] V. Dang, Q. Nguyen, and O. Kilic, "Fast multipole method for large-scale electromagnetic scattering problems on GPU cluster and FPGA-accelerated platforms," *Applied Computational Electromagnetics Society Journal*, vol. 28, no. 12, 2013.

[8] X. Q. Sheng, J.-M. Jin, J. Song, W. C. Chew, and C.-C. Lu, "Solution of combined-field integral equation using multilevel fast multipole algorithm for scattering by homogeneous bodies," *IEEE Transactions on Antennas and Propagation*, vol. 46, no. 11, pp. 1718-1726, 1998.

Vector Control of PMSM Take Over by Photovoltaic Source

T. Yuvaraja [1] and K. Ramya [2]

[1] Department of Electrical and Electronics Engineering
Sri Sairam College of Engineering, Bangalore City, Karnataka - 632001, India
yuvarajastr@gmail.com

[2] Department of Electrical and Electronics Engineering
Sri Sairam College of Engineering, Bangalore City, Karnataka - 632001, India
ramyaj.k14@gmail.com

Abstract ─ This article established with the modeling and the field oriented control of permanent magnet synchronous machines, with a focus on their applications in variable speed domain in photovoltaic source.

Index Terms ─ Field oriented control, inverter, permanent magnet synchronous machine, photovoltaic.

I. INTRODUCTION

The permanent magnet synchronous motor (PMSM) has three phases winding on stator represented by the three axes (a, b, c) phase-shifted of 120° with respect to each other (Fig. 1) and has permanent magnets in the rotor ensuring its excitation. Depending on how the magnets are placed, we can distinguish two types of rotors; in the first type, the magnets are mounted on the surface of the rotor with a homogeneous air gap, the motor is called "smooth air gap PMSM" and inductors are independent on the rotor position. In the second type, the magnets are mounted inside the rotor mass and the air gap will vary because of the salience effect. In this case, inductors are highly dependent on the rotor position. Synchronous motors have a remarkable feature; the speed is constant regardless of the load. The field oriented control (FOC) is used for many years. It implements Park transformation which shows, like a separately excited dc machine, the expression of the instantaneous torque as a product of magnetic flux and current. In addition, there is the possibility to reduce the oscillations for a desired torque, to save energy delivered, to reduce the current harmonics and to improve power factor. When the motor model used is correct, the FOC works well [1-6].

II. DYNAMIC MODEL OF SYNCHRONOUS MOTOR

A. Mathematical model of PMSM

The dynamic model of a permanent magnet synchronous motor with rotor reference frame can be described by the equations below, considering the conditions of non-saturation of the magnetic circuit and the magneto motive force MMF is a sinusoidal distribution created by the stator windings:

$$[V_{abc}] = [RI_{abc}] + \frac{d(\varphi_{abc})}{dt}, \qquad (1)$$

$$[V_{abc}] = \begin{bmatrix} V_a \\ V_b \\ V_c \end{bmatrix}; [i_{abc}] = \begin{bmatrix} i_a \\ i_b \\ i_c \end{bmatrix}; [\varphi_{abc}] = \begin{bmatrix} \varphi_a \\ \varphi_b \\ \varphi_c \end{bmatrix} [R] = \begin{bmatrix} R & 0 & 0 \\ 0 & R & 0 \\ 0 & 0 & R \end{bmatrix}, \qquad (2)$$

with $V_{abc}, i_{abc}, \phi_{abc}$ representing the stator phases' voltages, the stator phases' currents and the total flux produced by the stator currents. R indicates the resistance of a stator phase.
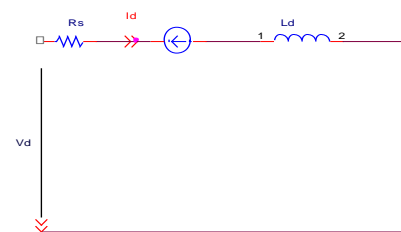
Fig. 1. Equivalent circuit of a permanent magnet synchronous motor in the dq frame.

Total fluxes are expressed by:

$$(\varphi_{abc}) = (L)(i_{abc}) + (\varphi'_{abc})$$

$$L = \begin{bmatrix} L_{ss} & M_s & M_s \\ M_s & L_{ss} & M_s \\ M_s & M_s & L_{ss} \end{bmatrix}, \qquad (3)$$

L_{ss} and M_s representing the self-inductance and the mutual inductance between stator windings. ϕ'_{abc} is the rotor flux seen by the stator windings. It represents the amplitudes of the voltages induced in the stator phases without load. Substituting (3) in (1):

$$[V_{abc}] = (R)(i_{abc}) + L\frac{d(i_{abc})}{dt} + \varphi'_{abc}. \qquad (4)$$

The electromagnetic torque is expressed by:

$$T_e = \frac{1}{\omega_r}(e_{abc})^t(i_{abc}), \qquad (5)$$

where $e_{abc} = \frac{d(\phi_{abc})}{dt}$ represents the electromotive forces generated by the stator phases. ω_r is rotation speed of the rotor in [rad/s]. Note that the system (4) leads to joined and highly non-linear equations. To simplify this problem, the majority of research in literature prefer to use the Park transformation which, by a transformation applied to real variables (voltages, currents and flux), provides fictive variables called dq components of Park's equations. Physically, this transformation is interpreted as a substitution for stationary windings (a,b,c) by rotating windings (d,q) which rotate with the rotor. This transformation makes the dynamic equations of AC motors simpler. The Park transformation is defined as follows:

$$(x_{dq0}) = (K_\theta)(x_{abc}), \qquad (6)$$

where X may be a current, a voltage or a flux and θ is the rotor position. X_{dq} represent longitudinal and transversal components of the stator variables (voltages, currents, fluxes and inductances). The transformation matrix K_θ is by:

$$[K_\theta] = \sqrt{\frac{2}{3}}\begin{bmatrix} \frac{1}{\sqrt{2}} & Cos\theta & -Sin\theta \\ \frac{1}{\sqrt{2}} & Cos\left(\theta - \frac{2\pi}{3}\right) & -Sin\left(\theta - \frac{2\pi}{3}\right) \\ \frac{1}{\sqrt{2}} & Cos\left(\theta - \frac{4\pi}{3}\right) & -Sin\left(\theta - \frac{4\pi}{3}\right) \end{bmatrix}. \qquad (7)$$

The inverse matrix:

$$[K_\theta]^{-1} = [K_\theta]^t = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ Cos\theta & Cos\left(\theta - \frac{2\pi}{3}\right) & Cos\left(\theta - \frac{4\pi}{3}\right) \\ -Sin\theta & -Sin\left(\theta - \frac{2\pi}{3}\right) & -Sin\left(\theta - \frac{4\pi}{3}\right) \end{bmatrix}. \qquad (8)$$

Applying the transformation (6) to the system (1), we have the electrical equations in the d_q reference:

$$[V_d] = [R_s I_{sd}] + \frac{d(\varphi_{sd})}{dt} - \omega_r \varphi_{sq}, \qquad (9)$$

$$[V_q] = [R_s I_{sq}] + \frac{d(\varphi_{sq})}{dt} - \omega_r \varphi_{sd}. \qquad (10)$$

The flux equation:

$$\varphi_{sd} = L_{sd}i_{sd} + \varphi_f, \qquad (11)$$

$$\varphi_{sq} = L_{sq}i_{sq}, \qquad (12)$$

ϕ_f is the flux created by the magnets in the rotor. By replacing (11) and (12) and in v_q, v_d we obtain the following equations:

$$V_d = R_s i_{sd} + L_{sd}\frac{di_{sd}}{dt} - \omega_r L_{sq}i_{sq}, \qquad (13)$$

$$V_q = R_s i_{sq} + L_{sq}\frac{di_{sq}}{dt} + \omega_r\left(L_{sd}i_{sd} + \varphi_f\right). \qquad (14)$$

Equations (13) and (14) form a second order differential equation system that models the electrical behaviour of the synchronous permanent magnet [1], [2].

III. FIELD ORIENTED CONTROL

We can determine the reference torque to impose on the motor and the speed reference from the electromagnetic torque equation expressed in terms of Park's components, if we impose the current. To preserve the torque Te proportional to the current isq, we must control the angle $\alpha = \pi/2$, and the angle α is determined by the following formula.

A. Inverter modeling

The inverter transforms a DC voltage into an alternating voltage with a varying amplitude and frequency. Its bridge structure is composed mostly of electronic switches such as IGBTs, power transistors or thyristors. Its operating principle is based on controlled switching in a suitable manner (usually a pulse width modulation), the source is modulated to obtain a wanted AC signal frequency. Two types of inverters are used; the voltage inverter and the current inverter. The voltage inverter with six switches, supplied by the photovoltaic generator and operating in pulse-width modulation (PWM) is commonly used for this application in Fig. 2.
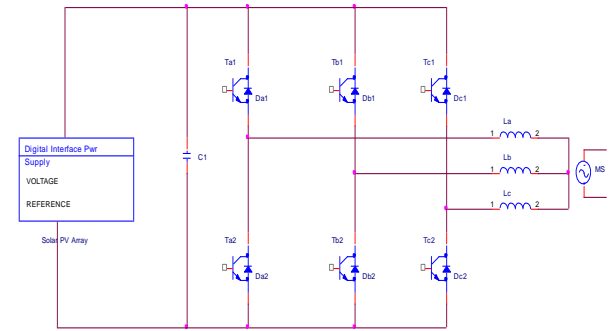
Fig. 2. PMSM supply using voltage inverter.

B. The hysteresis current control technique

It is a simple technique directly interested in current control; it limits the maximum current and is less sensitive to load variations. This method is used to control the current of a to follow a sinusoidal reference current calculated from the currents i_{sdref}, i_{sqref} and from the rotor position ϕ. If the error, which is the difference between the reference current of a phase and the same phase current reached the upper limit $\left(i_{ref} + \Delta I\right)$, the switch

arm of the inverter corresponding to the same phase is started and connected to the (-) pole of the power source to reduce the current; while if the error reached the lower limit $(i_{ref} - \Delta I)$, the switch connected to the positive terminal of the power source should be started to increase the flow of the corresponding phase. The lower and upper limits of the hysteresis band ΔI are set by the motor absorbed current and the maximum switching frequency of switches respectively. A narrow band of hysteresis implies a current more similar to the sine wave with a low harmonic content, and a switching frequency higher and higher, and vice versa.

The current references are given by voltage inverter in such to force the phase currents of the motor these currents are sinusoidal functions of rotor position. They create in the gap a field with magnetic axis in quadrature with the axis of the magnets' field. They are in phase with the electromotive forces induced in these windings by the magnets. When the reference current in a phase deviates from its reference, each controller requires switching the switches of each inverter arm and keeps it within the hysteresis band ΔI in Fig. 3.
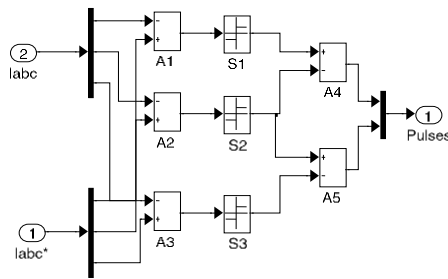

Fig. 3. Representation of hysteresis current control.

C. Modeling of the photovoltaic cell

The photovoltaic generator which produces a continuous electrical current is represented by a standard model with a single diode, established by Shockley for a single PV cell and generalized to a PV module by considering it as a set of identical cells connected in series-parallel [4].

IV. SPEED CONTROLLER DESIGN

The design of the speed controller is important from the point of view of imparting desired transient and steady state characteristics to speed controlled PMSM drive system. A proportional pulse integral controller is sufficient for many industrial applications. Selection of the gain and time constant of such a controller by using the symmetric optimum principle is straightforward if the d axis stator current is assumed to be zero in Fig. 4.

In the presence of a d axis stator current, the d and q current channels are cross couples, and the model is nonlinear, as a result of the torque term [5], [6]. A proportional plus integral (PI) controller is used to

process the speed error between the speed reference and filtered speed feedback signals, the transfer function of the speed controller is given as in Fig. 5.
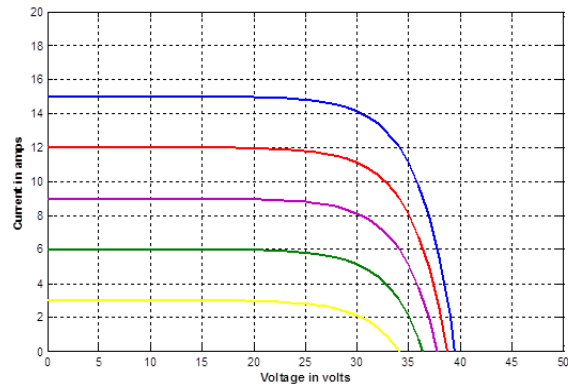
Fig. 4. Characteristics of the power function of the current and the voltage.
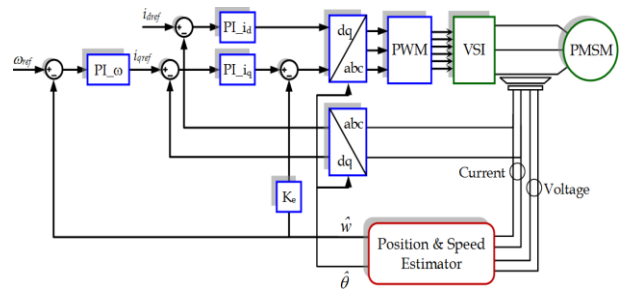
Fig. 5. Block diagram of a speed controlled PMSM drive.

A. Current loop

This induced emf loop crosses the q axis current loop, and it could be simplified by moving the pick-off point for the induced emf loop from speed to current output point. This gives the current loop transfer function in Fig. 6.
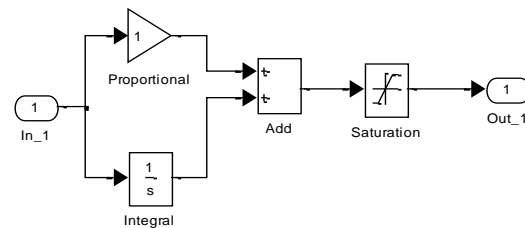
Fig. 6. Implantation the speed controller with saturation in Simulink.

V. LOAD TESTING WITH CONSTANT RESISTANCE TORQUE

The system established in Simulink for a drive system of PMSM with reference current hysteresis control method.

Figures 7-8 shows phase currents of the synchronous motor with permanent magnet. It is clear that the currents are not sinusoidal at startup and becomes sinusoidal when the motor reaches the steady state. The motor absorbs a high current at start-up. i_{dq} currents increase when the motor is controlled by oriented flux, the current i_d is zero ($i_d = 0$), while i_q current increases at start up then stabilizes in steady state. The torque Te developed by the motor follows the instructions properly; its value at startup is five times the value of the rated torque.

Steady speed is the same as that of the commanded speed reference (1790 tr/mn). Simulink program of Matlab is used for simulation tests The PMSM parameters used in the tests are as follows: stator Resistance Rs = 1.4 Ω, stator inductance $L_d = L_q = 0.0006$ Henry Magnet flux linkage $\varphi_f = 0.1679$ $Tesla/m^2$ System inertia J = 0.01176 kgm^2, viscous friction coefficient B = 0.00338818 N.s/rad, DC link voltage using the Lookup Table block in Simulink. We insert the values of the voltage at the PV generator. rated electrical speed ω_n = 1850 tr/mn, pole pairs np = 3 The phase voltages are reconstructed from DC-bus voltage and duty cycle; motor currents are filtered by a three-order low pass filter with pass band edge frequency equal to 12666 rad/s.
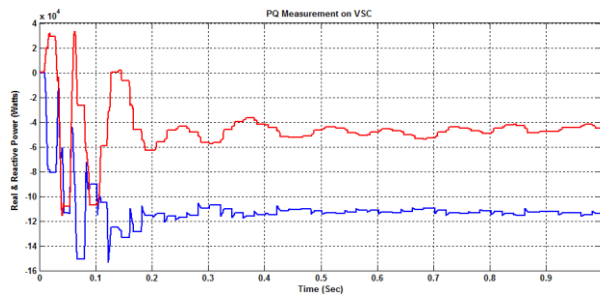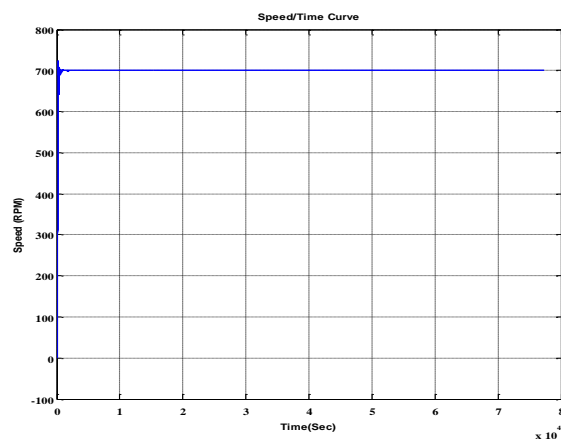

Fig. 7. Diagram of I_{abc} versus time.

Fig. 8. Diagram of speed versus time.

VI. CONCLUSION

The vector control is introduced in order to control the permanent magnet synchronous machine with maximum power. It is based on a transient model. It allows precise adjustment of the torque of the machine and can ensure torque at zero speed. In this paper, we have presented the principle of the permanent magnet synchronous motor field oriented control, fed by a voltage inverter in the presence of a speed loop with a PI corrector. We can conclude that the field oriented control has a good dynamic and static torque and flux results. View that the radiation and temperature are variable in the day it would be interesting in future work to introduce a buck or boost chopper between the PV generator and the inverter to extract maximum energy delivered by the photovoltaic generator.

REFERENCES

[1] A. Kaddouri, "Étude d'une Commande Non-Linéaire Adaptative d'une Machine Synchrone À Aimants Permanents," *Ph.D. Thesis*, Laval University, 2000.

[2] D. Grenier, F. Labrique, H. Buyse, and E. Matagne, *Electromécanique Convertisseur d'énergie et Actionneur*, Dunod, 2001.

[3] M. Pinard, *Commande Electronique Des Moteurs Electriques*, Dunod, 2004.

[4] A. Hansen, P. Sørensen, L. Hansen, and H. Bindner, "Models for a stand-alone PV system," *Risø National Laboratory*, Roskilde, December 2000.

[5] R. Krishnan, "Electric motor drives modeling, analysis, and control," Virginia Tech, Blacksburg, VA.

[6] J. C. Balda and P. Pillay, "Speed controller design for a vector-controlled permanent magnet synchronous motor drive with parameter variations," *Industry Applications Society Annual Meeting*, 1990. Conference Record of the 1990 IEEE.

Non-uniform Surface Impedance Absorbing Boundary Condition for FDTD Method

Yunlong Mao [1], Atef Z. Elsherbeni [2], Si Li [1], and Tao Jiang [1]

[1] Department of Information and Communication Engineering
Harbin Engineering University, Harbin, Heilongjiang, 150001, China
maoyunlong@hrbeu.edu.cn, lisiharbin@hrbeu.edu.cn, jiangtao@hrbeu.edu.cn

[2] Department of Electrical and Computer Engineering
Colorado School of Mines, Golden, CO, 80401, USA
aelsherb@mines.edu

Abstract ─ Recently, we reported a novel absorbing boundary condition (ABC), surface impedance absorbing boundary (SIABC). SIABC has a comparable absorbing performance compared to CPML, but requires a sufficient long distance between the boundary and the scatter. In this paper, we focus on this issue and introduce the non-uniform SIABC. Non-uniform SIABC archives a similar absorbing performance as the uniform SIABC at a same distance, while the number of the air buffer cells is much smaller. Therefore, it is possible for us to make it more efficient relative to uniform SIABC or CPML. An example of a patch antenna is discussed to explore the accuracy and efficiency of non-uniform SIABC. We also compare the memory usage for uniform SIABC, non-uniform SIABC, and 10-layers CPML. All the results indicate that non-uniform SIABC requires much less memory, needs much less time for simulations, which makes it a potential of being one of the most popular ABCs in FDTD method.

Index Terms ─ CPML, FDTD, Non-uniform, SIABC.

I. INTRODUCTION

Surface impedance absorbing boundary condition is first introduced by us in [1]. SIABC comes from the concept of surface impedance boundary condition (SIBC) proposed by Leontovich in 1940s [2]. It is implemented by setting the impedance the same as free space. Compared to CPML [3], SIABC is extremely easy for understanding and for programming, meanwhile, it has a comparable absorbing performance relative to 10 layers CPML. Despite of these advantages, the disadvantage of SIABC is also obvious. In order to absorb the incident waves efficiently, a sufficient long distance between SIABC and the scatter is always required. Therefore, the memory requirement enlarges with the increase of number of air buffer cells. In some situations, this increase may be significant.

On the purpose of reducing the memory usage, we propose the non-uniform SIABC. Non-uniform gridding is a good way to reduce the simulation time with an acceptable accuracy. Non-uniform gridding is proposed originally to deal with complex geometrically details by changing large grids into smaller grids. However, for non-uniform SIABC, the general purpose of non-uniform gridding is inverted. We build larger grids for the air buffer between the boundaries and the scattering objects, in order to reduce the number of air buffer cells. There are a few non-uniform methods proposed [4-6], and in this paper, we will apply the gradually changing grid method as described in [7].

In the first section of this paper, we briefly described non-uniform sub-gridding and how it is combined with SIABC method. Later, we test the accuracy and efficiency of non-uniform SIABC with a mircostrip antenna example. The results are compared with 8-layers CPML and show a good agreement. The memory requirements are also compared to 10-layers CPML for general simulations. It can be concluded that non-uniform SIABC is an excellent ABC, and has the potential of being one of the most popular ABCs in FDTD method.

II. NON-UNIFORM FORMATTING

The geometrical illustration of grid discretization of non-uniform grid SIABC is displayed in Fig. 1 [7].
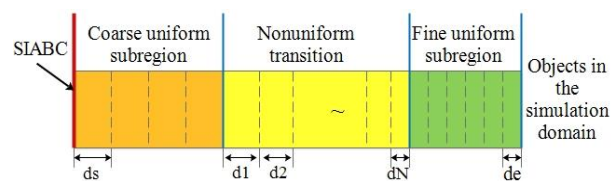

Fig. 1. Geometrical illustration of non-uniform SIABC.

In Fig. 1, the air buffer between the SIABC boundary and the objects inside the simulation domain is divided into 3 sub-regions: coarse uniform sub-region, non-uniform transition sub-region, and fine uniform sub-region. The objects in the simulation domain are located in the fine uniform sub-region. In the coarse uniform sub-region, the cell size is ds; in the non-uniform transition sub-region, the cell size changes from ds to de gradually with the same decreasing ratio, and in the fine uniform sub-region, the cell size is de. The length of the first cell in the transition sub-region is set as:

$$d1 = Rds, \qquad (1)$$

where R is the ratio of change between subsequent half cells. In the transition sub-region, the length of each cell is changing gradually, hence, the length of each cell in the transition sub-region can be represented as:

$$dM = R^M ds, \qquad (2)$$

where M is the index of the Mth cell in the transition sub-region. At the beginning of the fine uniform sub-region, the length of the fine cells should be:

$$de = R^{N+1} ds, \qquad (3)$$

where N is the number of the cells in the transition sub-region. If the transition is desired to happen on a given length ΔT, the total length of the transition sub-region, in addition to one cell at both ends from the uniform regions is:

$$\Delta T + ds + de = \sum_{k=0}^{N+1} ds R^k = \frac{ds - deR}{1 - R}. \qquad (4)$$

Therefore, ratio R can be calculated as:

$$R = \frac{\Delta T + de}{\Delta T + ds}. \qquad (5)$$

Then, the number of cells, N, can be determined using:

$$N = \frac{\log(de/ds)}{\log(R)} - 1. \qquad (6)$$

One should notice that N must be an integer number. It can be rounded to its closest appropriate integer. A smoother transition can be archived if the transition sub-region is selected to be appropriately long.

Another important thing is that the size of the coarse grid should be chosen properly. As is known to all, the larger the coarse grid is, the smaller number of air buffer cells there will be. However, larger grid will cause instability during the simulation [8-10]. Therefore, a proper selection of coarse grids will reduce the cells needed meanwhile keep the stability of the simulation. As time duration is determined by the fine grids, there is no need to consider it separately.

The format of updating equations for non-uniform SIABC are totally the same as those of uniform SIABC method, while the cell size, dx, dy and dz, should be replaced with coarse grids.

III. VERIFICATION EXAMPLES

A. Microstrip patch antenna

In this example, a patch antenna as described in [7] is used to examine the performance of non-uniform SIABC. A microstrip rectangular patch antenna is constructed, as shown in Fig. 2. The problem domain is identified with grid size Δx=2 mm, Δy=2 mm, Δz=0.95 mm, which are regarded as fine grids in this example. A rectangular brick represents the substrate of the antenna with $60mm \times 40mm \times 1.9mm$ dimension and 2.2 dielectric constant. A PEC plate for the ground of the antenna is placed at the bottom side of the substrate covering its entire surface area. A PEC patch is centered on the top surface of the substrate with 56 mm width and 20 mm length in the x and y directions, respectively. The feeding point to the top patch is at the center of the long edge of the patch. A voltage source with 50 internal resistance between the ground plane and the feeding point is defined. This patch antenna operates at 3.45 GHz. The simulation with 8-layers CPML, uniform SIABC and non-uniform SIABC are employed in turn. In each simulation, the number of time steps is 8000. These simulations are executed on a personal computer operating with Inter(R) Core(TM) i7-4700MQ, running at 2.4 GHz.
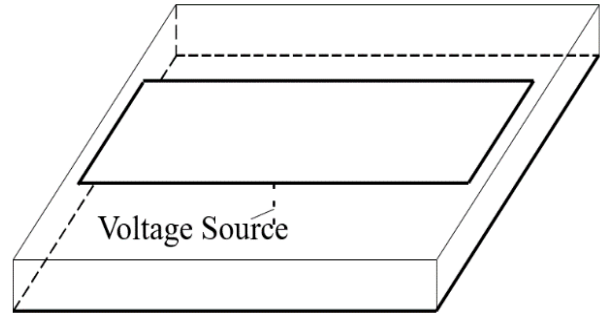
Fig. 2. A microstrip patch antenna configuration.

The performance of non-uniform SIABC is compared to uniform SIABC and 8-layers CPML. For the CPML case, the number of air buffer cells is 10. For the uniform SIABC case, the number of air buffer cells is 20. For the non-uniform SIABC cases, the number of air buffer is 20 and 30. The coarse grid is 5 mm in x and y directions, and 2.85 mm in z direction. The transition length is 21 mm in x and y directions and 8.55 mm in z direction. There are 2 fine grids before the transition region starts. A domain size comparison (in terms of number of cells) of 8-layers CPML, uniform SIABC, and non-uniform SIABC is shown in Table 1 along with the required CPU time. The comparison of the power reflection coefficient is shown in Fig. 3.

Table 1: Comparison of CPML, uniform SIABC and non-uniform SIABC in dimension size and simulation time

ABCs	nx×ny×nz	Domain Size	CPU Time (m)
CPML-8	66×56×38	140,448	8.01
SIABC-uniform	70×60×42	176,400	7.98
SIABC-non 20	52×42×22	48,048	1.35
SIABC-non 30	60×50×26	78,000	2.22

Fig. 3. Comparison of 8-layers CPML, uniform SIABC and non-uniform SIABC for power reflection coefficient.

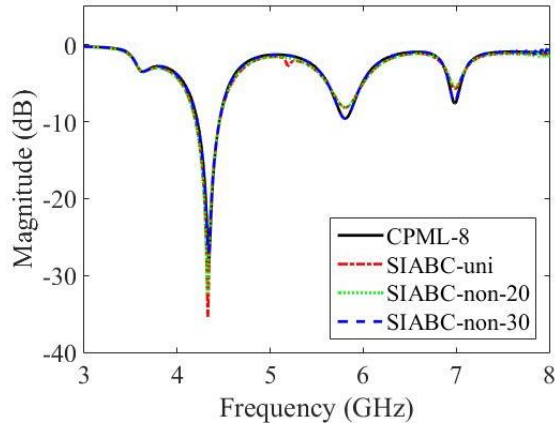From Table 1 and Fig, 3, one can figure out that compared to uniform SIABC or 8-layers CPML, non-uniform SIABC requires much less memory storage, and needs much less time for the simulation with the same reflection coefficient performance. For this example, the domain size for non-uniform SIABC is almost half of that for 8-layers CPML, while the time needed for the simulation is just around 1 fourth. That is because from the complexity aspect, SIABC is much easier than CPML.

B. Memory comparison

In this section, we compare the memory storage requirement for 10-layers CPML, uniform SIABC with its number of air buffer cells ranging from 20 to 50, and non-uniform SIABC with the same air buffer length as uniform SIABC has. Due to the fact that for non-uniform SIABC, the coarse grids will vary with different problems, we made an assumption that for all situations, the coarse grid is three times of the fine grid, and the transition length between coarse sub-region and fine sub-region is 3 times of a coarse grid. Also, there are 2 fine grids in front of the scattering object. Therefore, ratio R can be calculated as:

$$R = \frac{\Delta T + de}{\Delta T + ds} = \frac{9ds + ds}{9ds + 3ds} = \frac{5}{6}, \qquad (7)$$

and the number of cells in the transition sub-region, N, should be:

$$N = \frac{\log(de/ds)}{\log(R)} - 1 = 5.0. \qquad (8)$$

The air buffer gridding is illustrated in Fig. 4. There are 20 fine grids for uniform SIABC case and correspondingly, there are 10 cells for non-uniform SIABC case. There are 3 coarse grids and 2 fine grids. The other 5 cells comes from the transition sub-region, with their cell size decreasing from ds to de.

In order to reduce the number of cells for a certain length, the number of coarse grids should be as many as possible. In our discussion, as the number for fine grids is settled down, we should adjust the length of transition sub-region to ensure that we can have as many coarse grids as possible. Table 2 shows the number of air buffer cells comparison for uniform SIABC and non-uniform SIABC. In the first column, we list the number of air buffer cells for uniform SIABC, which is changing from 20 to 50. The rest of the columns list the corresponding number of cells for non-uniform SIABC: the number of course cells, the length for transition length in terms of de, the number of cells for the transition sub-region, N, the number of cells in the fine region, and the total number of cells used.
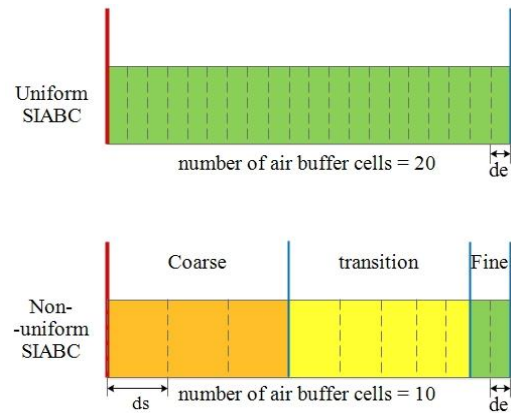
Fig. 4. Air buffer cells configurations for uniform SIABC and non-uniform SIABC with fixed air buffer length.

Table 2: Number of air buffer cells comparison

Uniform (cells)	Non-uniform (cells)				
	Coarse	Transition Length (de)	N	Fine	Total
20	3	9	5	2	10
30	6	10	6	2	14
40	9	11	6	2	17
50	13	9	5	2	21

The memory comparison result is shown in Fig. 5. Generally speaking, by using non-uniform SIABC, the memory requirement is significantly smaller than 10-

layers CPML, especially for problems with small domain size.
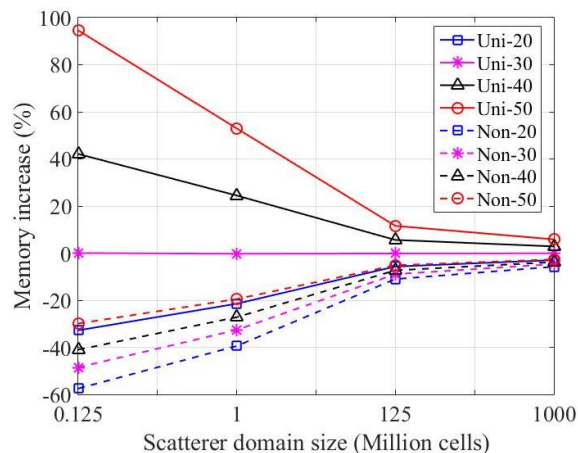
Fig. 5. Memory increase relative to 10-layers CPML.

IV. CONCLUSION

In this paper, an advanced absorbing boundary condition, non-uniform SIABC is proposed which is a combination of non-uniform grid and SIABC. By applying non-uniform grid, we significantly reduced the number of air buffer cells between the SIABC boundaries and the scattering objects leading to much smaller memory requirements relative to uniform SIABC or CPML with acceptable accuracy. Hence, non-uniform SIABC has the potential of being one of the most easy to implement with good performance and less memory requirements ABC for electromagnetic simulations.

REFERENCES

[1] Y. Mao, A. Z. Elsherbeni, S. Li, and T. Jiang, "Surface impedance absorbing boundary for terminating FDTD simulations," *Applied Computational Electromagnetics Society Journal*, vol. 29, pp. 1035-1046, 2014.

[2] M. Leontovich, "Approximate boundary conditions for the electromagnetic field on the surface of a good conductor," *Investigations on Radiowave Propagation*, vol. 2, pp. 5-12, 1948.

[3] J. A. Roden and S. D. Gedney, "Convolutional PML (CPML): An efficient FDTD implementation of the CFS-PML for arbitrary media," *Microwave and Optical Technology Letters*, vol. 27, pp. 334-338, 2000.

[4] S. S. Zivanovic, K. S. Yee, and K. K. Mei, "A subgridding method for the time-domain finite-difference method to solve Maxwell's equations," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 39, pp. 471-479, 1991.

[5] P. Thoma and T. Weiland, "A consistent subgridding scheme for the finite difference time domain method," *International Journal of Numerical Modelling: Electronic Networks, Devices and Fields*, vol. 9, pp. 359-374, 1996.

[6] K. Xiao, D. J. Pommerenke, and J. L. Drewniak, "A three-dimensional FDTD subgridding algorithm with separated temporal and spatial interfaces and related stability analysis," *IEEE Transactions on Antennas & Propagation*, vol. 55, pp. 1981-1990, 2007.

[7] A. Z. Elsherbeni and V. Demir, *The Finite-difference Time-domain Method for Electro-magnetics with MATLAB® Simulations*. 2nd ed., Edison, NJ: SciTech Publishing, an Imprint of the IET, 2016.

[8] J. S. Juntunen and T. D. Tsiboukis, "Reduction of numerical dispersion in FDTD method through artificial anisotropy," *IEEE Transactions on Microwave Theory & Techniques*, vol. 48, pp. 582-588, 2000.

[9] T. Martin and L. Pettersson, "Dispersion compensation for Huygens' sources and far-zone transformation in FDTD," *IEEE Transactions on Antennas & Propagation*, vol. 48, pp. 494-501, 2000.

[10] J. B. Schneider and R. J. Kruhlak, "Dispersion of homogeneous and inhomogeneous waves in the Yee finite-difference time-domain grid," *IEEE Transactions on Microwave Theory & Techniques*, vol. 49, pp. 280-287, 2001.

[11] J. H. Beggs, R. J. Luebbers, K. S. Yee, and K. S. Kunz, "Finite-difference time-domain implementation of surface impedance boundary conditions," *Antennas and Propagation, IEEE Transactions on*, vol. 40, pp. 49-56, 1992.

Time-Dependent Adjoint Formulation for Metamaterial Optimization using Petrov-Galerkin Methods

Xueying Zhang, James C. Newman III, Weiyang Lin, and W. Kyle Anderson

Simcenter: Center of Excellence in Applied Computational Science and Engineering
University of Tennessee, Chattanooga, TN 37403, USA
Xueying-Zhang@mocs.utc.edu, James-Newman@utc.edu, Weiyang-Lin@mocs.utc.edu, Kyle-Anderson@utc.edu

Abstract ─ A time-dependent discrete adjoint algorithm for electromagnetic problems is presented. The governing equations are discretized with a semi-discrete Petrov-Galerkin method. Time advancement is accomplished using an implicit, second-order backward differentiation formula (BDF2). An all-dielectric metamaterial is proposed and gradient-based shape design optimization is conducted. Hicks-Henne functions are utilized for shape parameterization to ensure smooth surfaces, and linear elasticity employed to adapt interior mesh points to boundary modifications. The cost function used in the design optimization attempts to widen the bandwidth of the metamaterial over a desired frequency range. Optimization results demonstrate an increase of the full width at half maximum (FWHM) of reflection from 111 THz to 303 THz.

Index Terms ─ Adjoint-based sensitivity analysis, design optimization, finite element method, metamaterial, Petrov-Galerkin.

I. INTRODUCTION

In recent years, the Petrove-Galerkin (PG) method has become a popular approach in solving a wide range of convection-dominated problems, such as computational fluid dynamics and electromagnetics [1, 2]. The primary benefits of using PG methods arise from their suitability for efficient parallel computing and their high performance of efficiency over DG methods [1]. Since sensitivity analysis is required for gradient-based optimization, the PG method has been extended to provide sensitivity derivatives for both steady-state and time-dependent problems. To this end, a discrete adjoint approach for time-dependent acoustic problems is described in [3, 4] for a Petrov-Galerkin method.

Metamaterials [5] are artificially structured materials with sub-wavelength scale building blocks. The optical properties of metamaterials depend on the constituent materials and geometries of the building blocks. The ability to design such metamaterials opens the pathway for creating materials with designer optical properties. Over the past decade, metamaterials have shown the ability of controlling light propagation [6], absorption and emission [7, 8]. The rapid development of metamaterial research has lead to many scientific breakthroughs such as negative refraction, invisible cloaking, and ultra-compact optical elements [5-9].

In this paper, the research described in [3, 10] is extended for optimization of metamaterials at optical frequencies. A time-dependent discrete adjoint method is employed to obtain sensitivity derivatives. An all-dielectric metamaterial is proposed and optimization is conducted. To follow, the governing equations and discretization method, the adjoint-based sensitivity analysis and shape optimization algorithm, and numerical results are presented.

II. GOVERNING EQUATIONS AND DISCRETIZATION METHOD

The governing equations considered are the two-dimensional source-free Maxwell's equations, which can be written in the conservative form:

$$\frac{\partial q(x,t)}{\partial t} + \frac{\partial F(q(x,t))}{\partial x} + \frac{\partial G(q(x,t))}{\partial y} = 0, \qquad (1)$$

where q, F and G are given by:

$$q = \{D_x, D_y, B_z\}^T,$$
$$F = \{0, B_z/\mu, D_y/\varepsilon\}^T, G = \{-B_z/\mu, 0, -D_x/\varepsilon\}^T, \qquad (2)$$

or

$$q = \{B_x, B_y, D_z\}^T,$$
$$F = \{0, -D_z/\varepsilon, -B_y/\mu\}^T, G = \{D_z/\varepsilon, 0, B_x/\mu\}^T, \qquad (3)$$

for a transverse-electric (TE) mode or a transverse-magnetic (TM) mode, respectively. In the equations above, μ and ε represent the relative permeability and permittivity, respectively. These parameters are assumed to be constants in the current work.

The Petrov-Galerkin discretization begins by formulating a weighted-integral statement of the governing equations by multiplying Eq. (1) by a set of weighting functions, and integrating within each element, as:

$$\int_{\Omega_k} [\phi] \left(\frac{\partial q}{\partial t} + \frac{\partial F}{\partial x} + \frac{\partial G}{\partial y} \right) d\Omega_k = 0, \qquad (4)$$

where ϕ is a weighting function defined by the Streamline Upwind/Petrov-Galerkn (SUPG) method given by:

$$[\phi] = N[I] + \left(\frac{\partial N}{\partial x} [A] + \frac{\partial N}{\partial y} [B] \right) [\tau]$$
$$= N[I] + [P], \qquad (5)$$

where $[\tau]$ represents the stabilization matrix and can be obtained using the following definitions:

$$[\tau]^{-1} = \sum_{k=1}^{n} \left| \frac{\partial N_k}{\partial x} [A] + \frac{\partial N_k}{\partial y} [B] \right|, \qquad (6)$$

$$\left| \frac{\partial N_k}{\partial x} [A] + \frac{\partial N_k}{\partial y} [B] \right| = [T][|\Lambda|][T]^{-1}, \qquad (7)$$

where $[T]$ and $[\Lambda]$ are the right eigenvectors and eigenvalues of the matrix on the left hand side of Eq. (7) respectively, and $[T]^{-1}$ represents the inverse of $[T]$.

Integrating Eq. (4) by parts, the weak formulation of the problem for each elementcan be written as:

$$\int_{\Omega_k} N_i \frac{\partial \boldsymbol{q_p}}{\partial t} d\Omega_k - \int_{\Omega_k} \left[\frac{\partial N_i}{\partial x} F(\boldsymbol{q_p}) + \right.$$
$$\left. \frac{\partial N_i}{\partial y} G(\boldsymbol{q_p}) \right] d\Omega_k + \int_{\Omega_k} [P] \left[\frac{\partial \boldsymbol{q_p}}{\partial t} + \frac{\partial F(\boldsymbol{q_p})}{\partial x} + \right. \qquad (8)$$
$$\left. \frac{\partial G(\boldsymbol{q_p})}{\partial y} \right] d\Omega_k + \int_{\Gamma_k} N_i H(\boldsymbol{q_p^+}, \boldsymbol{q_p^-}, \boldsymbol{n}) dS = 0,$$

where the solution is approximated as $q_p = \sum_{i=1}^{M} \tilde{q}_{p_i} N_i(x)$. In Eq. (8), $H(q_p^+, q_p^-, n)$ represents the flux on the element boundaries, which is determined from the data on either side of the interface using a Riemann solver described in [1]. Equation (8) can be written as an ordinary differential equation in time, which is integrated using an implicit, second-order backward difference formula (BDF2).

III. ADJOINT-BASED UNSTEADY SHAPE OPTIMIZATION

In gradient-based optimization, sensitivity derivatives of the objective function are utilized to construct an appropriate search direction for improving the design. For the direct approach a linear system is formed and solved for each design variable. Numerical evaluation, such as central finite-difference, requires two high converged solutions for each design variable. When the number of design variables is greater than the number of objective functions, adjoint-based sensitivity analysis is the most efficient option for obtaining these derivatives. The number of linear systems requiring solution is equal to the number of objective functions.

A. Design variables and shape parametrization

During a design cycle, the geometry is modified through surface node displacements according to a defined parameterization. The specific method will dictate the set of geometric design variables. A number of surface parameterization methods have been utilized for this purpose in the literature, such as Bezier, B-spline, Hicks-Henne functions, basis vectors, free-

form deformation, etc. In this paper, the Hicks-Henne sine bump function is utilized to ensure smooth surface shape, given by:

$$b_i(x_{si}, \beta_m) = \beta_m sin^4 \left(\pi x_{si}^{\ln(0.5)/\ln(x_{sm})} \right), \qquad (9)$$

where the design variables are set to be the magnitudes of the bump functions $\boldsymbol{\beta} = \{\beta_m, m = 1, \cdots, N_d\}$, where N_d represents the total number of design variables. In Eq. (9), b_i represents the surface node displacement at x_{si} due to the displacement of the surface node at x_{sm}, and β_m denotes the mth component of the design variables associated with the surface node at x_{sm}. The modified surface coordinates are computed by:

$$x_{si}^{new} = x_{si}^{old} + \sum_{m=1}^{N_d} b_i(x_{si}, \beta_m). \qquad (10)$$

As the surface mesh deformation is obtained, the interior mesh points are deformed using linear elasticity to prevent the generation of overlapping elements. This system of equations may be expressed as $[K]\Delta x = \Delta x_s$, where $[K]$ represents the stiffness matrix as found in solid mechanics applications.

For gradient-based optimization the function I refers to a scalar-valued objective function used for minimization. A general formulation for the objective function is expressed in terms of the design variables as $I = I\left(X(\boldsymbol{\beta}), \tilde{q}(X(\boldsymbol{\beta})) \right)$, where $\tilde{q}$ represents the computed unsteady solution, X represents the computational mesh and $\boldsymbol{\beta}$ represents the set of design variables, which control the modification of the surface geometry.

B. Adjoint-base sensitivity calculation

The unsteady residual for time step n can be expressed as:

$$R^n(\boldsymbol{\beta}, X, \tilde{q}^n, \tilde{q}^{n-1}, \tilde{q}^{n-2}) = 0. \qquad (11)$$

The sensitivity derivative can be computed using a forward mode direct differentiation by examining the function dependencies of the objective function. The total differential of I with respect to β can be expressed as:

$$\frac{dI}{d\beta} = \frac{\partial I}{\partial X} \frac{\partial X}{\partial \beta} - \sum_{n=1}^{ncyc} \frac{\partial I}{\partial q^n} \left[\frac{\partial R^n}{\partial q^n} \right]^{-1} \left(\frac{\partial R^n}{\partial X} \frac{\partial X}{\partial \beta} + \right.$$
$$\left. \frac{\partial R^n}{\partial q^{n-1}} \frac{\partial q^{n-1}}{\partial \beta} + \frac{\partial R^n}{\partial q^{n-2}} \frac{\partial q^{n-2}}{\partial \beta} \right). \qquad (12)$$

The adjoint method eliminates the computational overhead caused by repetitive calculations of the solution sensitivities by transposing the inverse of the Jacobian matrix. While a detailed derivation of the procedure is given in [3], the total differential of the objective function may be expressed in terms of the adjoint vector as:

$$\frac{dI}{d\beta} = \frac{\partial I}{\partial X} \frac{\partial X}{\partial \beta} + \sum_{n=1}^{ncyc} \left([\lambda_q^n]^T \left(\frac{\partial R^n}{\partial X} \frac{\partial X}{\partial \beta} \right) \right), \qquad (13)$$

where

$$\lambda_q^n = - \left[\frac{\partial R^n}{\partial q^n} \right]^{-T} \left(\left[\frac{\partial I}{\partial q^n} \right]^T + [\psi_1^{n+1}]^T + [\psi_2^{n+2}]^T \right), \qquad (14)$$

$$\psi_1^n = \left[\frac{\partial R^n}{\partial q^{n-1}}\right]^T \lambda_q^n, \quad \psi_2^n = \left[\frac{\partial R^n}{\partial q^{n-2}}\right]^T \lambda_q^n, \qquad (15)$$

are the adjoint variables.

C. Shape optimization algorithm

Once the sensitivity derivatives of the objective function are evaluated, they are utilized to predict an appropriate search direction. The basic algorithm can be written as:

Algorithm. A discrete adjoint formulation for time-dependent sensitivity derivatives:

(1) Set ψ_1^{n+1}, ψ_2^{n+1}, ψ_2^{n+2} to be zero. Set n to be ncyc.
(2) Solve Eq. (14) for the adjoint variable.
(3) Set the sensitivity derivatives by:

$$\frac{dI}{d\beta} = \frac{dI}{d\beta} + \frac{\partial I}{\partial X}\frac{\partial X}{\partial \beta} + \left[\lambda_q^n\right]^T \left(\frac{\partial R^n}{\partial X}\frac{\partial X}{\partial \beta}\right). \qquad (16)$$

(4) Set $n = n - 1$.
(5) Set $\psi_2^{n+2} = \psi_2^{n+1}$, solve Eq. (15) for ψ_1^{n+1} and ψ_2^{n+1}.
(6) If $n = 1$, stop; otherwise go to step 2.

IV. NUMERICAL RESULTS

A. All-dielectric metamaterial and objective function

All-dielectric metamaterials offer a potential low-loss alternative to plasmonic metamaterials at optical frequencies [11]. In the current work, an all-dielectric metamaterial made of silicon on SiO_2 substrate is proposed as the initial design model. Figures 1 (a-b) illustrate the schematic of metamaterial unit cell and array. The silicon resonators with dimension of W = 200 nm and H = 100 nm are placed on top of a SiO_2 substrate (regarded infinite) with periodicity of P = 300nm. As shown the Fig. 1 (a), the metamaterial is illuminated with polarized light. The electric field is polarized along the x-direction and the magnetic field along the y-direction with wave vector k in z-direction. In this case, the light transmits from the air to the SiO_2 through the silicon resonators.

The metamaterial proposed above is modeled and simulated by an in-house code developed at the Simcenter. The results of reflection over frequency range of 350-650 THz are shown in Fig. 1 (c), with full width at half maximum (FWHM) of 111 THz (479~590 THz) in reflection. For comparative purposes, the current results are shown with those from the commercial software ANSYS© HFSS [12]. The reflection indicates that the metamaterial has the maximum reflection at 516 THz. The electric field distribution at 516 THz, depicted in Fig. 1 (d), clearly illustrates this reflection.

The objective of the current design optimization is to widen the bandwidth of the metamaterial. Accordingly, an objective function is proposed as:

$$I = \int_{f1}^{f2} (1 - Reflection)^2 \, df, \qquad (17)$$

where f_1 and f_2 represent the lower and upper bound of

the desired frequency range.

B. Optimization results

In the current research the DAKOTA toolkit [13] was utilized. DAKOTA's optimization capabilities include a wide variety of optimization methods, and an interface to link with third-party routines. The optimization is performed using a quasi-Newton method (DAKOTA's OPT++ library [14]) based on the Broyden-Fletcher-Goldfard-Shanno (BFGS) variable-metric algorithm, and the line searching approach of More and Thuente [15].

Utilizing the objective function given in Eq. (17), with $f_1 = 300\ THz$ and $f_2 = 700\ THz$, the optimization was performed using different numbers of design variables. Increasing the number of design variables allows for greater geometric flexibility. Figure 2 illustrates the optimization results with 1, 3, and 9 design variables. As seen in Fig. 2 (a), at 426 THz no reflection can be observed from the electric field distribution for the initial model, while high reflection can be observed for the optimized geometries in Figs. 2 (b)-(d) using different number of design variables.

As shown in Fig. 3, the FWHM of reflection for the all-dielectric metamaterial increases from 111 THz to 277 THz, 285 THz, and 303 THz with 1, 3, and 9 design variables, respectively. For the optimized result with 9 design variables, the FWHM of reflection ranges from 376 to 679 THz. As shown in Fig. 4, the electric field distributions at 404 THz, 505 THz and 620 THz are simulated to demonstrate the high reflection property of the optimized metamaterial over the wide frequency range.
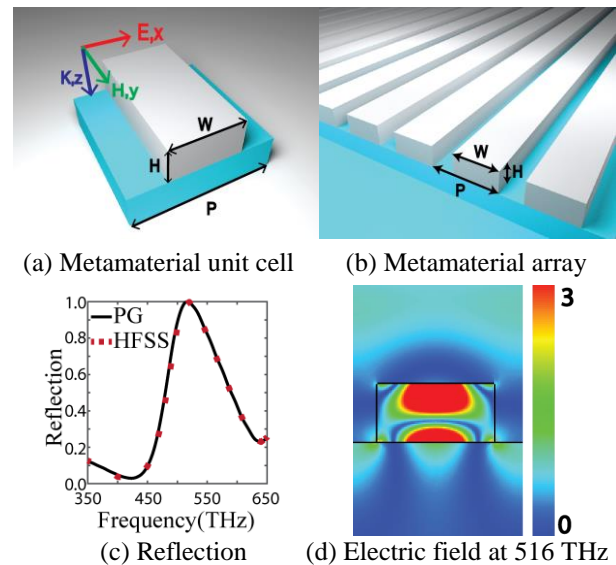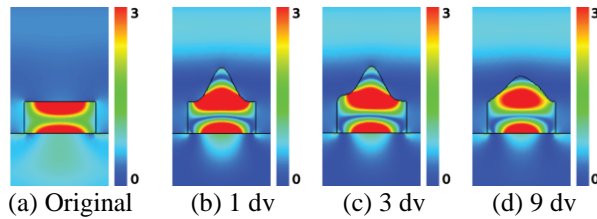

(a) Metamaterial unit cell   (b) Metamaterial array

(c) Reflection   (d) Electric field at 516 THz

Fig. 1. Proposed initial metamaterial model.

(a) Original   (b) 1 dv   (c) 3 dv   (d) 9 dv

Fig. 2. Electric field distribution at 426 THz.
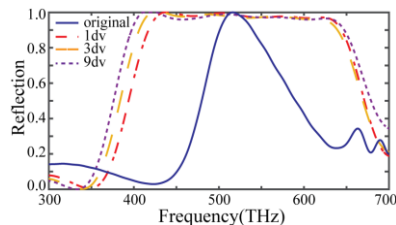
Fig. 3. Comparison of reflection over 300-700 THz.

(a) 404 THz   (b) 505 THz   (c) 620 THz
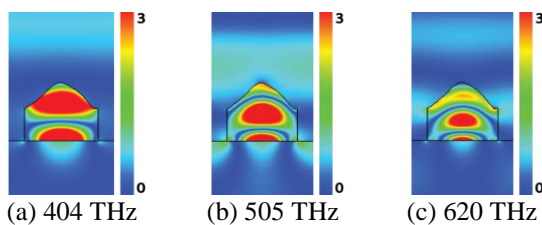
Fig. 4. Electric field distribution of model with 9dv.

V. CONCLUSION

This paper presents an unsteady discrete adjoint approach for performing sensitivity analysis as required by gradient-based optimization algorithms. The simulations are performed by discretizing the source-free Maxwell equations using a Petrov-Galerkin finite-element method. Temporal accuracy is achieved with an implicit, second-order backward differentiation formula (BDF2). Electromagnetic shape optimization is conducted on an all-dielectric metamaterial working at optical frequencies, and considers multiple numbers of design variables. Furthermore, utilizing the current shape optimization procedure the FWHM of reflection was increased from 111 THz to 303 THz.

ACKNOWLEDGMENT

REFERENCES

[1] W. K. Anderson, L. Wang, S. Kapadia, C. Tanis, and B. Hilbert, "Petrov-Galerkin and discontinuous-Galerkin methods for time-domain and frequency-domain electromagnetic simulations," *J. Comput. Phys.*, vol. 230, pp. 8360-8385, 2011.

[2] W. K. Anderson, L. Wang, J. C. Newman III, and S. Kapadia, "Extension of the Petrov-Galerkin time-domain algorithm for dispersive media," *IEEE Microw. Wirel. Compon. Lett.*, vol. 23, no. 5, pp. 234-236, 2013.

[3] W. Lin, W. K. Anderson, J. C. Newman III, and X. Zhang, "Shape optimization of two-dimensional acoustic metamaterials and phononic crystals with a time-dependent adjoint formulation," *AIAA Paper 2016-3830*, 2016.

[4] W. Lin, J. C. Newman III, W. K. Anderson, and X. Zhang, "Broadband shape and topology optimization of acoustic metamaterials and phononic crystals," *AIAA Paper* 2016-3216, 2016.

[5] R. A. Shelby, D. R. Smith, and S. Schultz, "Experimental verification of a negative index of refraction," *Science*, vol. 292, pp. 77-79, 2001.

[6] J. Valentine, S. Zhang, T. Zentgraf, E. Ulin-Avila, D. A. Genov, G. Bartal, and X. Zhang, "Three-dimensional optical metamaterial with a negative refractive index," *Nature*, vol. 455, no. 7211, pp. 376-379, 2008.

[7] W. Li and J. Valentine, "Metamaterial perfect absorber based hot electron photodetection," *Nano Lett.*, vol. 14, no. 6, pp. 3510-3514, 2014.

[8] X. Liu, T. Tyler, T. Starr, A. F. Starr, N. M. Jokerst, and W. J. Padilla, "Taming the blackbody with infrared metamaterials as selective thermal emitters," *Phys. Rev. Lett.*, vol. 107, no. 4:045901, 2011.

[9] W. Li, Z. J. Coppens, L. V. Besteiro, W. Wang, A. O. Govorov, and J. Valentine, "Circularly polarized light detection with hot electrons in chiral plasmonic metamaterials," *Nat. Commun.*, vol. 6:8379, 2015.

[10] L. Wang and W. K. Anderson, "Adjoint-based shape optimization for electromagnetic problems using discontinuous Galerkin methods," *AIAA J.*, vol. 49, no. 6, pp. 1302-1305, 2011.

[11] P. Moitra, B. A. Slovick, W. Li, I. I. Kravchencko, D. P. Briggs, S. Krishnamurthy, and J. Valentine, "Large-scale all-dielectric metamaterial perfect reflectors," *ACS Photonics*, vol. 2, no. 2, pp. 692-698, 2015.

[12] Ansoft High Frequency Structure Simulation (HFSS), ver. 12.0, Ansoft Corporation, Pittsburgh, PA, 2009.

[13] B. M. Adams, L. E. Bauman, W. J. Bohnhoff, K. R. Dalbey, M. S. Ebeida, J. P. Eddy, M. S. Eldred, P. D. Hough, K. T. Hu, J. D. Jakeman, L. P. Swiler and D. M. Vigil, "DAKOTA: version 5.2 user's manual," *Sandia Tech. Rep.*, SAND2010-2183, 2009.

[14] J. C. Meza, R. A. Oliva, P. D. Hough, and P. J. Williams, "OPT++: An object oriented toolkit for nonlinear optimization," *ACM Trans. Math. Software*, vol. 33, no. 2, pp. 126-136, 2007.

[15] J. J. More and D. J. Thuente, "Line search algorithms with guaranteed sufficient decrease," *ACM Trans. Math. Software*, vol. 20, pp. 286-307, 1994.

A Non-Focal Rotman Lens Design to Support Cylindrically Conformal Array Antenna

Toan K. Vo Dai [*], Tuan Nguyen, and Ozlem Kilic

Department of Electrical Engineering and Computer Science
The Catholic University of America, Washington, DC 20064, USA
[*]30vodai@cua.edu

Abstract ─ Rotman lenses offer broad bandwidth, and render to planar structures making them ideal for a variety of applications. However, a limitation of the Rotman lens is that it is based on the assumption of supporting linear arrays. In this paper, we develop a new design technique to enable the lens to feed a conformal array antenna.

Index Terms ─ Beam Forming Networks (BFNs), conformal array antennas, Particle Swarm Optimization (PSO), Rotman lens.

I. INTRODUCTION

Rotman lens, which was invented by Rotman and Turner in 1962, is an analog beam former that creates a specific phase taper at its output ports to feed an array antenna. Based on the input parameters (such as the scanning angle (φ), the focal angle (α), number of input and output ports; the lens equations can be solved to determine its receive contour, delay lines (w_i) and the focal arc to generate a desired phase taper [1]. In the past few decades, numerous Rotman lens designs have been developed [2-5]. However, a limitation of these designs is the assumption of a uniform linear array at the output ports. This prohibits their applications to conformal array antennas.

The ability to feed conformal array antennas with a Rotman lens would be of interest as these systems can be applied to a variety of applications such as antennas residing on the surface of an airplane's wing, the body of a missile or a high-speed train, which are used for communications or navigation purposes. In this paper, we develop a new design inspired by Rotman lens to feed a cylindrically conformal array. We solve for the lens equations to satisfy the constraints of the design by employing the Particle Swarm Optimization (PSO) [4]. Our approach applies to other conformal geometries as well. For the purposes of this paper, we will use our method to design a lens with 3 input ports feeding an 8-element conformal array operating at a center frequency of 10 GHz.

The remainder of the paper is organized as follows. Section II introduces the overview of our system design. Section III gives a brief overview of the PSO algorithm. Section IV discusses the optimization of phase values for the conformal array, and Section V reviews the proposed design procedure of the non-focal lens. The performance of the optimization process is discussed in Section VI. Simulation results are presented in Section VII, followed by the conclusions in Section VIII.

II. OVERVIEW OF THE SYSTEM DESIGN

In this paper, we demonstrate the design of a Rotman lens inspired, non-focal microwave lens to support an array conforming to a cylindrical surface as shown in Fig. 1.
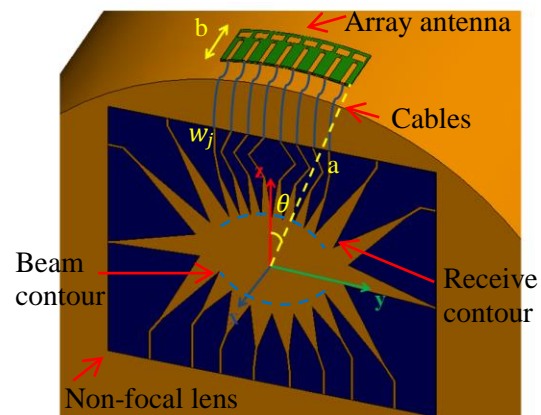

Fig. 1. Overall system design.

Two tasks are considered in order to design the lens. The first task is to optimize the required phase information to scan the conformal array in the desired directions. The second task is to use the phase information obtained in the first task as an objective function to design the non-focal lens that generates the desired phase values identified in the first task. Both tasks require an optimization method for which we choose PSO for its

ease of use and reliable performance. The optimization in the first task is straightforward while the second task can be challenging as it involves numerous variables to be optimized under a set of constraints.

III. OVERVIEW OF PSO

Particle Swarm Optimization (PSO) is a random search algorithm, which simulates the behavior of bees in their search for the best location in the field [6]. As with all random search algorithms, the objective is to minimize a cost function defined for the specific problem. In the optimization process, the bees sample the optimization space and decide where to go next as they search for the best location based on the collective intelligence of the swarm and their personal experiences. The velocity that determines the next position of an agent is based on these two factors and the user defined constants w, c_1, c_2, which correspond to the weights of velocity along the original direction of the bee and towards the personal and global best values, respectively. The user also specifies the total number of agents, maximum number of iterations for termination and the boundary conditions for the search space.

IV. OPTIMIZATION OF PHASE EXCITATION FOR CONFORMAL ARRAY

Our design of the cylindrically conforming array antenna is based on the parameters shown in Fig. 1, where a is the radius of the cylinder, b represents the width of the array, and θ is the angle measured with respect to the z-axis in radians, and determines the total length l of the array by the relation: $l = 2a\theta$. The phase distribution for each element of the array to acquire the desired scan direction depends on the curvature characteristics of the surface it resides on.

Our paper investigates two different array geometries as depicted in Table 1. Both arrays will be designed at a center frequency of 10 GHz; the center-to-center spacing between the radiating elements is $\lambda_0/2$ (where λ_0 is the wavelength at 10 GHz).

Table 1: Design parameters of array for two cases

Parameters	Case 1	Case 2
θ (°)	16	32
a (cm)	30	15
b (cm)	5	5
l (cm)	16.8	16.8

In this paper, we first employ PSO to optimize the phase distribution for the 8-element conformal array for a desired scanning angle. The desired scan positions are [-30°, 0, 30°], where the angle is calculated with respect to the z-axis. This results in 8 variables in the interval of [0, 2π] to be optimized for each scan position. We employ 100 agents, and terminate the search after 500 iterations.

The user defined parameters $c_1 = c_2$ and w are chosen to be 2 and 0.9, respectively. We minimize the cost function as given in (1) to achieve the desired solution:

$$F(\vec{\alpha}) = w(|SLL|_{dB} - |SLL_{max}|_{dB}) - Gain(\theta, \phi, \vec{\alpha}), (1)$$

where $Gain$ is the power gain in the (θ, ϕ) direction, $\vec{\alpha}$ is a 24 element vector corresponding to the phase distribution for each input port, SLL_{dB} refers to sidelobe levels achieved from the PSO algorithm and SLL_{maxdB} is the maximum level allowed. The PSO stops the search when F $\leq$ -15 dB (when Gain $\geq$ 15 dB and SLL_{dB} achieved is approximately equal SLL_{maxdB} required) or the maximum number of iterations is reached. The radiation patterns of arrays for both cases are shown in Figs. 2 (a) and 2 (b), respectively. The maximum gain we obtain is 15 dB for each scanning angle while the SLL is about 10 dB down from the peak.
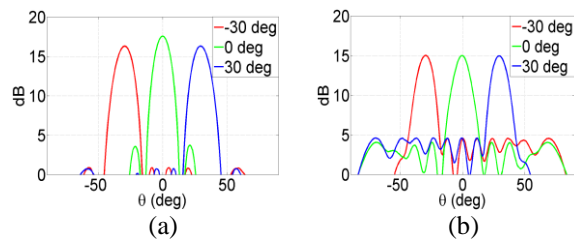

Fig. 2. Radiation pattern of conformal array in: (a) case 1 and (b) case 2.

V. OPTIMIZATION OF THE NON-FOCAL LENS

Figure 3 shows a general concept of the optimized lens. Due to the curvature dependence of the phase values feeding the conformal array, it is not possible to directly apply the Rotman lens formulation in [1] to solve for the receiving contour, delay lines and beam contour since the Rotman lens formulation assumes a linear array at its output. We modify the general equations in finding the path length from an input port on the beam contour to an output port on the receiving contour. By optimizing the locations of input and output ports as well as the delay lines, we obtain the expected phase value at the input of conformal array as described in (2):

$$\phi_{i,j} = k_r\sqrt{(x_j - x_{fi})^2 + (y_j - y_{fi})^2} + k_{eff}.w_j, (2)$$

where (x_{fi}, y_{fi}) is the coordinate of the input port, (x_j, y_j) is the coordinate of the output port, w_j is the length of the delay line corresponding to each output port, $k_r = \frac{2\pi}{\lambda}\sqrt{\varepsilon_r}$ and $k_{eff} = \frac{2\pi}{\lambda}\sqrt{\varepsilon_{eff}}$ where λ is the free space wavelength at the center frequency and ε_{eff} is the effective dielectric constant of the microstrip line feeding the array.

In our formulation, we assume that the receiving contour lies on an ellipse. This constraint ensures a

smooth curvature at the output. There are a variety of options in choosing the shape of the receiving contour. We choose an elliptical arc because it resembles the shape in a conventional Rotman lens. As shown in Fig. 3, we have 24 variables to optimize, which include the coordinates of input ports $(x_{f1} \dots x_{f3}, \ y_{f1} \dots y_{f3})$, y-coordinates of output ports $(y_1 \dots y_8)$, radii of elliptical receiving contour (l_1, l_2) and corresponding delay line lengths $(w_1 \dots w_8)$. For cases with higher number of input or output ports, the number of optimization variables increase, which would inevitably slow down the convergence of the PSO to good solutions.
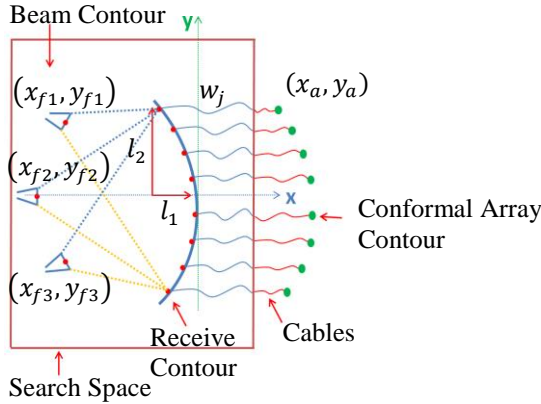

Fig. 3. Concept of non-focal lens design.

The fundamental challenges of this problem are how to design a lens so that it not only supports a conformal array, but is also practical for fabrication. Thus, we must implement a suitable cost function in PSO to guarantee a smooth curvature, and to position all output ports well for maximum power reception. Below are the constraints we choose for the PSO algorithm to design an appropriate lens:

i. Ports are randomly distributed on the receiving contour with spacing requirement $\boldsymbol{\lambda_r/2 \leq S_i \leq 3\lambda_r}$ where $\boldsymbol{\lambda_r = \lambda/\sqrt{\varepsilon_r}}$ is the effective wavelength inside the substrate, and $\boldsymbol{S_j}$ is the arc length between two adjacent output ports $\left(\boldsymbol{x_j, y_j}\right)$ and $\left(\boldsymbol{x_{j+1}, y_{j+1}}\right)$. The minimum spacing between two adjacent ports is $\boldsymbol{\lambda_r/2}$ to make sure that it is posible for the lens to be fabricated while the maximum spacing is $\boldsymbol{3\lambda_r}$ constrains the overall dimensions of the lens.

ii. Beam ports are positioned randomly with a constraint on the minimum distance between each pair, such that $\boldsymbol{d_i = \sqrt{\left(x_{fi} - x_{fi+1}\right)^2 + \left(y_{fi} - y_{fi+1}\right)^2} \geq \frac{\lambda_r}{4}}$. Unlike the conventional lens, our approach has no limitations for the beam ports to lie on a circular or elliptical arc.

iii. An important constraint is to make sure beam ports

do not block each other. We define an illuminating region from a beam port to the output ports (the regions are separated by the dashed blue and solid yellow lines as shown in Fig. 3). If the other input ports happen to lie in this region, the PSO algorithm will not accept these cases as solutions.

VI. PERFORMANCE OF PSO

After all the constraints are satisfied, the phases at the output ports are calculated and compared with the set of desired phase values to assess the cost function for this optimization problem. We use 600 agents and a maximum of 10000 iterations, where the user defined variables are chosen as $c_1 = c_2 = 2$ and $w = 0.9$. The optimization aims to minimize the cost function (3):

$$F = \sqrt{\sum_{i=1}^{N_{input}} \sum_{j=1}^{N_{output}} \left(\phi_{i,j} - \phi_{i,jref}\right)^2}, \qquad (3)$$

where $\phi_{i,j}$ is calculated in (2) and $\phi_{i,jref}$ is the desired phase values. To limit the overall dimensions of the lens, we set the search space for input ports, output ports and length of a delay line in appropriate intervals. For example, for a 3-input, 8-output Rotman lens, we would want $x_{fi} \in [-10\lambda_r, 0]$, $y_{fi}, y_j \in [-10\lambda_r, 10\lambda_r]$, $a, b \in [4\lambda_r, 10\lambda_r], w_j \in [\lambda_r, 10\lambda_r]$. Figure 4 shows the convergence rate of the optimized lens in case 1 and 2 over 10000 iterations.
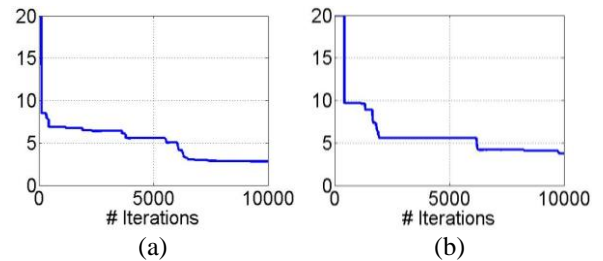

Fig. 4. Performance of PSO for optimized lens in case 1 and case 2.

VII. RESULTS FOR NON-FOCAL LENS

To verify the robustness of our lens design, we design a 3-input, 8-output used to feed a slightly bent conformal array with parameters specified in case 1 and case 2, Table 1. The substrate we are using is Duroid 5880 with $\varepsilon_r = 2.2$. Figures 5, 6 show the simulation model of the lens using commercial software package FEKO and its performance over a bandwidth from 9 GHz to 11.5 GHz. Figures 7, 8 show the radiation pattern of the array fed by the output of the lens at 10 GHz.

The overall dimensions of the non-focal lens design in each case are greater compared with the conventional Rotman lens working at 10 GHz and using the same substrate.
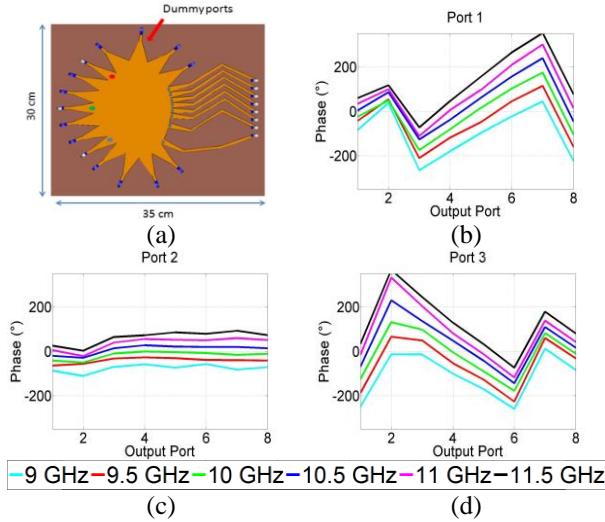
Fig. 5. 3-input, 8-output lens to feed conformal array in case 1: (a) CAD model, phase performance, (b) Port 1, (c) Port 2, and (d) Port 3.
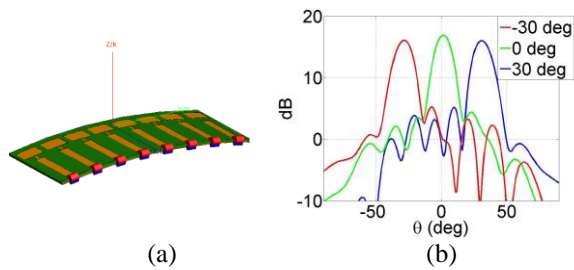

Fig. 6. (a) Bent array in case 1, and (b) radiation pattern of the array at -30°, 0°, 30°.
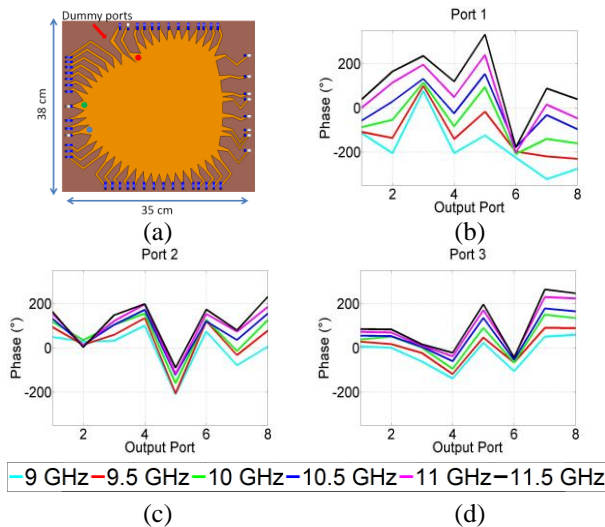

Fig. 7. 3-input, 8-output lens to feed conformal array in case 2: (a) CAD model, phase performance, (b) Port 1, (c) Port 2, and (d) Port 3.
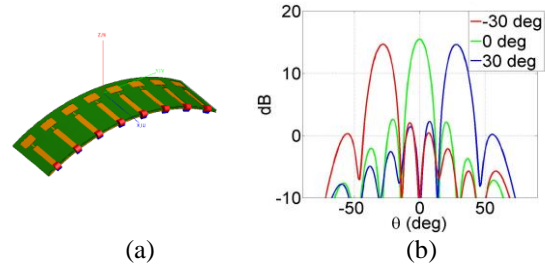

Fig. 8. (a) Bent array in case 2, and (b) radiation pattern of the array at -30°, 0°, 30°.

VIII. CONCLUSIONS

In this work, we investigated a new design technique to extend the design equations of Rotman lens to feed conformal array antennas. The applicability was shown using a cylindrical curvature, but the method applies to other curvatures as well. The results show that with similar total dimensions compared with conventional Rotman lens working at the same frequency and substrate, this lens design is able to support a cylindrical array. The proposed design works with any array antenna on different singly curved surfaces. While the concept has been shown for a 3-input port lens, the optimization can be run for higher number of input and output ports. However, for such cases, the number of optimization variables increase, which would inevitably slow down the convergence of the PSO algorithm to good solutions.

REFERENCES

[1] W. Rotman and R. Turner, "Wide-angle microwave lens for line source applications," *IEEE Transactions on Antennas and Propagation*, vol. 11, pp. 623-632, 1963.

[2] M. S. Smith, "Design considerations for Ruze and Rotman lenses," *Radio and Electronic Engineer*, vol. 52, no. 4, pp. 181-187, 1982.

[3] R. C. Hansen, "Design trades for Rotman lenses," *IEEE Transactions on Antennas and Propagation*, vol. 39, no. 4, pp. 464-472, 1991.

[4] O. Kilic and S. Weiss "Rotman lens designs for military applications," *Radio Science Bulletin*, no. 333, pp. 10-24, 2010.

[5] J. Dong and A. I. Zaghloul, "Non-focal 2D Rotman lens design," *IEEE Transactions on Antennas and Propagation*, 2008.

[6] J. Kennedy and R. C. Eberhart, "Particle swarm optimization," in *Proc. IEEE Conf. Neural Networks IV*, Piscataway, NJ, 1995.

Efficient MCF Evaluation in a Turbulent Atmosphere over Large Structure Constant Interval

Naser A. Abu-Zaid

Department of Electrical and Communications Engineering
An-Najah National University, Nablus, P.O. Box 7, West Bank
naser_res@yahoo.com, naserzaid@najah.edu

Abstract ─ A fast and accurate method is derived and simulated to compute the mutual coherence function (MCF) of an electromagnetic beam wave propagating through atmospheric turbulence over a large interval of structure constant. This method is based on expanding the integral equation solution for one value of structure constant into its Pade' approximation to cover large fluctuation interval. The expansion is compared with numerical results, and a very good agreement was obtained. Such computations are important in practical fields as remote sensing, imaging systems, and optical communications.

Index Terms ─ Atmospheric scattering, atmospheric turbulence, coherence, mutual coherence function, Pade' approximation.

I. INTRODUCTION

Electromagnetic wave scattering problem by a turbulent atmosphere was solved several decades ago [1], and does not represent a challenging problem anymore. Solutions were over and done with the Rytov method or by perturbation theory for several moments such as the second order statistical moment known as mutual coherence function (MCF) [1, 2]. MCF within Rytov approximations were compared to that which resulted from the parabolic equation for the case of weak and strong fluctuations, including cases of plane, spherical and beam waves [3]. In recent years, the advent of advanced processors and modern applications in astronomy, remote sensing, free space optical communications, and imaging [4] renewed the interest in atmospheric scattering of beam electromagnetic waves. Consequently, finite element methods [5] and finite difference methods [6] were applied extensively to solve the scattering problem [7].

In all mentioned methods, field solutions were obtained for a single value of structure constant or at a single or double excitation frequencies, and none were obtained over an interval of structure constant. Several authors succeeded in obtaining the solution over wide range of frequency [8], and over a range of complex permittivity [9]. Nevertheless, those methods dealt with deterministic media. In this work, solution for MCF is found over a continuous interval of structure constant in a turbulent atmosphere. The solution method is based on computing the MCF integral at a single value of structure constant, expanding integral solution into its truncated power series, finding power series coefficients from which Pade' approximants [10] are calculated, then approximations are established.

Generally, solving Maxwell's equations or the wave equations for non-symmetric geometries or obstacles usually requires resorting to numerical techniques. Domain discretization in such techniques, in either finite difference methods or finite element methods, is a necessity. Such procedures consume computer time and memory, especially for large scale problems, which put limits on the electrical size of problems under study, specifically when the solution is required for several values of some parameter of interest. Pade' approximation offers a method of reducing computer CPU time and memory, while maintaining the high accuracy of the solution.

II. FORMULATION

Consider a beam electromagnetic wave propagating along the z direction in a randomly turbulent atmosphere. The wave length is assumed to be much smaller than eddies forming fluctuating permittivity field. Such wave satisfies the stochastic Helmholtz equation [1-3]:

$$\nabla_t^2 U(\boldsymbol{\rho}) + 2jk\frac{\partial U(\boldsymbol{\rho})}{\partial z} + k^2 \tilde{\varepsilon}(\boldsymbol{\rho})U(\boldsymbol{\rho}) = 0, \quad (1)$$

where $\tilde{\varepsilon}$ is the stochastic permittivity field, ∇_t^2 is the transverse scalar Laplacian, $\boldsymbol{\rho}$ is the transverse radial vector, j is the imaginary unit, $U(\boldsymbol{\rho})$ is the transverse electric field component, $k = 2\pi/\lambda$ is the free space wave number, and λ is the free space wave length. The important parameter of interest characterizing such waves, is the mutual coherence function MCF, defined [1, 2]:

$$\Gamma_2(\boldsymbol{\rho}_1, \boldsymbol{\rho}_2, z) = \langle U(\boldsymbol{\rho}_1, z)U^*(\boldsymbol{\rho}_2, z)\rangle, \quad (2)$$

where $\langle \cdot \rangle$ represent ensemble average, $\boldsymbol{\rho}_1 = x_1\hat{x} + y_1\hat{y}$, $\boldsymbol{\rho}_2 = x_2\hat{x} + y_2\hat{y}$, and U^* indicates conjugation. Assuming

a beam with Gaussian amplitude distribution at the transmitting aperture $z = 0$, with a waist radius w_o and a phase front with radius of curvature R_o, an expression for MCF within the second order Rytov approximation was derived and evaluated for a beam wave propagating through a turbulent atmosphere [1-3]:

$$\Gamma_2(\boldsymbol{\rho}_c, \boldsymbol{\rho}_d, z) = \frac{w_o^2}{w^2} exp(g_1) exp(g_2),$$

$$g_1 = -\frac{k}{2} \left(\frac{w_o^2}{w^2}\right) (g_3 - j2g_4),$$

$$g_2 =$$
$$-4.352 k C_n^2 \int_0^{L_z} \left(\gamma_I(z) \frac{L_z - z}{k}\right)^{5/6} {}_1F_1\left(-\frac{5}{6}, 1; g_5\right) dz,$$

$$g_5 = -\frac{k|\gamma_R \rho_d - j2\gamma_I \rho_c|^2}{4\gamma_I(z)(L_z - z)},$$

$$g_3 = 2\alpha_1 \left(\rho_c^2 + \frac{\rho_d^2}{4}\right),$$

$$g_4 = [\alpha_2 - (\alpha_1^2 + \alpha_2^2) L_z](\boldsymbol{\rho}_c \cdot \boldsymbol{\rho}_d),$$

$$w^2 = w_o^2[(1 - \alpha_2 L_z)^2 + \alpha_1^2 L_z^2],$$

$$\boldsymbol{\rho}_c = \frac{\boldsymbol{\rho}_1 + \boldsymbol{\rho}_2}{2},$$

$$\boldsymbol{\rho}_d = \boldsymbol{\rho}_2 - \boldsymbol{\rho}_1,$$

$$\alpha = \alpha_1 + j\alpha_2 = \frac{2}{kw_o^2} + j\frac{1}{R_o},$$

$$\gamma(z, L_z) = \frac{1 + j\alpha z}{1 + j\alpha L_z} = \gamma_R - j\gamma_I, \qquad (3)$$

where ${}_1F_1$ is the confluent hypergeometric fuction, C_n^2 is the refractive index structure constant, L_z is the distance from the output aperture. Expanding the unknown MCF in Eq. (3) into its power series about an arbitrary structure constant C_{no}^2, as:

$$\Gamma_2 = \sum_{i=0}^{\infty} a_i (C_n^2 - C_{no}^2)^i, \qquad (4)$$

where

$$a_i = \frac{1}{i!} \frac{\partial \Gamma_2}{\partial C_n^2}\bigg|_{C_n^2 = C_{no}^2}$$

$$= \frac{1}{i!} \frac{w_o^2}{w^2} \left(\frac{1}{C_n^2} g_2\right)^i exp(g_1) exp(g_2)\bigg|_{C_n^2 = C_{no}^2}. \qquad (5)$$

$[L/M]$ Pade' approximants are obtained by truncating the power series at N, then matching to a rational:

$$\sum_{i=0}^{N} a_i (C_n^2 - C_{no}^2)^i \approx \frac{\sum_{l=0}^{L} p_l (C_n^2 - C_{no}^2)^l}{1 + \sum_{m=1}^{M} q_m (C_n^2 - C_{no}^2)^m}. \qquad (6)$$

$N + 1$ equations reached from expansion of Eq. (6). The $q's$ are attained from the last M of these equations:

$$\begin{bmatrix} q_1 \\ q_2 \\ \vdots \\ q_M \end{bmatrix} =$$

$$-\begin{bmatrix} a_L & a_{L-1} & \cdots & a_{L-M+1} \\ a_{L+1} & a_L & \cdots & a_{L-M} \\ \vdots & \vdots & & \vdots \\ a_{L+M-1} & a_{L+M-2} & \cdots & a_L \end{bmatrix}^{-1} \begin{bmatrix} a_{L+1} \\ a_{L+2} \\ \vdots \\ a_{L+M} \end{bmatrix}, \quad (7)$$

whereas the $p's$ are found from the rest of equations:

$$\begin{bmatrix} a_o & 0 & \cdots & 0 \\ a_1 & a_o & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ a_L & a_{L-1} & \cdots & a_o \end{bmatrix} \begin{bmatrix} 1 \\ q_1 \\ \vdots \\ q_L \end{bmatrix} = \begin{bmatrix} p_o \\ p_1 \\ \vdots \\ p_L \end{bmatrix}. \qquad (8)$$

III. NUMERICAL APPLICATIONS AND VALIDATION

With the derivation of asymptotic expansion coefficients accomplished, an investigation must now be made with regard to their applicability and validity. To demonstrate the efficiency of the technique, three simulations will be carried out on a $1.6\ GHz$ personal computer. In all simulations, 41 structure constant values and 9 transverse space points are implemented for direct numerical solution. As a first check, power series expansion around $C_{no}^2 = 10^{-16}$ with [2/5] Pade' approximants used, and both expansions compared with direct numerical solution as shown in Fig. 1. The propagation parameters given by $w_0 = 0.05\ m$, $L_z = 2.5\ km$, $\lambda = 630\ nm$. It requires 145.9531s to obtain the solution with direct solution, though, it only takes 4.3281s for single point expansion. It may be inferred that the power series approximates well for small values of C_n^2, but drops sharply for larger values, while Pade' expansions shows very good approximation even deep in lager C_n^2 values region.
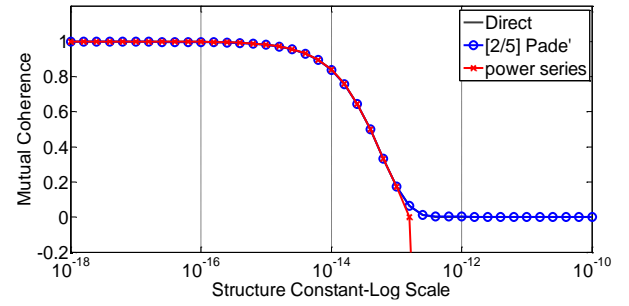

Fig. 1. Normalized mutual coherence function $\Gamma_2(0,0,L_z)$ versus C_n^2.

In Fig. 2, simulation parameters are $C_{no}^2 = 10^{-16}$ $w_0 = 0.005\ m$, $L_z = 2.5\ km$, $\lambda = 630\ nm$. To study the effect of numerator and denominator degrees on quality of Pade' expansions, three different values are taken, namely, $[L/M] = [3/3]$, $[L/M] = [2/4]$ and $[L/M] = [4/2]$. Simulations show that when the denominator degree is larger than numerator degree, Pade' expansions approximated the solution very well. In the contrary, other orders did not for larger values of C_n^2, the approximation skyrockets when numerator degree is larger, and drops steeply when numerator and denominator degrees are equal. The difference in time between the three cases is a fraction of a second and can be neglected. However, it should be noted that as the number of power series coefficients is increased to larger integers, their values become very prohibitive and Pade' matrix becomes close to singular. This is expected since their values increase as powers of wave number.
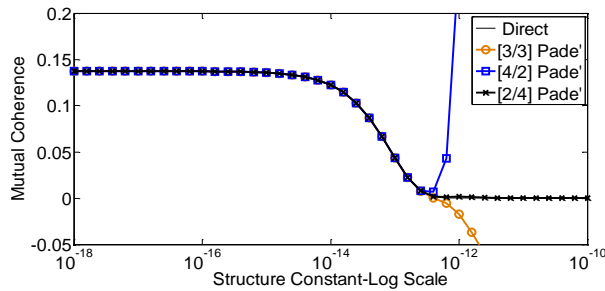
Fig. 2. Normalized mutual coherence function $\Gamma_2(\rho_c = 0.1, 0, L_z)/\Gamma_2(0, 0, L_z)$ versus structure constant C_n^2 for different Pade' orders.

Turning into another case, and shifting the expansion point toward larger fluctuations, namely, $C_{n0}^2 = 10^{-14}$, as shown in Fig. 3. An obvious observation is that power series does not approximate in larger fluctuations region, and has an almost constant error in smaller fluctuations region, which is expected, since the solution has almost constant slopes for smaller fluctuations. Amazingly enough, Pade' expansions agrees very well with direct solution for both small and large fluctuations. Simulation parameters given by $w_0 = 0.05\ m$, $L_z = 2.5\ km$, $\lambda = 630\ nm$, $[L/M] = [2/5]$.

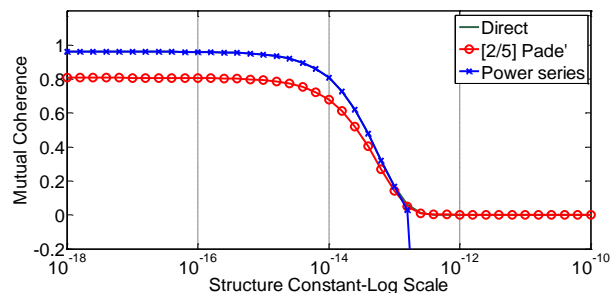Fig. 3. Mutual coherence function $\Gamma_2(0, 0, L_z)$ versus structure constant C_n^2 for a different expansion point $C_{n0}^2 = 10^{-14}$.

IV. CONCLUSION

Numerical methods offer a strong technique to compute the mutual coherence function of electromagnetic beam waves scattered by atmosphere. The efficiency is extremely increased when combined with Pade' approximation. It is shown that direct solution takes about 32 times the required time for Pade' expansion in presented simulations. In this research, it is also shown that a rational expansion with denominator degree larger than numerator gives much more accurate solution. Finally, rational asymptotes work very fine for small fluctuations region as well as for larger fluctuations.

REFERENCES

[1] V. Tatarski, *Wave Propagation in a Turbulent Medium*. McGraw-Hill Translation, 1961.

[2] A. Ishimaru, *Wave Propagation and Scattering in Random Media*. Academic Press, 1978.

[3] R. M. Manning, "Beam wave propagation within the second Rytov perturbation approximation," *Radiophysics and Quantum Electronics*, pp. 287-295, 1996.

[4] X. Zhu and J. Khan, "Performance bounds for coded free-space optical communications through atmospheric turbulence channels," *IEEE Transactions on Communications*, pp. 1233-1239, 2003.

[5] J. Jin, *The Finite Element Method in Electromagnetics*. Wiley, 2002.

[6] A. Taflove and S. Hagness, *Computational Electrodynamics: The Finite-Difference Time-Domain Method*. Artech House, 2005.

[7] T. Wang and D. Zhao, "Scattering theory of stochastic electromagnetic light waves," *Optics Letters*, vol. 35, 2010.

[8] M. Kuzuoglu and R. Mittra, "Finite element solution of electromagnetic problems over a wide frequency range via the Pade' approximation," *Comput. Methods Appl. Mech. Eng.*, pp. 263-277, 1999.

[9] N. Abu-Zaid and H. Tosun, "Finite element analysis of two dimensional electromagnetic scattering via Pade' approximation for complex permittivity," *Radio Science*, vol. 37, no. 1, 2002.

[10] G. Baker and B. Graves-Morris, *Pade' Approximants*. Cambridge Univ. Press, 1996.

Multi-Frequency T-Slot Loaded Elliptical Patch Antenna for Wireless Applications

S. Murugan, B. Rohini, P. Muthumari, and M. Padma Priya

Department of ECE, K.L.N. College of Engineering, Pottapalayam-630612, Sivagangai District, India
murugans1976@gmail.com, rohinibala17@gmail.com, muthumariengg@gmail.com,
padma.nmanickam@gmail.com

Abstract – In this paper, a multi frequency microstrip antenna (MSA) for wireless applications is designed. The proposed MSA comprised of elliptical patch antenna with T-slot. This antenna is fed by coaxial probe. The design parameters are major and minor axis of elliptical patch, length and width of T-slot and feeding point of probe. The proposed antenna can provide optimized multi frequency by varying the above design parameters. FR-4 substrate with dielectric constant 4.4 is chosen. The multi frequencies are 1.57 GHz, 1.96 GHz and 3.4 GHz, which covers the applications such as GPS and 4G LTE. The simulation of the antenna is performed using the ANSOFT HFSS and it is analyzed for S_{11} (dB) and radiation pattern. The prototype antenna is fabricated for optimized dimensions and tested using vector network analyzer. Simulation and experimental results are compared with each other.

Index Terms ─ Coaxial feed, elliptical patch antenna, multi-frequency, T-slot.

I. INTRODUCTION

The microstrip antenna (MSA) is one of the most preferred antenna structures for wireless applications and handheld devices. They are small in size, light weight and low volume. Generally, the multi-frequency MSA are divided into two categories: i) multi-resonator antennas and reactively loaded antenna. In the first category, the multi-frequency operation is achieved by means of multiple radiating elements, each supporting strong currents and radiation at its resonance. It includes the multilayer stacked-patch antennas using circular, annular, rectangular and triangular patches [1], [2]. A multi-resonator antenna in coplanar structures can also be fabricated by using aperture-coupled parallel microstrip dipoles [3]. As these antenna structures usually involve multiple substrate layers, they are of high cost. Large size is another drawback of the multi-resonator antenna, which makes it difficult for the antenna to be installed in hand-held terminals. The second category is reactively load MSA, to obtain multi frequency operation of the antenna such as multi-slotted patch, rectangular patch

with two T-slots, truncated circular patch with double U-slot, square spiral patch antenna and pi-shaped slot on rectangular patch [4-9]. These structures involve complex calculation, design, higher frequency ratio and lower bandwidth as compared to proposed antenna. Therefore, the proposed antenna consists of a simple T-slot which is loaded on the elliptical patch antenna and it is fed by coaxial probe. The dimensions of the proposed antenna are optimized using HFSS in such a way that it provides multi-frequency. The paper is organized as, proposed antenna design is discussed in Section 2, followed by simulation and experimental results in Section 3, and Section 4 concludes the paper.

II. ANTENNA DESIGN

The proposed antenna is shown in Fig. 1 (top view) and Fig. 2 (side view). The elliptical patch of semi major axis 'a' and semi minor axis 'b' is printed on the FR-4 substrate (ε_r=4.4). A T-slot of length 'l' and width 'w' is slotted in the elliptical patch.
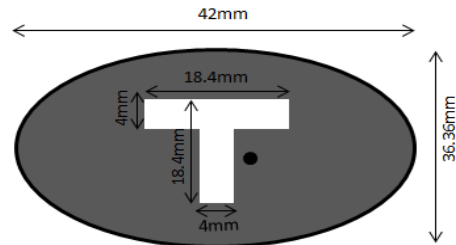

Fig. 1. Top view of proposed antenna.

Fig. 2. Side view of proposed antenna.

The resonant frequency of elliptical patch is given as [10]:

$$fr = \frac{c\sqrt{q}}{\pi a e \sqrt{\varepsilon_r}},\tag{1}$$

where c is velocity of light, $(3*10^8$ m/s), e is eccentricity of elliptical patch as:

$$e = \sqrt{1 - (b|a)^2}, \qquad (2)$$

where 'a' is the semi major axis of the elliptical patch, 'b' is semi minor axis of the elliptical patch, ε_r is dielectric constant of the substrate, q is the approximated Mathieu function of the given mode and eccentricity is calculated from [11]. The order of few modes of elliptical patch antenna is TM_{11} and TM_{21} and is based on the q value function. In this paper, TM_{11} mode is chosen. Then the q value for TM_{11} mode is given.

For the e values between 0 and 0.4:

$$q_{11} = 0.847e^2 - 0.0013e^3 + 0.0379e^4, \qquad (3)$$

For the e value between 0.4 and 1.0:

$$q_{11} = -0.0064e + 0.8838e^2 - 0.0696e^3 + 0.082e^4, \quad (4)$$

Here, the eccentricity of 0.5 is chosen and the canter frequency is taken as 2 GHz. By substituting the canter frequency and eccentricity values in the above equations, the dimensions of the elliptical patch can be calculated. The T-slot length and width can be determined by parametric study.

For TM_{11} mode, the theoretical value of resonant frequency for elliptical patch of semi major axis 21 mm is found to be 2 GHz. This is the theoretical resonant frequency value for elliptical patch without T-slot. The multi frequency resonance can be obtained by properly designing the length and width of the T-slot and also the feed point of the probe. This plays a major role in optimizing the frequency.

III. SIMULATION AND EXPERIMENTAL RESULTS

The simulation of the above designed antenna was performed using ANSOFT HFSS software. The FR-4 substrate size of 10 0mm*100 mm*1.6 mm is chosen as a dielectric material. Coaxial probe is used for exciting the patch. Return loss (dB) is defined as that the difference in dB between power sent towards antenna under test (AUT) and power reflected [12]. The requirement for reflection co-efficient of wireless devices specifies 10 dB return loss bandwidth.

The parametric study of the antenna is performed using Ansoft HFSS. Table 2 shows the S_{11} (dB) values for different lengths of T slot, keeping the width constant. As the length of the T-slot is increased, the resonant frequency is decreased. The length and width of the T-slot is chosen as 17.8 mm and 4 mm to obtain the optimized desired multi-frequency. A prototype is fabricated for the dimensions given in Table 1. The photograph of the antenna is shown in Fig. 3. The fabricated antenna is tested using vector network analyzer. Figure 4 shows the comparison of both simulated and measured S_{11} (dB) vs. frequency (GHz).

Table 3 shows the comparison of simulated and measured output.

Table 1: Design specifications

Design Parameters	Values
Semi major axis 'a'	21 mm
Semi minor axis 'b'	18.18 mm
Eccentricity 'e'	0.5
Substrate thickness h	1.6 mm
Dielectric constant 'ε_r'	4.4
Length of the T-slot 'l'	18.4 mm
Width of the T-slot 'w'	4 mm
Feed point	(6,3)

Table 2: Parametric study (w = 4 mm)

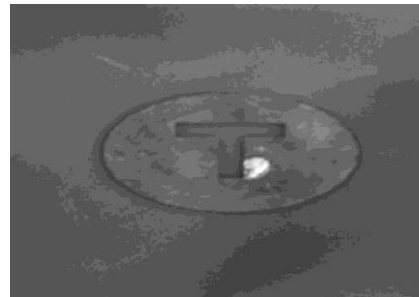| Length (mm)	Resonant Frequencies (GHz)	Corresponding S_{11} (dB)
17.8	1.57	-31.06
	1.96	-19.83
	3.43	-21.76
17	1.58	-18.73
	1.94	-24.6
	3.25	-15.73
16.5	1.63	-16.61
	1.99	-21.06
	3,41	-15.34
16	1.67	-14.26
	1.99	-20.79
	3.4	-12.28

Fig. 3. Photograph of fabricated antenna.
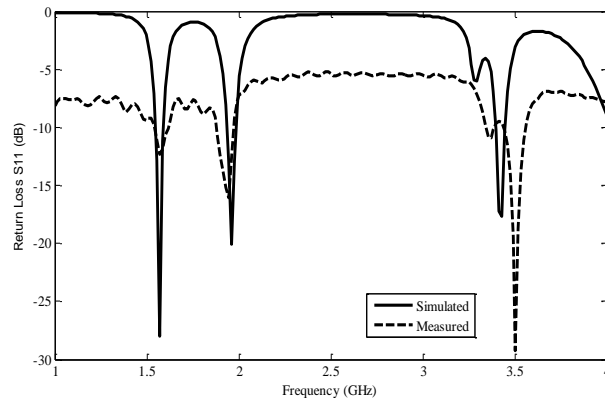

Fig. 4. S_{11} (dB) vs. frequency (GHz).

Table 3: Simulated and measured S_{11} (dB)

Parameters	Simulation Output	Measured Output
Frequency (GHz)	1.57 1.96 3.4	1.58, 1.946 3.5
S_{11} (dB)	-31.06, -19.83 and -21.76 respectively	-11.26, -16.36 and -19.09 respectively

Figure 5 shows the radiation pattern at 1.57 GHz. It is simulated using HFSS for φ=0° E plane (XZ plane) and φ=90° (YZ plane). The coordinate system is XYZ and the antenna is placed in XY plane. It has very low gain of -4.53 dB and HPBW of 84 degrees. Figure 6 shows the radiation pattern at 1.96 GHz has a very low gain of -0.3357 dB and HPBW of 84 degrees. Figure 7 shows radiation pattern at 3.43 GHz, which has butterfly pattern. FR4 is lossy substrate. The loss tangent of the substrate is high, which affects the performance of the antenna. It may be the reason for low gain. Moreover, size of the patch is very small. The gain of the antenna is directly proportional to cross sectional area of the patch. The T shaped slot is etched from this patch, which also reduces the area of the patch; hence, further reduces the gain. In [13], a truncated elliptical patch is discussed, which also has a low value of gain.
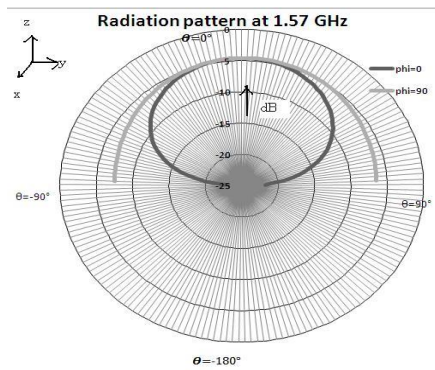

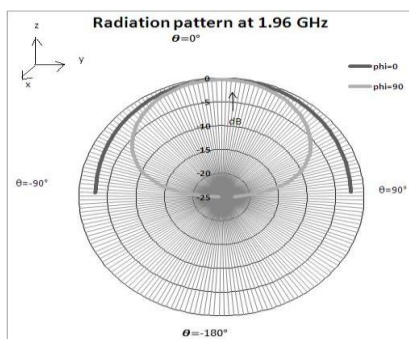
Fig. 5. Radiation pattern at 1.57 GHz.
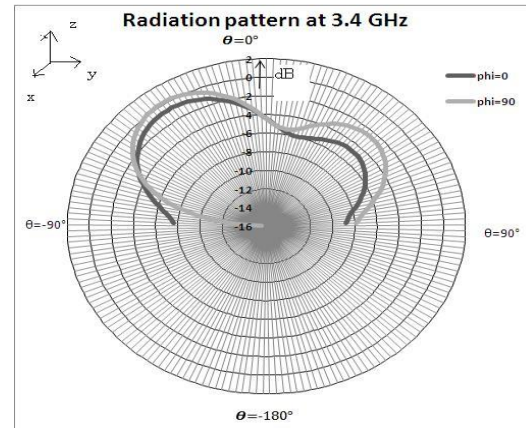

Fig. 6. Radiation pattern at 1.96 GHz.

Fig. 7. Radiation pattern at 3.4 GHz.

The proposed antenna may be compared with some of antennas given in references. It is described below. In [4], multi-slotted antenna of size 39.6 mm*47.9 mm is patched on the FR-4 substrate. The antenna is very large size and of very complicated structure than proposed antenna. The square patch of size 30 mm*30 mm with T-slot and defective ground structure is discussed in [5]. The antenna resonates at multi-frequency. The structure is simple but it does not provide better return loss than proposed antenna. In [6], the truncated circular patch of radius 40 mm is patched on the FR-4 substrate with double U slot. Air gap is introduced between the substrate and ground plane. Use of air gap may increases the size of the antenna. This air gap is avoided in the proposed antenna. In [7], the antenna of rectangular patch with size 23.4 mm*18.2 mm is patched on the substrate and two T-slots are made on this patch. The antenna is fed by microstrip feed. But this antenna uses two T-slots to produce multi frequency. In [8], the square spiral patch antenna with size 33.7 mm*33.7 mm is patched on the FR-4 substrate. The antenna is of very large size and complicated design as compared to proposed antenna.

The proposed antenna is an elliptical shape of semi major axis 21 mm and semi minor axis 18.18 mm with T-slot on the elliptical patch. The antenna is fed by coaxial feed. There is no air gap. FR4 substrate of thickness 2 mm with dielectric constant 4.4 is chosen. The antenna structure is simple and provides better return loss.

IV. CONCLUSION

From the analysis, it is concluded that the proposed antenna resonates at three different frequencies. The frequencies are 1.57 GHz, 1.96 GHz and 3.4 GHz, which has S_{11} (dB) of -31.06, -19.83 and -21.76 respectively. The frequencies can cover applications such as GPS and LTE. Simulation and measurement results are presented for validation of the design and slight deviation is observed, which is below the tolerable limit of 5%. It

is due to substrate, connector losses and fabrication tolerances.

REFERENCES

[1] J. S. Dahele, K. F. Lee, and D. P. Wong, "Dual-frequency stacked annular- ring microstrip antenna," *IEEE Trans. Antennas Propagat.*, vol. AP-35, no. 11, pp. 1281-1285, 1987.

[2] J. Wang, R. Fralich, C. Wu, and J. Litva, "Multifunctional aperture-coupled stacked antenna," *Electron. Lett.*, vol. 26, no. 25, pp. 2067-2068, 1990.

[3] F. Croq and D. M. Pozar, "Multifrequency operation of microstrip antennas using aperture coupled parallel resonators," *IEEE Trans. Antennas Propagat.*, vol. 40, no. 11, pp. 1367-1374, 1992.

[4] S. Natarajamani, S. K. Behera, and R. K. Mishra, "Design of multi slotted and multfrequency patch antenna," *Applied Electromagnetic Conference*, Kolkata, 2009.

[5] S. De, P. Samaddar, S. Sarkar, S. Biswas, et al., "Compact high gain multi-frequency microstrip antenna," *International Journal of Soft Computing and Engineering*, vol. 2, no. 6, 2013.

[6] S. Murugan, E. Sathish Kumar, and V. Rajamani, "Design and analysis of double U slot loaded dual frequency microstrip antenna," *Progress In Electromagnetics Research C*, vol. 45, pp. 101-112, 2013.

[7] D. Ramya Keertana, M. V. S. D. N. N. Murthy, B. Yeswanth, et al., "A novel multi frequency rectangular microstrip antenna with dual T shaped slots for UWB applications," *IOSR Journal of Electronics and Communication Engineering*, vol. 9, no. 1, 2014.

[8] A. Ghosal, A. Majumdar, S. Kumar Das, A. Das, "Wideband and multi-frequency square spiral microstrip patch antenna," *International Journal of Innovative Research in Computer and Communication Engineering*, vol. 3, no. 2, 2015.

[9] S. Das, P. P. Sarkar, and S. K. Chowdhury, "Modified π-shaped slot loaded multi-frequency microstrip antenna," *Progress In Electromagnetics Research B*, vol. 64, pp. 103-117, 2015.

[10] J. A. Ansari, K. Kumari, A. Singh, and A. Mishra, "Ultra-wideband co-planer microstrip patch antenna for wireless applications," *Wireless Pers. Commun.*, vol. 69, pp. 1365-1378, 2013.

[11] I. J. Bahl and P. Bhartia, *Broadband Microstrip Patch Antennas*. Dedham: Artech House, 1980.

[12] T. S. Bird, "Definition and the misuse of return loss," *IEEE Antenna Propagation Magazine*, vol. 51, no. 2, pp. 166-167, 2009.

[13] P. Sekra, S. Shekhawat, et al., "Design of circularly polarised truncated elliptical patch antenna with improved performance," *Indian Journal of Radio & Space Physics*, vol. 40, pp. 227-233, 2011.