APPLIED COMPUTATIONAL ELECTROMAGNETICS SOCIETY JOURNAL

Special Issue on Antennas and Big Data in Electromagnetics

Guest Editors: Wenhua Yu, Rushan Chen, Lei Zhao,
Wenxing Li, and Mingyao Xia

May 2017
Vol. 32 No. 5
ISSN 1054-4887

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

The illustrations on the front cover have been obtained from the research groups at the Department of Electrical Engineering, The University of Mississippi.

THE APPLIED COMPUTATIONAL ELECTROMAGNETICS SOCIETY
http://aces-society.org

EDITOR-IN-CHIEF

Atef Elsherbeni
Colorado School of Mines, EE Dept.
Golden, CO 80401, USA

ASSOCIATE EDITORS-IN-CHIEF

Sami Barmada
University of Pisa. ESE Dept.
Pisa, Italy, 56122

Mohamed Bakr
McMaster University, ECE Dept.
Hamilton, ON, L8S 4K1, Canada

Antonio Musolino
University of Pisa
56126 Pisa, Italy

Mohammed Hadi
Kuwait University, EE Dept.
Safat, Kuwait

Abdul Arkadan
Marquette University, ECE Dept.
Milwaukee, WI 53201, USA

Marco Arjona López
La Laguna Institute of Technology
Torreon, Coahuila 27266, Mexico

Alistair Duffy
De Montfort University
Leicester, UK

Paolo Mezzanotte
University of Perugia
I-06125 Perugia, Italy

EDITORIAL ASSISTANTS

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

Shanell Lopez
Colorado School of Mines, EE Dept.
Golden, CO 80401, USA

EMERITUS EDITORS-IN-CHIEF

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

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

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

Ahmed Kishk
Concordia University, ECS Dept.
Montreal, QC H3G 1M8, Canada

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

EMERITUS ASSOCIATE EDITORS-IN-CHIEF

Yasushi Kanai
Niigata Inst. of Technology
Kashiwazaki, Japan

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

Levent Gurel
Bilkent University
Ankara, Turkey

Mohamed Abouzahra
MIT Lincoln Laboratory
Lexington, MA, USA

Ozlem Kilic
Catholic University of America
Washington, DC 20064, USA

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

Fan Yang
Tsinghua University, EE Dept.
Beijing 100084, China

EMERITUS EDITORIAL ASSISTANTS

MAY 2017 REVIEWERS

THE APPLIED COMPUTATIONAL ELECTROMAGNETICS SOCIETY

JOURNAL

Vol. 32 No. 5 May 2017

TABLE OF CONTENTS

Symbolic Regression for Derivation of an Accurate Analytical Formulation Using "Big Data": An Application Example

Peyman Mahouti [1], Filiz Güneş [1], Mehmet A. Belen [2], and Salih Demirel [1]

[1] Department of Electronic and Communication Engineering
Yıldız Technical University, Istanbul, Turkey

[2] Department of Electric and Electronic Engineering
Artvin Çoruh University, Istanbul, Turkey
pmahouti@yildiz.edu.tr, gunes@yildiz.edu.tr, mehmetalibelen@artvin.edu.tr, salihd@yildiz.edu.tr

Abstract **─** With emerging of the Big Data era, sample datasets are becoming increasingly large. One of the recently proposed algorithms for Big Data applications is Symbolic Regression (SR). SR is a type of regression analysis that performs a search within mathematical expression domain to generate an analytical expression that fits large size dataset. SR is capable of finding intrinsic relationships within the dataset to obtain an accurate model. Herein, for the first time in literature, SR is applied to derivate a full-wave simulation based analytical expression for the characteristic impedance Z_0 of microstrip lines using Big Data obtained from an 3D-EM simulator, in terms of only its real parameters which are substrate dielectric constant ε, height h and strip width w within 1-10 GHz band. The obtained expression is compared with the targeted simulation data together with the other analytical counterpart expressions of Z_0 for different types of error function. It can be concluded that SR is a suitable algorithm for obtaining accurate analytical expressions where the size of the available data is large and the interrelations within the data are highly complex, to be used in Electromagnetic analysis and designs.

Index Terms **─** Big Data application, characteristic impedance, microstrip line, Symbolic Regression.

I. INTRODUCTION

The emerging of the big data era poses a challenge to traditional machine learning algorithms. In many research area, big data is a collection of datasets so large and complex that they become difficult to process using available database management tools or traditional data processing applications. Big data is usually composed of datasets with sizes beyond ability of the commonly used software tools, which are unable to capture, curate, manage, or process such data within a reasonable elapsed time [1-2]. Thus, a challenging problem is being faced in terms of hardware capabilities for the traditional algorithms where the training time for the algorithm becomes very lengthy which would significantly decreases the efficiency of the system design optimization process. To compensate the requirements of today's data analytics, many methods and tools have been lately established. Various systems have been developed mainly by the industry to support Big Data analysis, including Google's Map Reduce, Yahoo's PNUTS, Microsoft's SCOPE, and Walmart Labs Muppet. Each of these tools possess a unique algorithm for classification, clustering, and collaborative filtering, dimension reduction, miscellaneous approximations. Recently, a proposal has been put forward for a new machine learning tool based on Symbolic Regression (SR): Eureqa for Big Data applications [3-10]. Eureqa is a breakthrough technology that uncovers the intrinsic relationships hidden within complex data. Traditional machine learning techniques like neural networks and regression machines are capable tools for prediction, but become impractical when "solving the problem" involves understanding how you arrive at the answer. Eureqa uses a machine learning technique called Symbolic Regression to unravel the intrinsic relationships in data and describe them in simple mathematical equations. By using Symbolic Regression (SR), Eureqa can create incredibly accurate predictions that are easily explained and shared with others.

SR is a type of regression analysis that performs a search within mathematical expressions domain to create a model that fits to a given dataset. Commonly in the starting point of a search, there is not a particular model, not unless one is given by the user. Instead, initial expressions are formed by randomly combining mathematical blocks such as mathematical operators, analytic functions, constants, and variables. New equations are then formed by recombining previous equations, using genetic programming. Since SR does not require a specific model, it is not affected by users or unknown gaps in problem domain. The software uses

Submitted On: December 30, 2016
Accepted On: March 19, 2017

evolutionary search to determine mathematical equations that describe sets of data in their simplest form. This class of algorithms are based on Darwinian Theory of evolution and one of its main attributes is that there is no calculated single solution, but a class of possible solutions at once. This class of possible and acceptable solutions is called "population". Members of this population are called "individuals" and mathematically said, they represent possible solution, i.e., solution which can be realized in real world application. Main aim of evolutionary algorithms is to find the best solution of all during evolutionary process. Evolutionary algorithms differ among themselves in many points of view like individual representation (binary, decimal) or offspring creation (standard crossover, arithmetic operations, vector operations, etc.).

In this work, the aim is briefly to use Symbolic Regression SR in processing Big Data to obtain a full - wave Simulation-based analytical expression for the characteristic impedance Z_0 of a microstrip line in terms of only its real parameters to the contrary of the counterpart analytical expressions which require additional calculations for an effective dielectric constant ε_{eff} for different w/h ratios. In fact, there have been many different proposed equation sets for the calculation of Z_0. The most commonly used expressions are proposed by Pozar [11] and Balanis [12]. However, in all of these expressions, Z_0 is calculated using an effective dielectric permittivity ε_{eff} through the helping relationships obtained with the empirical methods.

To obtain "Big Data", an efficient and fast 3D simulation tool is needed for gathering the required data for training and validation in the modelling process. Sonnet's suites [13] of high-frequency electromagnetic (EM) software are aimed at today's demanding design challenges involving predominantly planar (3D planar) circuits and antennas such as microstrip, stripline, coplanar waveguide, multi/single layer PCB and combinations with vias, vertical metal sheets (z-directed strips), and any number of layers of metal traces embedded in stratified dielectric material. By using Sonnet, it is possible to obtain a high number of samples for modelling the Z_0 of a given substrate (h, ε_r) for different values of the lines' widths in a very short period of time. After obtaining the required data from Sonnet's suites, the data were inputted into the SR in the "Eureqa" environment for modelling characteristic impedance of microstrip lines.

In the study case section, all the data obtained by using Sonnet will be applied to the expressions given by [13-14] and SR of the Eureqa, respectively. Then, their performances are compared according to four commonly used error metrics, which are Mean Absolute Error (MAE), Relative Mean Absolute Error (RMAE), Maximum Error within the whole test data (MXE) and Root Mean Squared Error (RMSE). Furthermore, comparison between the approaches is also presented as figures for different values for h, ε_r and w within 1 GHz-10 GHz bandwidth. The flowchart of the methodology for obtaining a full-wave simulation based analytical expression for Z_0 is presented in Fig. 1. The paper is organized as follows: In Section 2, a brief explanation about Z_0 and the current mostly used analytical expressions in literature are presented. The Section 3 presents a general overview of the electromagnetic simulations and the obtained data. In Section 4, a brief explanation about working mechanism of SR and some example of its usage in science and engineering fields is presented alongside of the methodology that is implemented to forming the full - wave Simulation-based analytical expression for Z_0. Section 5 gives results of the study case. Finally, the paper ends with the conclusions in the Section 6.

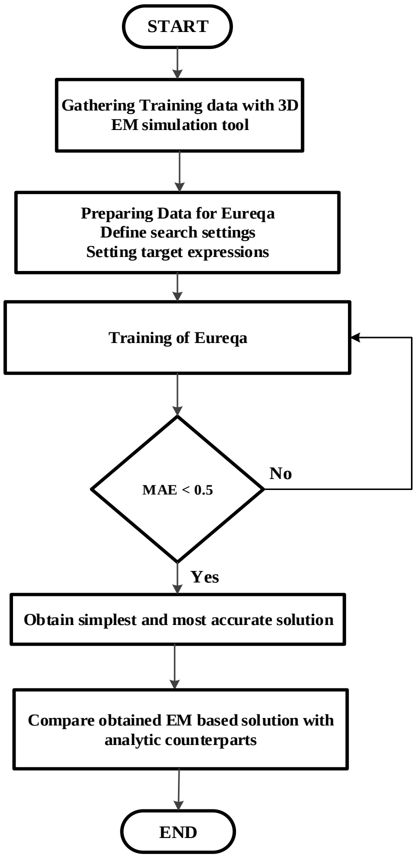


Fig. 1. Flowchart of the proposed methodology.

II. ALTERNATIVE FORMULATION OF THE CHARACTERISTIC IMPEDANCE OF MICROSTRIP LINES

Characteristic impedance is an important design parameter in microstrip-based microwave circuit designs such as filter, antenna, power divider, coupler, oscillator, and amplifier applications. Also, it is essential in the design of interdigital capacitance and stub inductance.

There are many analytical expressions for the calculation of microstrip transmission lines of Z_0. The well-known equations which are credited to Pozar [11] and Balanis [12] are the most commonly used expressions. These expressions require pre-calculation of ε_{eff} and some sub-equations, where the main equation of Z_0 changes depending on the ratio of transmission lines widths (w) and substrate's height (h). In [11], the solution only depends on the pre-calculation of ε_{eff} and the condition of w/h ratio, the effect of the frequency is not taken into account, either. On the other hand, in addition to the pre-calculation of ε_{eff} and the condition of w/h ratio, the effect of the transmission line thickness T and the frequency were taken into the account in [12]. In our study, the expressions in [11] and [12] are presented in 4 different cases for benchmarking purposes.

Case 1

The analytical expression in [11] is given in Eqs. (1-3), where the expression depends on the ratio of w/h and requires the pre-calculation of Eq. (1). For the sake of simplicity, Eqs. (1-3) will hereafter be named Case 1:

$$\varepsilon_{eff} = \frac{\varepsilon_r + 1}{2} + \frac{\varepsilon_r - 1}{2}\left(\left(1 + 12\frac{h}{w}\right)^{-\frac{1}{2}}\right). \tag{1}$$

For $w/h \leq 1$,

$$Z_0 = \frac{60}{\sqrt{\varepsilon_{eff}}}\ln\left(\frac{8h}{w} + \frac{w}{4h}\right). \tag{2}$$

For $w/h \geq 1$,

$$Z_0 = \frac{120\pi}{\sqrt{\varepsilon_{eff}}\left(\mathrm{w/h} + 1.393 + 0.667\ln\left(\mathrm{w/h} + 1.444\right)\right)}. \tag{3}$$

Case 2

In [12], a more detailed expression set is provided for analytical calculation of Z_0. Firstly, all the expressions are solved for f=10 GHz conditions, then the frequency-dependent expression is calculated by using previous values as sub-expressions. In Eqs. (4-5), the effect of the transmission line thickness is considered in the calculation of the effective width value of microstrip transmission line for f=10 GHz. Hereafter, Eqs. (4-12) will be named Case 2:

for $w/h \leq 1$,

$$\frac{w_{eff}\left(0\right)_{f=0}}{h} = \frac{w}{h} + \frac{1.25}{\pi}\frac{T}{h}\left(1 + \ln\left(\frac{2h}{T}\right)\right). \tag{4}$$

For $w/h \geq 1$,

$$\frac{w_{eff}\left(0\right)_{f=0}}{h} = \frac{w}{h} + \frac{1.25}{\pi}\frac{T}{h}\left(1 + \ln\left(\frac{4\pi w}{T}\right)\right). \tag{5}$$

After calculating the effective width value of the transmission line, Z_0 (f=0) can be obtained depending on the ratio of $w_{eff}(0)/h$ as follows:

For $\frac{w_{eff}\left(0\right)}{h} \leq 1$,

$$\varepsilon_{eff}\left(0\right) = \frac{\varepsilon_r + 1}{2} + \frac{\varepsilon_r - 1}{2}\left(\left(1 + 12\frac{h}{w_{eff}\left(0\right)}\right)^{-\frac{1}{2}} + 0.04\left(1 - \frac{w_{eff}\left(0\right)}{h}\right)^2\right), \tag{6}$$

$$Z_0\left(0\right)_{f=0} = \frac{60}{\sqrt{\varepsilon_{eff}\left(0\right)}}\ln\left(\frac{8h}{w_{eff}\left(0\right)} + \frac{w_{eff}\left(0\right)}{4h}\right), \tag{7}$$

For $\frac{w_{eff}\left(0\right)}{h} \geq 1$,

$$\varepsilon_{eff}\left(0\right) = \frac{\varepsilon_r + 1}{2} + \frac{\varepsilon_r - 1}{2}\left(1 + 12\frac{h}{w_{eff}\left(0\right)}\right)^{-\frac{1}{2}}, \tag{8}$$

$$Z_0\left(0\right) = \frac{\dfrac{120\pi}{\sqrt{\varepsilon_{eff}\left(0\right)}}}{\dfrac{w_{eff}\left(0\right)}{h} + 1.393 + 0.677\ln\left(\dfrac{w_{eff}\left(0\right)}{h} + 1.444\right)}. \tag{9}$$

Finally, the effect of the frequency is taken into the account in Eq. (12) by using sub-equations given in Eqs. (10-11):

$$\varepsilon_{eff}\left(f\right) = \varepsilon_r - \frac{\varepsilon_r - \varepsilon_{eff}\left(0\right)}{1 + \dfrac{\varepsilon_{eff}\left(0\right)}{\varepsilon_r}\left(\dfrac{f}{f_t}\right)^2}, \tag{10}$$

$$f_t = \frac{Z_0\left(0\right)}{2\mu_0 h}, \tag{11}$$

$$Z_0\left(f\right) = Z_0\left(0\right)\sqrt{\frac{\varepsilon_{eff}\left(0\right)}{\varepsilon_{eff}\left(f\right)}}. \tag{12}$$

Case 3

In examples given in [12], the T/h ratio is commonly considered equal to zero. In our study, we also ignored the effect of T/h ratio in Case 3 for the sake of better benchmarking results. Thus, Eqs. (4-5) are simplified as follows:

for $w/h \leq 1$ and T/h=0,

$$\frac{w_{eff}\left(0\right)}{h} = \frac{w_{eff}\left(f=0\right)}{h} = \frac{w}{h}. \tag{13}$$

For $w/h \geq 1$ and T/h=0,

$$\frac{w_{eff}\left(0\right)}{h} = \frac{w_{eff}\left(f=0\right)}{h} = \frac{w}{h}. \tag{14}$$

In our studies, we have observed that when an EM-based simulation result is taken as the reference point, Case 2 can be improved by considering T/h ratio equal to zero. The results can be clearly seen in Figs. 4-6 and Tables 4 and 5.

Case 4

As it can be seen in Figs. 4-6, commonly, the result of Sonnet for Z_0 slightly increases with the frequency. However, in Cases 2 and 3, $Z_0(f)$ is inversely proportional with the frequency. Thus, by ignoring the frequency effect in Eq. (12), Case 4 can be expressed as follows where its performance is better than Cases 2 and 3:

$$Z_0(f) = Z_0(0) \quad for \quad \sqrt{\frac{\varepsilon_{eff}(0)}{\varepsilon_{eff}(f)}} = 1, \; \frac{T}{h} = 1. \tag{15}$$

In Eqs. (1-15) the analytical expressions of two well-known references had been given in 4 different cases. Although these have admissible results, they require:

- Pre-calculation of ε_{eff} ;
- Condition of w/h ratio;
- Additional sub-equations.

In the next sections, by using Sonnet and Eureqa, a new EM based analytical expression for calculation of Z_0 is presented, which is independent from the ratio of w/h and does not require pre-calculation of ε_{eff} as well as being more simple and accurate than the counterpart analytical expressions given in Cases 1-4.

III. ELECTROMAGNETIC SIMULATIONS WITH SONNET'S SUITS

Sonnet's suites of high frequency EM software are aimed at today's demanding design challenges involving predominantly planar (3D planar) circuits [13]. The software requires a physical description of your circuit and employs a rigorous Method-of-Moments EM analysis based on Maxwell's equations that includes all parasitic, cross-coupling, enclosure and package resonance effects. Thus, it is a very effective tool for calculating the Z_0 of a microstrip lines.

In our study, a series of simulations are performed in Sonnet's suites with a range of parameters for the combination of each microstrip parameters such as substrate height h, dielectric permittivity ε_r, widths of the microstrip line w and frequency f. The values given in Tables 1 and 2 were simulated with Sonnet to obtain a high accuracy rate in the training process. Training data is given in Table 1 while Table 2 contains the test data including the most commonly used substrates (FR4, Rogers, Duroid). Tables 3-6 presents the variations of characteristic impedance Z_0 according to the microstrip parameters of width w, operation frequency f, substrate height h and the dielectric constant ε_r. As it can be observed from these tables, the most effective parameters can be ordered as: w, h, ε_r while the effect of operation frequency f can be ignored.

Some typical examples of the training data are given in Table 7. Due to the training features of Eureqa normalized with 50 ohm are used. Also, in the study case, all the data in Tables 1 and 2 were used for performance benchmarking of the equations given in Cases 1-4 and the expression obtained by Eureqa. The results of the benchmarking are given in Tables 8 and 9.

Table 1: Training data

Parameters	Value	Sample Number
h (mm)	0.3, 0.5, 0.7, 1, 1.5, 2, 2.5	7
ε_r	1, 2, 3, …., 6	6
w (mm)	0.2, 0.25, 0.3,…, 1, 1.2, 1.4, 1.6, 1.8, 2, 2.2, 2.5, 2.7, 3, 3.3, 3.5, 3.7, 4, 4.3, 4.5, 5	33
Thickness (µm)	35	1
f (GHz)	1, 2, 3, …., 10	10
Total samples	7x7x33x1x10=13860	

Table 2: Test data

Parameters	Duroid 5880	Duroid 6006	FR4	Roger 3003	Roger 3006
h (mm)	0.254 0.784 3.175	0.254 1.27 2.5	1.6	0.25 0.75 1.52	0.25 0.64 1.27
ε_r	2.2	6.15	4.6	3	6.15
w (mm)	0.2, 0.25, 0.3,…, 11.2, 1.4, 1.6, 1.8, 2, 2.2, 2.5, 2.7, 3, 3.3, 3.5, 3.7, 4, 4.3, 4.5, 5				
Thickness T (µm)	35				
f (GHz)	1, 2, 3, 4, 5, 6, 7, 8, 9, 10				
Sample number	990	990	330	990	990
Total samples	4290				

Table 3: Variation of Z0 due to the effect of microstrip width w for (h=1.5 mm, εr=3, f=5 GHz)

h (mm)	ε_r	w (mm)	f (GHz)	$Z_0/50$ (Sonnet)
1.5	3	0.2	5	3.417
1.5	3	0.3	5	3.058
1.5	3	0.5	5	2.614
1.5	3	1	5	2.024
1.5	3	1.6	5	1.639
1.5	3	3.5	5	1.058
1.5	3	5	5	0.834

Table 4: Variation of Z0 due to the frequency f for (h=1.5 mm, εr=3, w=1 mm)

h (mm)	ε_r	w (mm)	f (GHz)	$Z_0/50$ (Sonnet)
1.5	3	1	1	2.015
1.5	3	1	3	2.019
1.5	3	1	5	2.024
1.5	3	1	7	2.029
1.5	3	1	10	2.035

Table 5: Variation of Z0 due to the effect of substrate height h for (εr=3, w=1 mm, f=5 GHz)

h (mm)	ε_r	*w* (mm)	*f* (GHz)	Z_0/50 (Sonnet)
0.3	3	1	5	0.836
0.5	3	1	5	1.164
0.7	3	1	5	1.410
1	3	1	5	1.691
1.5	3	1	5	2.024
2	3	1	5	2.263
2.5	3	1	5	2.446

Table 6: Variation of Z0 due to the effect of substrate dielectric constant εr for (h=1.5 mm, w=1 mm, f=5 GHz)

h (mm)	ε_r	*w* (mm)	*f* (GHz)	Z_0/50 (Sonnet)
1.5	1	1	5	2.940
1.5	2	1	5	2.359
1.5	3	1	5	2.024
1.5	4	1	5	1.801
1.5	5	1	5	1.637
1.5	6	1	5	1.511

Table 7: Example of the data input into Eureqa

ID	*f* (GHz)	*h* (mm)	ε_r	*w* (mm)	Z_0 (Sonnet)	Z_0/50
1	1	0.3	1	0.4	111.729	2.235
2	1	1.5	5	1.4	70.272	1.405
3	1	2.5	6.15	3	56.5	1.130
4	2	1	5	3.5	32.120	0.642
5	3	3.175	2.2	2	112.75	2.255
6	4	2.5	6	5	43.321	0.866
7	6	0.75	3	1	73.2	1.464
8	7	0.5	2	0.75	81.149	1.623

IV. OBTAINING THE EM-SIMULATION BASED ANALYTICAL EXPRESSION WITH SYMBOLIC REGRESSION

In this section, firstly, a brief introduction about SR and its usage in various applications in science and engineering are provided. After that, the implementation of the SR's training process for obtaining the EM-based analytical expression of Z_0 is presented.

A. Symbolic Regression with Eureqa

As it has been mentioned before, Eureqa [14] is a breakthrough technology that uncovers the intrinsic relationships hidden within complex data using SR. As referenced in [15], Schmidt and Lipson developed a computing system that could observe a phenomenon in nature and then automatically identify various laws of nature and invariant equations that it obeyed. For example, they had the system to observe a double pendulum swinging chaotically using motion-tracking. Without any knowledge of physics or geometry, the system identified the exact energy conservation and momentum relations that governed its dynamics. In [16], Ceperic, Bako and Baric presented the use of Symbolic Regression for the black-box modelling of non-linear dynamic behaviour in the AC/DC rectifiers of radio-frequency identification (RFID) circuits. Also, there are many recent publications in the literature which used Eureqa's SR algorithm in order to solve or obtain solutions for various problems in astronomy, biology, chemistry, computer science, physics and many other fields [15-17].

Symbolic Regression is based on evolutionary algorithms. This type of algorithms are inspired from Darwinian theory of evolution. In evolutionary algorithms there is no calculated single solution. Instead, there is a class of possible solutions called as "population" at once. Individuals of this population represents the possible solution of real world problem. Main aim of these type of algorithms is to find the optimal solution during evolutionary process.

SR is a process in which the given dataset fitted by suitable mathematical formulas. This process is commonly used when some data of unknown process are obtained. For the first time, the idea of how to solve various problems using SR by means of evolutionary algorithms (EAs), come from Koza who used genetic algorithm (GA) in so called genetic programming (GP) [15-16]. GP is basically SR which is done by evolutionary algorithms. Main principle of GP is such that expressions are represented in chromosomes like syntactic trees. The syntactic tree form of Eq. (16) is given in Fig. 2 for a better understanding. Based on GA principles, new individuals (Childs) whose representations are in fact new expressions which are evaluated by fitness are created either by random generators or by exchanging parent's parts by crossover operators, or mutation. This process can be simply seen in Fig. 3, where Eqs. (16-17) are taken as parents and Eqs. (18-20) are the children of these parents through cross over or mutation operations:

$$P_1 = f(a,b,c,d,e) = \cos(ab) + \sqrt{c-d}, \qquad (16)$$

$$P_2 = f(a,b,c,d,e) = a^2 + \frac{c}{b-d} - 12e, \qquad (17)$$

$$C_1 = \cos(ab) - 12e, \qquad (18)$$

$$C_2 = a^2 + \frac{c}{b-d} + \sqrt{c-d}, \qquad (19)$$

$$C_3 = \sin\left(\frac{a}{b}\right) + \sqrt{c-d}. \qquad (20)$$

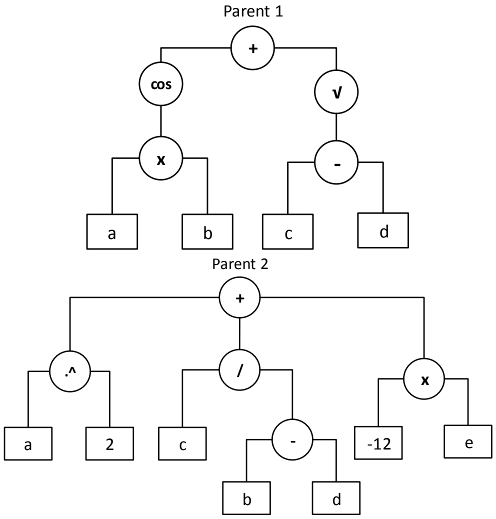
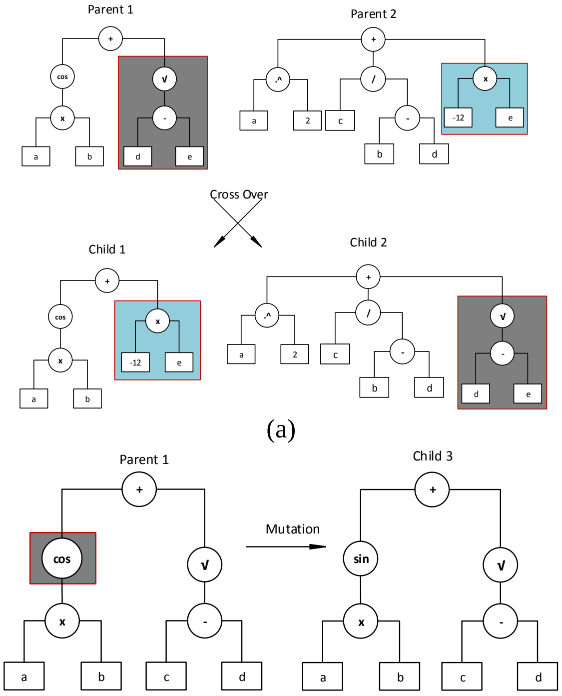


Fig. 2. Examples for Syntactic Tree.



(a)

(b)

Fig. 3. An example of (a) crossover and (b) mutation in syntactic tree form.

B. Implementation process in Eureqa for Z_0

As mentioned before, SR is a process in which the given dataset is fitted by suitable mathematical formulas. For this mean, firstly the data in Tables 1 and 2 were input into Eureqa as the training data with the following commands to start the search settings for creating a fitted mathematical /analytical expression of Z_0:

- Shuffle and split data points equally for training and test purposes;
  - Use all basic formula blocks;
  - Use all trigonometry blocks;
  - Use all exponential blocks;
  - Use all squashing blocks;
  - Use mean absolute error as error metric (default).
  - Target expression is expressed as follows:

$$Z_0 = f\left(h, \varepsilon_r, w, f\right). \tag{21}$$

Unfortunately, Eureqa did not provide an appropriate expression as it was expected. By checking the results of Sonnet given in Table 4, one can observed that the effect of frequency can be ignored as compared with the effects of *h*, ε_r and *w*. Thus, frequency is discarded from the target expression in Eq. (21) and some changes are made in Eureqa's searching commands as follows:

- Treat all data equally both for training and test;
- Use all basic formula blocks;
- Use all trigonometry blocks;
- Use all exponential blocks;
- Use only Gaussian from squashing blocks;
- Use Mean Absolute Error MAE as error metric (default).
- In order to increase the complexity of the predicted expression user can define the target expression as the sum of two or more functions in case of need. Here Z0 is defined as the sum of the two ingredients as follows as target expression:

$$Z_0 = f_0\left(h, \varepsilon_r, w\right) + f_1\left(h, \varepsilon_r, w\right). \tag{22}$$

Thus, by using SR algorithm Eureqa builds up an expression consisting of summation of the two different functions as given in Eq. (22), which is expected to give more accurate results than a single function. The training stage of Eureqa was performed by the Intel Core i5 CPU, 3.3 GHz Processor, 4 GB RAM. After approximately 10 hours of training, the MAE value dropped by less than 0.5, which was chosen as the target error value, therefore, we stopped the training process. Eureqa sorts the candidate's expression according to their complexity and fitness value of expressions. Eureqa's sorting algorithm suggests that the expression given in Eq. (23) is the optimal expressions:

$$Z_0 = 3.5 - 0.725\varepsilon_r + \frac{112.5}{w/h + 2.47} + \frac{300}{1 + 0.217\varepsilon_r + (\varepsilon_r + 0.57)(w/h) + 0.87e^{-\left(0.22 + \frac{0.75}{\varepsilon_r + w} + \frac{0.09}{w/h}\right)^2}}. \tag{23}$$

V. CASE STUDY

For the case study purposes, in Table 9, performance benchmarking of each analytic expression are presented for the data examples given in Table 7. Also, approximately 20.000 samples given in Tables 1 and 2 were taken as testing instances for each of the Cases 1-4 and the solution in Eq. (12). Performance results given in Table 5 were calculated by using the following error metrics:

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|T_i - P_i\right|, \tag{24}$$

$$RMAE = \frac{1}{N}\sum_{i=1}^{N}\frac{\left|T_i - P_i\right|}{\left|T_i\right|}, \tag{25}$$

$$MXE = \max\left(\left|T_i - P_i\right|\right), \tag{26}$$

$$RMSE = \frac{1}{N}\sum_{i=1}^{N}\sqrt{\left(T_i - P_i\right)^2}, \tag{27}$$

where, *T*: Target, *P*: Predicted, *N*: total sample.

Table 8: Z_0 results for samples in Table 3

ID	Z_0	Eureqa	Case 1	Case 2	Case 3	Case 4
1	111.7	112.3	110.6	103.9	110.6	110.6
2	70.2	70.2	70.4	67.6	69.3	69.3
3	56.5	56.5	55.7	54.8	55.6	55.7
4	32.1	32.1	31.8	31.3	31.7	31.8
5	112.7	115.6	116.3	113.8	115.7	116.2
6	43.3	42.8	42.3	40.5	40.9	42.3
7	73.2	73.1	72.6	70.1	72.4	72.6
8	81.1	80.9	80.8	77.4	80.7	80.8

Table 9: Performance benchmarking of Cases 1-4 and Eureqa

	MAE	RMAE	MXE	RMSE
Eureqa	0.4942	0.0078	7.1105	0.0061
Case 1	1.0197	0.0117	8.3851	0.0116
Case 2	4.2904	0.0513	24.292	0.0404
Case 3	1.6788	0.0214	14.642	0.0178
Case 4	1.1455	0.0128	9.9481	0.0127

Furthermore, in Figs. 4-6, benchmarking results of the 4 cases and SR applied to the 3 different substrates are given as compared to targeted values obtained with Sonnet within 1 GHz-10 GHz bandwidth for 11 different w values. Here the test data given in the Table 2 is used. Figure 4 depicts all the Z_0 results for FR4, while Fig. 5 gives the results for Duroid 6006 and Fig. 6 shows the results for Roger 3003. In order to depict all the results of the different w values for the different cases within 1-10 GHz bandwidth in a single plane, all the results are compressed in x-axis for each value of w. On the x-axis, each sample in the w-band is taken with 1 GHz intervals using Sonnet suit.

As it can be seen from the figures, the values of Z_0 obtained by Sonnet slightly increase with the frequency on the contrary to Case 3 and 4. As a consequence, the proposed EM-based analytical expression for Z_0 is better than the other expressions. Also, it should be noted that the proposed expression is not only better in performance benchmarking, but it is also, simpler than the other expressions because it does not require pre-calculation of ε_{eff} as well as it does not require different expressions for the ratio of the *w*/*h*.

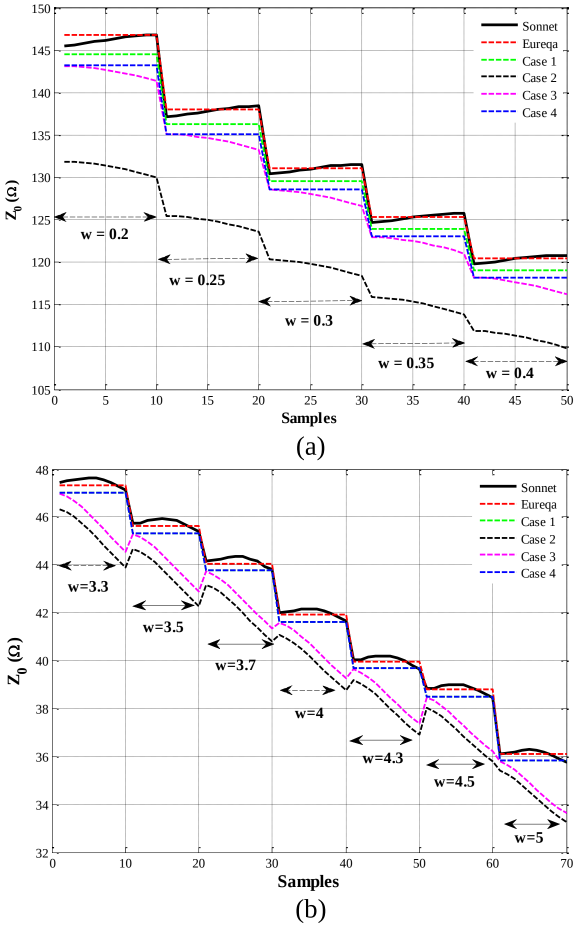


Fig. 4. Z_0 benchmarking results of FR4: (a) *w*=[0.2 0.25 0.3 0.35 0.4], and (b) *w*=[3.3 3.5 3.7 4 4.3 4.5 5], *h*=0.64 mm.

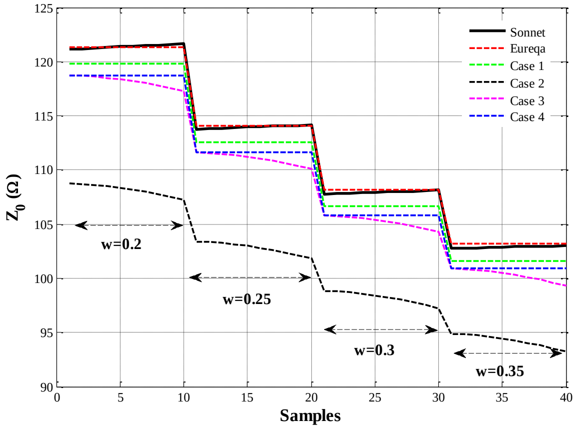


Fig. 5. Z_0 benchmarking results of Duroid 6006 (h=1.52 mm) *w*=[0.2 0.25 0.3 0.35].

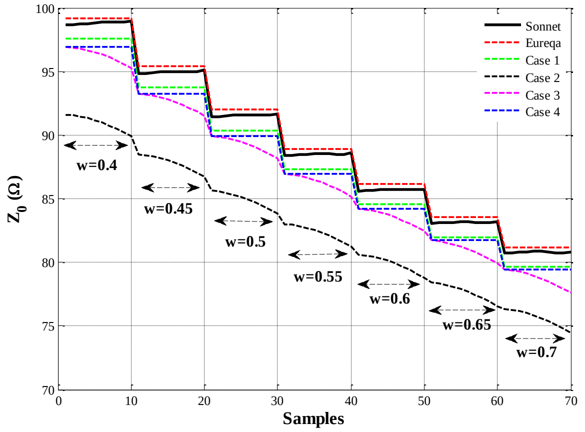


Fig. 6. Z_0 benchmarking results of Roger 3003 (h=0.64 mm) *w*=[0.4 0.45 0.5 0.55 0.6 0.65 0.7].

VI. CONCLUSION

In this study, by using 3D EM simulation and Symbolic Regression tools, an EM-based analytical expression for the characteristic impedance Z_0 of a microstrip transmission line is obtained and compared with the other analytical expressions in literature for both simulation and measurement results. The obtained expression is simpler and does not require additional calculation of ε_{eff} and does not require to use different expressions with respect to the ratio of the height to the width of the microstrip transmission line.

Furthermore, the proposed expression can easily be deployed on similar microstrip-based microwave circuit designs such as filter, amplifier, and antenna designs to acquire more accurate results, where the precise value of Z_0 is an important intrinsic design parameter.

Also, all the data obtained by Sonnet in Tables 1 and 2 are shared with detailed descriptions about the data set in [18] for all the readers who are interested in this field and would like to make contributions to this subject.

In future studies, the analytical expressions for more complex design parameters such as gain, far field, near field, scattering parameters, extraction of LC parameters for interdigital capacitance and stub inductance will be studied.

ACKNOWLEDGMENT

We would like to express our special appreciation and gratitude to Sonnet Software, Inc. [13] and Nutonian, Inc. [14] for the researcher’s licenses they provided for our use. With their kind support, we have managed to offer this study to the interest of readers.

REFERENCES

[1] C. Doulkeridis and K. Norvag, “A survey of large-scale analytical query processing in Map Reduce,” *VLDB Journal*, vol. 23, no. 3, pp. 355-380, 2014.

[2] Big Data: A Parallel Particle Swarm Optimization-Back-Propagation Neural Network Algorithm Based on Map Reduce.

[3] M. F. Korns, “Extremely accurate symbolic regression for large feature problems,” Genetic Programming Theory and Practice XII, Genetic and Evolutionary Computation, pp. 109-131, 2015.

[4] N. Allgaier and R. McDevitt, *Reverse Engineering the Brain with Eureqa*. UVM, 2011.

[5] I. Icke and C. Joshua, “Improving genetic programming based symbolic regression using deterministic machine learning,” *Bongard. Congress on Evolutionary Computation (CEC), 2013 IEEE*, pp. 1763-1770, 2013.

[6] E. Magalhaes, et al., “Experimental-based subsystem models for simulation of heterogeneous optical networks,” *J. Microw. Optoelectron. Electromagn. Appl*. [online], vol. 13, no. 2, [cited 2016-10-04], pp. 197-213, 2014.

[7] F. Renner, A. Schwab, R.-P. Kapsch, Ch. Makowski, and D. Jannek, “An approach to an accurate determination of the energy spectrum of high-energy electron beams using magnetic spectrometry,” *JINST*, vol. 9, p. P03004, 2014.

[8] G. D. Rodriguez, I. Velásquez, D. Cachi, and D. Inga, “Surrogate modeling with genetic programming applied to satellite communication and ground stations,” *Aerospace Conference, 2012 IEEE*, Big Sky, MT, pp. 1-8, 2012. DOI: 10.1109/AERO.2012.6187326

[9] S. T. Goh, S. A. Zekavat, and O. Abdelkhalik, “Space-based wireless solar power transfer via a network of LEO satellites: Doppler effect analysis,” *Aerospace Conference, 2012 IEEE*, Big Sky, MT, pp. 1-7. 2012. DOI: 10.1109/AERO.2012.6187081

[10] H. Kharbanda, M. Krishnan, and R. H. Campbell, “Synergy: A middleware for energy conservation in mobile devices,” *2012 IEEE International Conference on Cluster Computing*, Beijing, pp. 54-62, 2012. DOI: 10.1109/CLUSTER.2012.64

[11] D. M. Pozar, *Microwave Engineering*. 4th edition, Willey, 2011. ISBN 978-0-470-63155-3.

[12] C. A. Balanis, *Advanced Engineering Electromagnetics*. 2nd edition, John Wiley & Sons, 2012.

[13] www.sonnetsoftware.com

[14] www.nutonian.com

[15] M. Schmidt and H. Lipson, “Distilling free-form natural laws from experimental data,” *Science*, vol. 324, no. 5923, 2009, pp. 81-85, 2009.

[16] V. Ceperic, N. Bako, A. Baric, “A symbolic regression-based modelling strategy of AC/DC rectifiers for RFID applications,” *Expert Systems with Applications*, vol. 41, iss. 16, pp. 7061-7067, 2014.

[17] M. M. Hasan, A. Sharma, F. Johnson, G. Mariethoz, A. Seed, “Correcting bias in radar Z-R relationships due to uncertainty in point rain gauge

networks," *Journal of Hydrology*, 2014. DOI: http://dx.doi.org/10.1016/j.jhydrol.2014.09.060
[18] www.neuralnetworkdb.com

Peyman Mahouti received his Ph.D. degree in Electronics and Communication Engineering from the Yıldız Technical University in 2016. He has been currently working as a Teaching Assistant in the Yıldız Technical University Department of Electronic and Communication Engineering. The main research areas are optimization of microwave circuits, broadband matching circuits, and device modeling, computer-aided circuit design, and microwave amplifiers.

Filiz Güneş received her M.Sc. degree in Electronics and Communication Engineering from the Istanbul Technical University. She attained her Ph.D. degree in Communication Engineering from the Bradford University in 1979. She is currently a Full Professor in Yıldız Technical University. Her current research interests are in the areas of multivariable network theory, device modeling, computer-aided microwave circuit design, monolithic microwave integrated circuits, and antenna designs.

Mehmet Ali Belen received his Ph.D. degree in Electronics and Communication Engineering from Yıldız Technical University. His current research interests are in the areas of multivariable network theory, device modeling, computer aided microwave circuit design, monolithic microwave integrated circuits, and antenna arrays. He has been currently working as an Assistance Professor in the Artvin Çoruh University Department of Electric and Electronic Engineering.

Salih Demirel has received M.Sc. and Ph.D. degrees in Electronics and Communication Engineering from Yıldız Technical University, Istanbul, Turkey in 2006 and 2009, respectively. He has been currently working as an Associated Professor in the same department. His current research interests are among of microwave circuits especially optimization of microwave circuits, broadband matching circuits, device modeling, computer-aided circuit design, and microwave amplifiers.

Artificial Intelligence Technology and Engineering Applications

Xiuquan Li [1] **and Hongling Jiang** [2]

[1] Institute of Science and Technology Foresight and Evaluation
Chinese Academy of Science and Technology for Development, Beijing, 100038, China

[2] Internet of Things Technology Application Institute
China Aerospace Science and Technology Corporation, Beijing 100094, China

Abstract – There has been sixty-year development of the artificial intelligence (AI) and the maturation of AI techniques is now leading to extensive applications and industrialization. In this paper, authors review the connotation and evolution of AI techniques and engineering applications. A four-layer framework of the AI technology system is summarized in this paper to help readers understand AI family. Engineering applications of AI techniques have made remarkable progress in the recent years, for instance, applications in fault diagnosis, medical engineering, petroleum industry and aerospace industry. By introducing the state-of-the-art of AI technologies, it can help the researchers in both engineering and science fields get ideas on how to apply AI techniques to solve application-related problems in their own research areas.

Index Terms – Artificial Intelligence (AI), engineering applications, technology framework.

I. INTRODUCTION

Artificial intelligence (AI) that originated from computer science now becomes a fast growing topic in many different fields. The terminology of artificial intelligence was firstly proposed by John McCarthy et al. on Dartmouth Conference in 1956, which was originally inspired by Turing test [1]. Since AI initially refers to the creation of "humanoid" machine, it is hoped to have the ability of human-like perception and cognition and acting in complex environment. However, the definition of intelligence is still relatively vague, AI has not yet formed a unified definition. It is generally believed that AI is a discipline that studies the process of computer simulation of certain human intelligent behaviors such as perception, learning, reasoning, communicating, and acting, etc. [2, 3].

In fact, the general objective mentioned above is still far from realization because of the limitations of technology. Currently, the aim of AI mainly focuses on training machine to do things which humans can do, even in a better and more efficient manner. AI is developed as a powerful tool to make people's life better, lessen workload of humans, and improve work experience as well. Human could thus be released from many repetitive, physical and dangerous tasks.

The development of AI has experienced several ups and downs. After AI was first proposed in 1956, it went into its first golden age of 1956-1974. Many governments invested in this new research area and a lot of research projects and programs related to AI techniques were initiated worldwide during that time. In the late 1970s, AI was hit by its first “winter” (1974-1980). In 1973, British government stopped funding undirected research, subsequently, AI was under pressure. Japanese investors also withdrew funding provided to AI research. Although AI went through hard times, a few of areas were explored, such as logic programming, commonsense reasoning and so on [4].

After 1980, AI became booming again with the incentive of development of expert system. Many expert systems were adopted in the industry and governments were attracted and started to fund AI research. However, in the late 1980s and early 1990s, AI suffered the second AI winter (1987−1993). It seemed that AI machines could not satisfy what people want. Anyway, AI researchers did not give up their efforts and AI kept its development.

In 1990s, the machine learning was proposed and developed very fast. New research booms were fueled one after another in this field, for example, artificial neural network and supported vector machine. Now the most popular method is the deep neural networks, also known as deep learning, which was firstly established by Hinton in 2006 [5].

In addition to deep learning model, the development of big data and high-performance parallel computing chip together help foster the third AI boom. AI is employed to solve complex problems in various fields of engineering, business, medicine, weather forecasting, and becomes more powerful in improving performance of manufacturing and service systems.

Submitted On: March 4, 2017
Accepted On: March 23, 2017

II. A FRAMEWORK OF TYPICAL AI TECHNIQUES

Artificial intelligence is a collective of advanced computation techniques. It not only contains basic techniques of pattern classification, machine learning, knowledge understanding and expression, but also includes the application techniques of image recognition, natural language processing, human-computer interaction, expert system, anomaly detection, and so on.

In this review study, a technology framework of AI is designed as illustrated in Fig. 1, in which AI related techniques are divided into four layers, namely, supporting basic theory layer, AI model or algorithm layer, AI general technique layer and AI application technique layer. This illustration could not thoroughly depict each fields and techniques in AI research, but it will help give an overall view of the structure and main achievements of this technology.

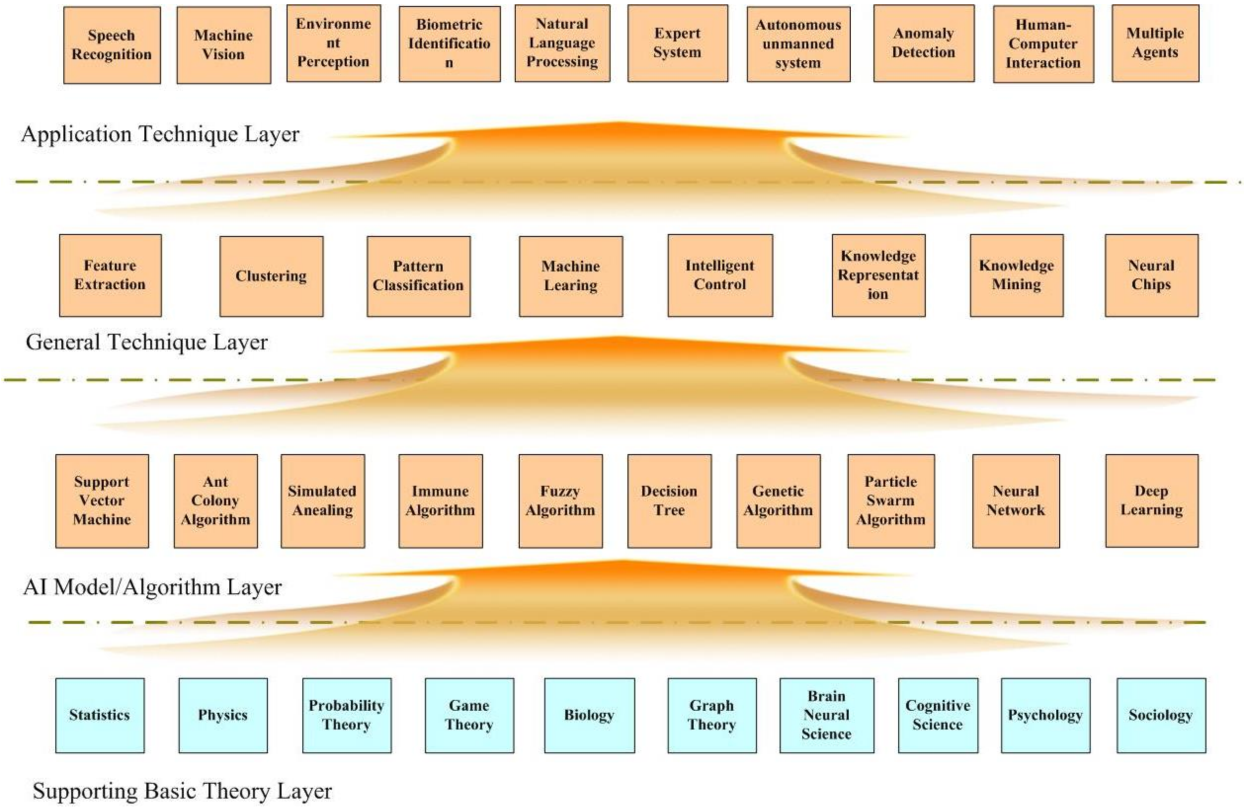


Fig. 1. A framework of typical artificial intelligence techniques.

Artificial intelligence is a typical comprehensive inter-discipline technology whose development was inspired by multiple basic disciplines. Statistics, physics, probability theory, game theory, biology, graph theory, brain neural science, cognitive science, psychology and sociology are all supporting AI development. Some selected techniques in other three layers, which are the main body of AI technology system, will be introduced in the subsections.

2.1 AI models and algorithms

Models and algorithms are fundamental techniques for artificial intelligence, most of which are designed based on scientific findings of the disciplines in supporting basic theory layer, especially statistics, brain neural science or biology. Popular models and algorithms in AI include support vector machine, ant colony algorithm, immune algorithm, Fuzzy algorithm, decision tree, genetic algorithm, particle swarm algorithm, neural network and deep learning.

Support vector machine (SVM) is a typical statistical learning model associated with supervised learning algorithms. SVM is usually used for classification and regression analysis. A SVM training algorithm builds a classification model by finding an optimal hyperplane based on a set of training examples [6, 7]. Support vector machine has been utilized for pattern classification and trend prediction in many application fields, such as power transformer fault diagnosis, disease diagnosis and treatment optimization.

Artificial neural network (ANN) is a representative model of connectionism methodology to realize artificial intelligence. ANN builds mathematical models to imitate natural biological activity from the perspective of brain neural information processing. An artificial neural network usually contains lots of neurons that are connected with each other, and models the mechanism that a biological brain solves problems by using many biological neurons connected by axons. There are various ANN models that have been proposed, such as perceptron, BP neural network, radial basis function (RBF), Hopfield neural network, self-organizing feature map (SOM) etc. [8, 9]. ANN has been employed for both supervised learning and unsupervised learning and has been applied to solve a wide variety of problems.

Deep learning is a newly developed multilayer perception feed forward artificial neural network model that has drawn a lot of attentions in artificial intelligence field [10]. Benefiting from advance in IT environment, especially large amount of available data, we have sufficient computation capacity, and more hidden layers and neurons could be employed to model high level abstractions in data, which could help get closer complicated function and lower optimization difficulty [11].

Another excellent advance of deep learning models is that they need less manual interference, for example, feature selection, weights initialization and network structure learning could be completed with minimum human contribution and less training time as compared with traditional methods. For that reason, deep learning models are not fixed for any specific task, but can be used for more general application.

Great performance of deep learning methods in object recognition competitions gained reputation. Deep learning is now widely used in various fields, in which abstract representation is useful, including speech recognition, bioinformatics, fault diagnosis, drug discovery, genomics, image classification, semantic segmentation, human pose estimation and so on [12, 13].

2.2 Typical AI general techniques

Some AI techniques are not tightly linked with specific applications, but are designed for more general purpose, such as feature extraction, clustering, pattern recognition, machine learning, intelligent control, knowledge representation, knowledge mining and so on.

2.2.1 Feature extraction

Feature extraction or feature selection is the process of selecting the most effective ones from large number of original features, constructing the feature vector for pattern recognition and modeling. One important function of feature extraction is to reduce the resources requirement for describing a large set of data. Extracted features contain sufficient information but are more non-redundant. Feature extraction could optimize the subsequent modeling processes, and bring better human interpretations in some cases. Many methods of feature extraction have been proposed, for example, linear transformation, principal components analysis (PCA), linear discriminant analysis (LDA), wavelet analysis and so on [14,15]. The selection of feature extraction method could affect the ultimate performance in AI applications such as face recognition and speech recognition.

2.2.2 Pattern recognition

Pattern recognition focuses on identifying specific patterns or regularities in data, which enables AI to make judgment like human. Bayesian classification, decision tree, linear discriminant function method, neighborhood classification method, nonlinear mapping method are frequently used in pattern recognition [16]. The subjects in pattern recognition include voice waveform, seismic wave, ECG, EEG, photos, text, symbols, biosensors and other objects, for example, computer-aided diagnosis (CAD) systems, using pattern recognition approaches, have the potential to assist radiologists in the detection and classification of breast cancer [17].

2.2.3 Machine learning

Machine learning refers to the AI technology that can enhance the performance of AI system only depending on data without following the instructions of the program. Supervised learning, unsupervised learning, and semi supervised learning are three main types of machine learning according to mode of learning [18]. Machine learning is a way of intelligent lifting of AI system and is the supporting technology of many AI applications, so machine learning has developed as the mainstream in research of artificial intelligence. Great endeavors are made currently by many top AI researchers to enhance the ability of machine learning through various methods, such as transfer learning, small sample learning, reinforcement learning, interactive learning, open learning and so on [19 20].

2.2.4 Knowledge mining

Knowledge mining is the computational process of discovering underlying knowledge from a huge amount of data and makes it understandable for further use [21]. Knowledge mining is to search for hidden information through automatic or semi-automatic algorithm, such as association rule mining, sequential pattern mining. It could be used in customer relationship management, merchandise recommendation, marketing decision and so on [22].

2.3 Typical AI application techniques

There are a series of application-oriented AI techniques, which are designed to meet specific intelligent demands. Typical AI application techniques include speech recognition, machine vision, environmental perception, biometric identification, natural language processing, expert system, trend prediction, anomaly detection, human-computer interaction, multiple agents and so on.

2.3.1 Speech recognition

Many research efforts have been put on the AI technology that can improve the experience of human-computer interaction, among which speech recognition is one of the most useful techniques. It transcribes human speech into text automatically and accurately [23]. AI general techniques of feature extraction and pattern recognition play important role in speech recognition. Other techniques are also needed, such as the description of sound and acoustic model that appears in a particular sequence and language [24].

2.3.2 Natural language processing

Teaching computer to understand from text, especially from natural language, is one of the most attractive and also most challenging tasks for AI researchers. Natural language processing (NLP) refers to an AI technique that has text processing ability like a human, for example, to extract main ideas or semantic meaning from the text that is readable, natural and grammatical. With help of NLP, an AI system could communicate with people through language to answer questions and also could learn from text to accumulate knowledge by itself. Machine learning, pattern recognition and knowledge computing are frequently used in natural language processing research [25].

2.3.3 Machine vision

Machine vision technique refers to the ability of a computer that can recognize objects, scenes, and activities from an image or video. In machine vision research, the large image analysis tasks are usually broken down into multiple controllable tasks, which can be handled by various AI general techniques. For example, some feature extraction techniques can detect edge and texture of objects from the image. Patent classification techniques can be used to determine whether the identified features represent a class of objects known to the system. Machine learning is also one of the main research methods in machine vision. It can improve the ability of object recognition by training and improving the visual model [26]. Machine vision technique has wide application fields such as robotics, intelligent factory, fault detection, security monitoring system, etc. [27, 28].

2.3.4 Expert system

Research of expert system started in the early stage of artificial intelligence development, which simulates the problem solving and reasoning ability of human experts by designing intelligent models and algorithms. It follows a knowledge driving methodology that realizes intelligence by embedding human knowledge and experience in AI system. In the past several decades, researchers have scored remarkable achievements in this field such as automatic theorem proof and medical diagnosis [29, 30]. Nowadays, this technique gets fast development with the application to knowledge engineering projects in many fields. Knowledge acquisition, human-computer interaction, knowledge representation, reasoning and decision are key techniques in designing of expert system.

2.3.5 Anomaly detection

There are widespread needs in real world, especially in engineering field, for anomaly detection technology. Compared with expert system, anomaly detection technique is a data driven AI method whose performance rely heavily on feature extraction and pattern recognition of the data. There is usually abnormal signal in monitoring data before the fault of operating machines arising, which is so weak that is hard to be noticed by people. In banking sector, insurance and investment trading, mistakes are often made by investors due to their human shortcomings, emotions, and biases. When network crime is ongoing, there is also signal that could be detected. AI has advantage in this area by utilizing pattern classification and machine learning methods.

III. APPLICATIONS OF AI TECHNOLOGY IN ENGINEERING FIELD

As an enabling technology, AI could reconstruct the mode of production, distribution, exchange and consumption in real economy, especially in engineering field. AI technology not only has advantage in reducing cost, improving efficiency and ensuring safety, but also could provide machine with man-like ability to reduce labor intensity of workers. Broad application scenes such as environmental perception, fault diagnosis, biometrics, medical diagnosis, intelligent control have attracted much research attention and many successful industrial applications have been realized.

3.1 AI in power industry

Ensuring device health is a key issue in power industry because heavy loss might be caused due to unnecessary interruption and downtime induced by an even small device fault. Many research efforts have been made to utilize AI techniques in fault diagnosis such as power transformer fault diagnosis and machinery fault diagnosis.

Tran et al. [31] proposed an approach based on decision trees and adaptive neuro-fuzzy inference to diagnose fault of induction motor. Feng et al. [32] proposed a deep neural networks (DNNs) to diagnose rotating machinery, which could extract available fault characteristics and classify fault types accurately. Samanta et al. [33] used the ANNs and SVM methods to diagnose faults of bearings. Fisher [34] applied AI to failure detection system and introduced a fault detection approach that is advanced in processing flawed data. Fischer applied the approach in many fields of power industry, such as static security assessment of electric power systems, oil leak detection in underground power cables and the stator overheating detector.

Souahlia et al. [35] presented an AI approach to conduct fault classification for power transformer dissolved gas analysis (DGA). The AI techniques used in DGA include fuzzy logic, ANN and support vector machine classifiers. Ismail et al. [36] introduced ANN and genetic algorithms into fault detection and diagnosis to establish intelligent monitoring systems in power plant. Song et al. [37] designed a fault diagnosis system based on machine vision technique. The fault diagnosis system was used for a heliostat field of a solar power plant to detect the fault heliostats in a large field.

3.2 AI in medical engineering

Medical diagnosis is a process of both knowledge intensity and experience intensity. A tiny change in medical image or medical signal is hard to be recognized by human eyes. It is usually not easy for doctors especially juniors to give an accurate diagnosis. Advances in image recognition and pattern recognition contribute to medical diagnosis based on medical image, EEG, MRI-based image data, and even speech data.

Stoitsis et al. [38] introduced a fuzzy c-means method and genetic algorithm based method to extract features from medical images. They testified AI technology to be a very useful tool in medical diagnosis for accurate quantitative analysis and qualitative evaluation of medical data. Magnetic resonance imaging (MRI) is an important diagnostic tool for early detection of cancer. Machine vision techniques have been used in deciding whether a given tumor is benign or malignant by MRI image recognition [17]. Deep learning methods can also be efficient when processing MRI-based image data to classify breast lesions [39].

AI has been applied in risk stratification of cardiovascular diseases. Zygmunt et al. [40] introduced a system based on ANN to make diagnosis of brain dysfunctions through assessment of speech quality of patient suffering from speech motor disorder. Adeli et al. [30] designed a fuzzy expert system to diagnose heart disease by employing 13 medical data as inputs, like chest pain type, blood pressure, cholesterol, maximum heart rates et al., and the output of the system is the possibility one may suffer from heart disease. Sikchi et al. [41] proposed a fuzzy expert system that be used for liver disease diagnosis using fuzzy model.

3.3 AI in petroleum industry

AI technology application in petroleum industry has also made notable progress these years. Different AI techniques have been utilized to optimize drilling operations or give an earlier detection of oil-spill in oil fields. AI systems have been testified superior than traditional methods, such as hardware based methods and biological methods.

Manshad et al. [42] proposed a two-model method to optimize drilling penetration rate based on feed forward two-layer perception neural network. The first model is proposed to choose the drilling bit and the second model is designed to predict the maximum drilling penetration rate. The models proved much efficient and accurate for optimization of drilling penetration rate. Singha et al. [43] demonstrated that Neural Networks can be used in oil spill classification systems based on image segmentation and feature classification. The approach they proposed uses two different ANNs. One is used to segment SAR images to identify pixels form oil features and the other to classify objects into oil spills based on their features.

Leak location and leak rate are two main factors that need to be taken into consideration in the process of detection of pipeline leaking. Sukarno et al. [44] developed a transmission pipeline model and leak detection model to recognize patterns of pressure distribution using ANN. After training, the artificial neural network model can predict the position of leak based on input information.

3.4 AI in aerospace industry

There are many complex tasks in aerospace industry, which are well suited for the AI technology to complete. The applications of AI in aerospace field include diagnosis for aero-engine, wear condition aid design for aircraft, optimization of key parameter of aerospace alloy and so on.

In the process of preliminary designing aircraft, various disciplines should be involved, such as aerodynamics, structure and propulsion. These different disciplines are related to each other and should be satisfied simultaneously. Oroumieh et al. [45] introduced AI models into aircraft design. Their work demonstrated that fuzzy logic and neural network can aid selecting suitable association of key parameters of aircraft, and the AI tools were effective to reduce aircraft design cycle time.

Ma et al. [46] constructed an immune algorithms based method to diagnose the wear condition of aero-engine. Negative selection principle was used, and the detectors were trained by using fault samples data. Three types of wear faults were detected, including gear overload fatigue, wear of bearing fatigue and gear agglutination or scratches. AI models can also be used to optimize key parameter of aerospace alloy. Devarasiddappa et al. [47] applied ANN model to predict the surface roughness in wire-cut electrical discharge machining of aerospace alloy.

IV. CONCLUSIONS

The new generation of information technology, represented by AI, might induce a new round of technological revolution. As a typical enabling technology, the maturation and application of AI technology lead to extensive penetration into other important industries of national economy, and will change transportation, manufacturing, medical care, business, and other industry over time.

A review on artificial intelligence technology is given in this study. We established a technology framework of four layers including theory, model and algorithm, general technology and application technology. This framework would be helpful to obtain a clear understanding of relations between various AI techniques, especially helpful for researchers of other fields. Artificial Intelligence techniques have already been applied in a variety of fields. The study summarizes some examples of AI applications in industry, including power industry, medical engineering, petroleum industry, and aerospace industry. Some of these applications have achieved surprising results.

Furthermore, today's successful of artificial intelligence attributes in some degree to the growth of many supporting technologies, such as internet of things, sensor, and big data, which also advanced at high speed in recent years. We believe that along with further maturation of AI and related technique cluster, application of artificial intelligence would be widely extended in the near future.

ACKNOWLEDGMENT

This work was supported by research project of "Research on the China Artificial Intelligence 2.0 Development Strategy" funded by reform and development special project of the Ministry of Science and Technology of China.

REFERENCES

[1] A. Turing, "Computing machinery and intelligence," *Mind*, vol. 59, pp. 433-460, 1950.

[2] A. Pannu, "Artificial intelligence and its application in different areas," *International Journal of Engineering and Innovative Technology (IJEIT)*, vol. 4, no. 10, pp. 79-84, 2015.

[3] J. Gabriel, *Artificial Intelligence: Artificial Intelligence for Humans*. Createspace Independent Publishing Platform, USA, 2016.

[4] M. Flasiński, *History of Artificial Intelligence. Introduction to Artificial Intelligence*. Springer International Publishing, pp. 1-20, 2016.

[5] G. Hinton and R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science*, vol. 313, pp. 504-507, 2006.

[6] B. Cortes and V. Vapnik, "Support vector networks," *Int. Machine Learning*, vol. 20, pp. 273-297, 1995.

[7] V. Vapnik, *The Nature of Statistical Learning Theory*. Springer-Verlag, New York, 1995.

[8] J. Hopfield, "Neural networks and physical systems with emergent collective computational abilities," *Proc. Natl. Acad. Sci.*, vol. 79, pp. 2554-2558, 1982.

[9] D. Rumelhart, G. Hinton, and R. Williams, "Learning internal representations by error propagation," in *Parallel Distributed Processing*, vol. 1, ed. D. E. Rumelhart and J. L. McClelland, Cambridge, MA: MIT Press, pp. 318-362, 1986.

[10] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436-444, 2015.

[11] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural Networks*, vol. 61, pp. 85-117, 2014.

[12] A. Saiz, "*Deep learning review and its applications*," Department of Computer Science, Artificial Intelligence Computer Science Faculty, Universidad del Pais Vasco Universidad del Pais Vasco, 2015.

[13] O. Russakovsky, J. Deng, H. Su, et al., "ImageNet large scale visual recognition challenge, arXiv preprint arXiv," *International Journal of Computer Vision*, vol. 115, pp. 211-252, 2015.

[14] S. Bharti and S. Preet, "An enhanced feature extraction method and classification method of EEG signals using artificial intelligence," *International Journal of Computer Applications*, vol. 126, pp. 19-24, 2015.

[15] B. Agarwal and N. Mittal, "Prominent feature extraction for review analysis: an empirical study," *Journal of Experimental & Theoretical Artificial Intelligence*, pp. 1-14, 2014. Doi:10.1080/0952813X.2014.97783.

[16] A. Rosenfeld and H. Wechsler, "Pattern recognition: Historical perspective and future directions," *International Journal of Imaging Systems & Technology*, vol. 11, pp. 101-116, 2015.

[17] F. Roberta, S. Mario, F. Salvatore, et al., "Pattern recognition approaches for breast cancer DCE-MRI classification: A systematic review," *Journal of Medical & Biological Engineering*, vol. 36, pp. 449-459, 2016.

[18] E. Alpaydin, *Machine Learning: The New AI*. The MIT Press Essential Knowledge Series, MIT Press, 2016.

[19] V. Mnih, K. Kavukcuoglu, D. Silver, et al., "Human-level control through deep reinforcement learning," *Nature*, vol. 518, no. 7540, pp. 529, 2015.

[20] Z. Wang, T. Schaul, M. Hessel, et al., "Dueling network architectures for deep reinforcement learning," *Proceedings of the 33th International Conference on Machine Learning*, New York, NY,

USA, 2016.

[21] M. Shaw, C. Subramaniam, G. Tan, et al., "Knowledge management and data mining for marketing," *Decision Support Systems*, vol. 31, pp. 127-137, 2001.

[22] T. Nasukawa and T. Nagano, "Text analysis and knowledge mining system," *Ibm Systems Journal*, vol. 40, pp. 967-984, 2009.

[23] M. Johnson, S. Lapkin, V. Long, et al., "A systematic review of speech recognition technology in health care," *BMC Medical Informatics and Decision Making*, vol. 14, pp. 94, 2014.

[24] L. Besacier, E. Barnard, A. Karpov, et al., "Automatic speech recognition for under-resourced languages: A survey," *Speech Communication*, vol. 56, pp. 85-100, 2014.

[25] E. Cambria and B. White, "Jumping NLP curves: A review of natural language processing research," *IEEE Computational Intelligence Magazine*, vol. 9, pp. 48-57, 2014.

[26] Y. Guo, Y. Liu, A. Oerlemans, et al., "Deep learning for visual understanding," *Neurocomputing*, vol. 187, pp. 27-48, 2015.

[27] D. Luzuriaga and M. Balaban, "Application of computer vision and electronic nose technologies for quality assessment of color and odor of shrimp and salmon," *Tijdschrift Voor Seksuologie*, vol. 38, pp. 68-76, 2014.

[28] F. Ali, H. Mohsen, A. S. Mohammad, et al., "Every picture tells a story: Generating sentences from images," *European Conference on Computer Vision (ECCV 2010)*, pp. 15-29, 2010.

[29] W. J. Wu, "Some remarks on mechanical theorem-proving in elementary geometry," *Acta Math. Scientia*, vol. 3, pp. 357-360, 1983.

[30] A. Adeli, and M Neshat, *A fuzzy expert system for heart disease diagnosis*, Lecture Notes in Engineering & Computer Science, vol. 1, pp. 1-6, 2010.

[31] V. Tran, B. Yang, M. Oh, and A. Tan, "Fault diagnosis of induction motor based on decision trees and adaptive neuro-fuzzy inference," *Expert Syst. Appl*., vol. 36, pp. 1840-1849, 2009.

[32] J. Feng, Y. Lei, J. Lin, et al., "Deep neural networks: A promising tool for fault characteristic mining and intelligent diagnosis of rotating machinery with massive data," *Mechanical Systems & Signal Processing*, vol. 72-73, pp. 303-315, 2016.

[33] B. Samanta and C. Nataraj, "Use of particle swarm optimization for machinery fault detection," *Eng. Appl. Artif. Intell*., vol. 22, pp. 308-316, 2009.

[34] D. Fischer, *Artificial intelligence techniques applied to fault detection systems*, Ph.D. Dissertation, Electrical and Computer Engineering, McMaster University, Ontario, Canada, 2004.

[35] S. Souahlia, K. Bacha, and A. Chaari, "Artificial intelligence tools aided-decision for power transformer fault diagnosis," *International Journal of Computer Applications*, vol. 38, pp. 1-8, 2012.

[36] R. I. B. Ismail, F. B. Ismail Alnaimi, and H. F. Alqrimli, "Artificial intelligence application in power generation industry: initial considerations," *IOP Conference Series Earth and Environmental Science*, vol. 32, pp. 1-4, 2016.

[37] Y. Song, W. Huang, and X. Zhu, "A vision-based fault diagnosis system for heliostats in a central receiver solar power plant," *10th World Congress on Intelligent Control and Automation (WCICA 2012)*, Beijing, pp. 3417-3421, 2012.

[38] J. Stoitsis, et al., "Computer aided diagnosis based on medical image processing and artificial intelligence methods," *Nuclear Instruments & Methods in Physics Research*, vol. 569, pp. 591-595, 2006.

[39] A. Işın, C. Direkoğlu, and M. Şah, "Review of MRI-based brain tumor image segmentation using deep learning methods," *Procedia Computer Science*, vol. 102, pp. 317-324, 2016.

[40] C. Zygmunt and A. Napieralski, "Artificial intelligence in medical diagnosis of some brain dysfunctions," *International Journal of Microelectronics and Computer Science*, vol. 6, pp. 1-5, 2015.

[41] S. S. Sikchi, S. Sikchi, and M. Ali, "Artificial intelligence in medical diagnosis," *International Journal of Applied Engineering Research*, vol. 7, pp. 1539-1543, 2012.

[42] A. Manshad, H. Rostami, H. Toreifi, et al., *Optimization of Drilling Penetration Rate in Oil Fields Using Artificial Intelligence Technique*. Nova Science Publishers, Inc., NY, USA, 2016.

[43] S. Singha, J. Bellerby, and O. Trieschmann, "Detection and classification of oil spill and look-alike spots from SAR imagery using an artificial neural network," *IEEE Geoscience and Remote Sensing Symposium*, pp. 5630-5633, 2012.

[44] P. Sukarno, et al., "Leak detection modeling and simulation for oil pipeline with artificial intelligence method," *Itb Journal of Engineering Science*, vol. 39, pp. 1-19, 2007.

[45] M. Oroumieh, S. Malaek, M. Ashrafizaadeh, et al., "Aircraft design cycle time reduction using artificial intelligence," *Aerospace Science & Technology*, vol. 26, pp. 244-258, 2013.

[46] A. Ma, Y. Li, Y. Cao, et al., "Intelligent diagnosis for aero-engine wear condition based on immune theory," *IEEE Prognostics and System Health Management Conference*, pp. 678-682, 2014.

[47] D. Devarasiddappa, J. George, M. Chandrasekaran, et al., "Application of artificial intelligence approach in modeling surface quality of aerospace Alloys in WEDM process," *Procedia Technology*, vol. 25,

pp. 1199-1208, 2016.

Xiuquan Li was born in 1976. He received his Ph.D. degree in the field of Intelligent Data Processing from Tsinghua University in 2010. Before that, he received his master degree in the field of Artificial Intelligence from East China University of Science and Technology.

He is now an Associate Research Fellow in Chinese Academy of Science and Technology for Development, Beijing, China. His current Research area is big data and artificial intelligence technology foresight, technology roadmap and academic hotspot discovery. He is Member of Emergent Technologies Technical Committee of IEEE Computational Intelligence Society.

Hongling Jiang graduated from Computer Science College of Nankai University and received her doctor degree in 2013. She currently works in Internet of Things Technology Application Institute, China Aerospace Science and Technology Corporation since graduation. Her research interests include artificial intelligence, big data, data mining, internet of things and so on.

Enhanced Energy Localization with Wideband Hyperthermia Treatment System

N. Nizam-Uddin and Ibrahim Elshafiey

Department of Electrical Engineering
King Saud University Riyadh, Kingdom of Saudi Arabia
enizamuddin@ksu.edu.sa, ishafiey@ksu.edu.sa

Abstract **─** This paper presents a tool to enhance hyperthermia treatment based on multichannel wideband system. The potential of this system in enhancing energy localization is investigated. A model is developed for the hyperthermia treatment plan taking into account cylindrical phantom of human head. Dispersive modeling of tissue properties is used to allow wideband simulation. Problem formulation for optimization of wideband system is presented. Simplified time-delay tool based on coherent phased-array approach is developed and presented. The results reveal the potential of wideband system compared to conventional narrow band systems in enhancement of energy focus. Investigation of advanced optimization processes is introduced that allow real time shaping of the signal waveform in wideband systems. The research addresses the importance of incorporating recent techniques in processing big data to facilitate adopting wideband hyperthermia in clinical systems.

Index Terms **─** Big data, energy localization, hyperthermia treatment, wideband systems.

I. INTRODUCTION

Big data is predicted to contribute considerably in healthcare applications by enhancing the performance at lower cost. Big data is characterized in different dimensions including volume, variety, velocity, and veracity. In biomedical and healthcare domains volume indicates the exponentially growing size of data [1, 2]. As an example the healthcare data is predicted to increase form 500 petabytes in 2012 to 25,000 petabytes by the year 2020 [3]. Variety of biomedical data signifies the heterogeneous nature of data types characterizing patients. The big data revolution is thus shown to accelerate value and innovation in health care [4] .

The tools associated with big data allows the move from population based treatment plans to patient specific treatment plans. Plans can be established to enhance the efficacy and specificity of the treatment by optimizing the system capabilities to the patient. These directions motivated this research to enhance hyperthermia treatment (HT). A multi-channel wideband system is proposed to overcome limitations of conventional systems in energy localization to tumor regions.

Hyperthermia is found to be effective in eradicating cancerous tissues in various regions including breast and head and neck. During hyperthermia, the temperature of malignant tissue is raised to 40-45C° [5, 6]. Hyperthermia reduces the vitality of malignant tissues and increases their sensitivity to radiotherapy and anticancer drugs. Hyperthermia can thus reduce the required doses and the associated side effects of treatment plan.

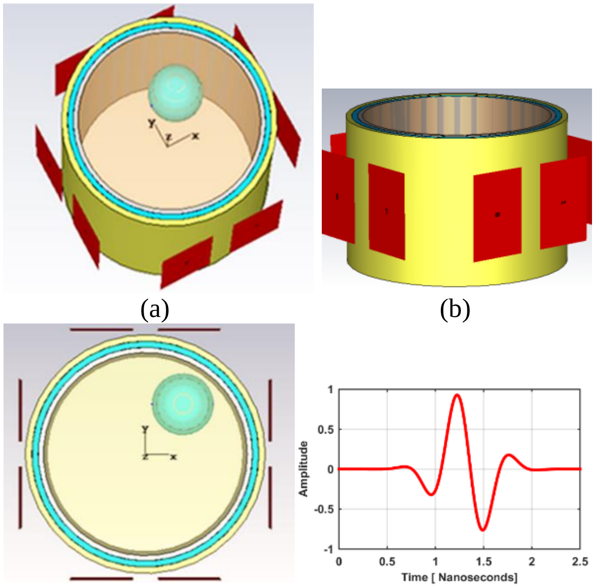
External HT provides an attractive non-invasive therapeutic plan [7, 8]. The potential of effective treatment however depends on the capability of localizing energy onto tumor regions without affecting the healthy tissues. This requires a high level of precision and accuracy in the treatment planning [9]. With differences in age, size, and tumor locations of patients, success of the treatment plan depends on the development of patient-specific treatment plan derived from patient-model-based data. The technical capabilities of the treatment system should also be enhanced to provide more flexibility and adaptivity in enhancing the energy localization schemes. This research thus aims at enhancing the degrees of freedom available to the system in terms of increasing the number of channels and the operation bandwidth. Computational tools are developed to allow handling of big data associated with patient interaction with energy sources.

II. INVESTIGATION OF WIDEBAND HYPERTHEMRIA SYSTEM

Feasibility of wideband hyperthermia is investigated by building a model of the interaction between tissues and energy sources. A simplified head phantom model is built using CST microwave studio [10], implementing a four-layered cylindrical phantom of radius 10 cm is chosen to represent a human head. The inner 8-cm radius cylinder is used for depicting brain tissue and the outer three cylinders for gray matter, cerebrospinal fluid (CSF)

Submitted On: December 31, 2016
Accepted On: April 21, 2017

and skull of radii of 8.4, 8.9 and 9.4 cm respectively. This model is shown in Fig. 1 (a). The applicator ports are arranged in counter clockwise direction around the phantom shown in Fig. 1 (b). The tumor is chosen to be of spherical shape of radius 2.5 cm, and is located at x = 3 cm, y = 4 cm and z = 0, as shown in Fig. 1 (c), where origin is set at the center of the phantom. For an acceptable compromise of resolution of reconstructed image, penetration of EM signals and the antenna size, a Gaussian pulse of bandwidth (0.3 to 2.5 GHz) is applied to each excitation port as shown in Fig. 1 (d).



(a) (b)

(c) (d)

Fig. 1. (a) Perspective view of proposed wideband hyperthermia treatment, (b) side view of proposed model to indicate arrangement of ports, (c) front view of proposed model to indicate the location of tumor, and (d) wideband excitation signal.

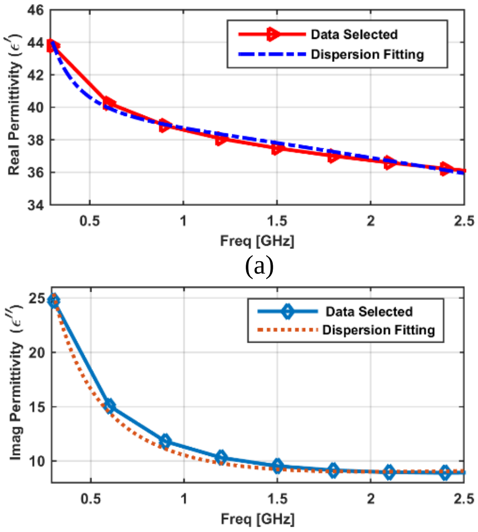
The big data characteristics of our proposed model are discussed below.

A. Volume of system

The model consists of multi-channels under wideband excitation to heat a certain sensitive tissue region. Eight channels are shown in this simulation for treatment of brain tumors. This model presents complexity in various dimensions. The first dimension is the bandwidth of operation. Multichannel wideband treatment allows treatment of both superficial as well as deep-seated tumors with appropriate spatial resolution. Optimization of excitation waveform at each port is essential to allow energy localization. The waveform can be modeled as excitation of a number of subcarriers. Each subcarrier is optimized for magnitude and phase. In addition, the problem depends on tumor data. The number of tumors and the size and location of each tumor contribute to the volume of data associated with this problem.

B. Veracity of system

The second dimension is the accurate modeling of human head tissue. The heterogeneous nature of human head representing different tissues levels such as brain, gray matter, cerebrospinal fluid (CSF) and skull along with tissue characterization of tumor in terms of its location, shape and size need to be modeled accurately under a wideband excitation signal. This makes the veracity of our problem. For this purpose, the dispersive dielectric properties of brain tissue and tumor are chosen in accordance with [11-13] and are shown in Figs. 2 (a) and (b).



(a)

(b)

Fig. 2. Behavior of human brain tissue in terms of permittivity under wideband excitation for: (a) real permittivity and (b) imaginary permittivity values.

C. Variety of system

The third and important dimension of the undertaken case study is the variety of the problem as we are analyzing the problem in two different domains under different simulations environment thus making the validation of the problem indispensable. For this purpose, we devised a time-delay tool to focus energy at cancerous region. This tool is incorporated in both configurations to investigate the equivalency of transformation process while shifting from model-to-data transformed configuration.

D. Velocity of system

With the increase of number of channels of the system and the number of subcarriers in each channel, the number of degrees of freedom becomes high and the optimization process becomes computationally complex. Devising an appropriate optimization tool can lead to minimum utilization of system resources and expediting treatment process. Theoretically, an appropriate problem formulation can lead to effective optimization. For this reason, we investigate the optimization problem in different scenarios as follows.

III. OPTIMIZATION PROBLEM

Advanced optimization with the aim to achieve an effective optimization, several techniques have been reported in recent research particularly in the domain of wideband. Three-dimensional time reversal technique in hyperthermia treatment to accumulate energy in head and neck tumors [14]. The waveform transmitting diversity approach for wideband MIMO radar [15, 16] can be implemented to enhance energy localization in hyperthermia treatment. Beamforming methods can also help the implementation of energy localization. It includes non-uniform and short inverse FFT based wideband beamforming method presented in [17]. Compress sensing beamforming proposed in [18] and sparse array based wideband beamforming discussed in [19]. For speed and flexibility a refined eigen value optimization is adopted in [20].

The optimization of model can be formulated in two scenarios: narrowband and wideband.

A. Narrowband case

Assume that an array of M applicators is located in a coupling or cooling medium surrounding the head region at locations $\mathbf{r}_m (m = 1,2, \ldots, M)$ as shown in Fig. 1 (a). Let $\mathbf{x}_m(n)$ $(n = 1,2, \ldots, N)$ represent the discrete-time baseband signal transmitted by the m^{th} applicator. The baseband signal at a location r inside the head can be described as:

$$y(r,n) = \sum_{m=1}^{M} \Psi(f_0,\ \mathbf{r}_m, \mathbf{r})\, \mathbf{x}_m(n), \tag{1}$$

where f_0 is the baseband frequency and $y(r,n) = \Psi(f_0, \boldsymbol{r}_m, \boldsymbol{r})$ is the tissue interaction function that accounts of propagation attenuation of the electromagnetic (EM) signal inside tissue along with the corresponding phase delay due to the travelling of energy.

Let **a**(**r**) be the transpose of steering vector such that:

$$\mathbf{a}(\mathbf{r}) = [\Psi^*(f_0, \boldsymbol{r}_1, \boldsymbol{r}), \Psi^*(f_0, \boldsymbol{r}_1, \boldsymbol{r}), \ldots \Psi^*(f_0, \boldsymbol{r}_1, \boldsymbol{r})]^{\boldsymbol{T}}, \tag{2}$$

while,

$$\mathrm{X}(n) = [\mathbf{x}_1(n)\ \ \mathbf{x}_{\mathbf{2}}(n)\ \ \ldots\ \mathbf{x}_M(n)]^T, \tag{3}$$

represents the transmitted signal. Equation (1) can be written as:

$$Y(\mathbf{r}, n) = \mathbf{a}^*(\mathbf{r})\mathrm{X}(n). \tag{4}$$

The beam pattern representing the signal power at locations **r** and is given by:

$$P(r) = E\{Y(r,n)Y^*(r,n)\} = \mathbf{a}^*(\mathbf{r})\ R\ \mathbf{a}(\mathbf{r}), \tag{5}$$

where $(\cdot)^*$ represents conjugate transpose and E the expectation respectively. R is the covariance matrix of $x(n)$ denoted by:

$$R = E\{\mathrm{X}(n)\mathrm{X}^*(r,n)\}. \tag{6}$$

From Equation (6) it is noted that transmitting beam pattern is a function of target location only. Under the constraint that power form the applicator elements is uniform, we reach at an optimization problem as:
Maximize the function,

$$\xi = \frac{P_i}{P_o}. \tag{7}$$

Subject to,

$$a^*(r)\ a(r) = 1. \tag{8}$$

P_i and P_o represent the spatial integration of absorbed power inside and outside the intended tumor location, respectively.

B. Wideband case

Now consider the case when f_0 is the baseband frequency in the interval $\left[-\frac{B}{2}, \frac{B}{2}\right]$. In such a case we can take the Fourier transform of Equation (1) to get:

$$y(r, \mathrm{n}, f) = \sum_{m=1}^{M} \Psi(f_0,\ \mathbf{r}_m, \mathbf{r})\, \mathbf{x}_m(n, f). \tag{9}$$

The beam pattern at a location r and frequency $(f_0 + f)$ can be written as:

$$P(r, f_0 + f) = |Y(r,n,f)|^2 = |a^*(r,f)\ Y(r,n,f)|^2, \tag{10}$$

where $f \in \left[-\frac{B}{2}, \frac{B}{2}\right]$.

In this case optimization problem is conducted in terms of maximization of ξ such that,

$$\xi = \int_{-B/2}^{B/2} \frac{P_i(f)}{P_o(f)} df. \tag{11}$$

P_i (f) and P_o (f) represent the spatial integration of absorbed power inside and outside the intended tumor location, respectively for a given band of frequencies.

Wideband excitation thus offers higher degree of computational complexity than its counterpart narrowband. Initial investigation is introduced of this optimization problem based on simplified optimization tool, which utilizes the phase shifting characteristics of transmission line model. This tool utilizes the information from the model to focus energy at sensitive region making it able to operate as an off-line standalone module even though the actual model and data are inaccessible. This signifies one of the important big data characteristics known as model-to-data and data-to-information transformation of our undertaken case study.

IV. DATA ANALYTICS

In order to investigate different aspects of our proposed case study a model-to-data based approach is adopted as shown in Fig. 3.

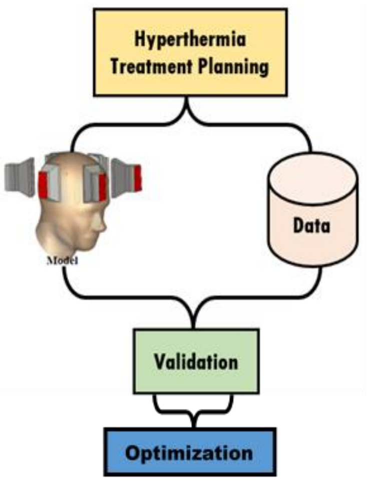


Fig. 3. A model-to-data validation and optimization for wideband HT system.

The data analytics for the proposed case study is divided into three steps as shown in Fig. 4. In the first step patient modeling is done in time domain by choosing appropriate tissue properties of human head and excitation ports using CST microwave studio [10].

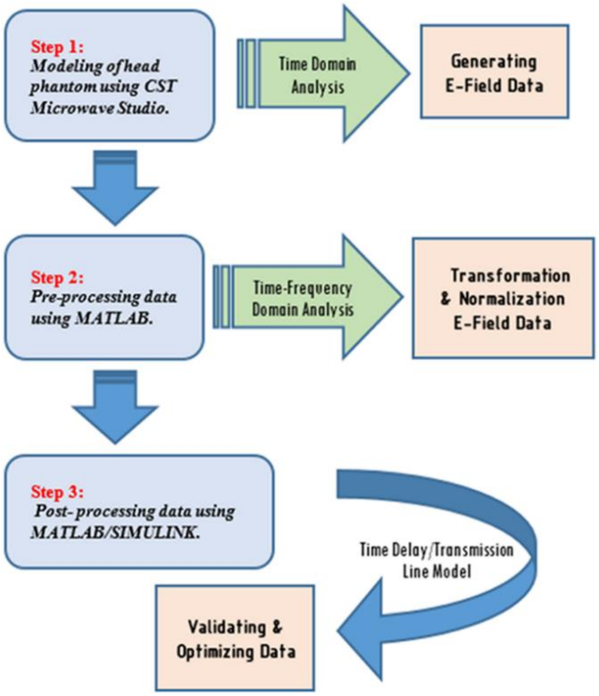


Fig. 4. Flow chart of big data analytics for wideband HT system.

In the second step, this model is transformed to data for pre-processing. The data is analyzed in time and frequency domains and results are visualized in the form of E-field mapping. In the third step, validation of model-to-data transition is achieved using a time delay tool based on coherent phased-array concept. Additionally, an optimization technique using transmission line model is proposed using MATLAB/ SIMULINK. The subsequent sections explain the procedures and techniques adopted in each step.

A. Generating E-field data

As a first step, electromagnetic (EM) simulation of the model shown in Fig. 1 (a) is performed using CST microwave studio. The solver discretizes the model. Time domain field monitors are placed to acquire the E-field data inside the head phantom in three planes as shown in Figs. 5 (a), (b), and (c), covering the tumor in x, y and z directions respectively.

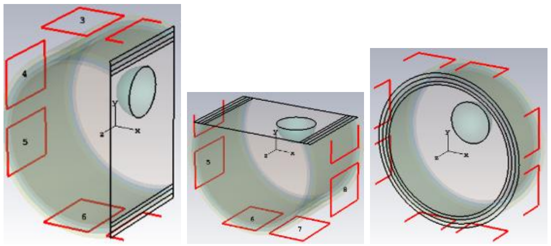

(a) (b) (c)

Fig. 5. Representation of proposed model in three normal planes traversing tumor region: (a) x-plane, (b) y-plane, and (c) z-plane.

To reduce the simulation run, the processor hardware is extended with graphical processing unit (GPU) C2070 from NVIDIA, Tesla. With this GPU configuration, the solver took about an hour to complete a run for a total number of 4.63344e6 mesh cells. The time domain E-field data corresponding to these three normal planes is arranged for processing in the next step.

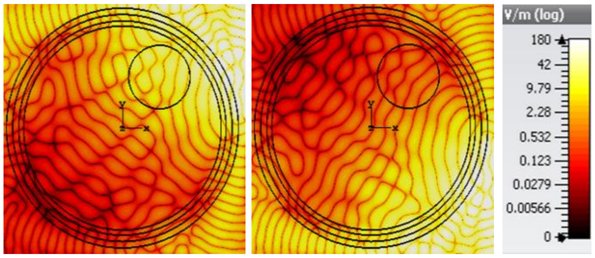
B. Transformation and normalization of E-field data

The time domain E-field data from step 1 is imported to a MATLAB [21] environment for pre-processing. The objective of this model-to-data transformation is to investigate the time-frequency behavior of the model. Additionally, it would allow us more degrees of freedom in terms of processing, organizing, and analyzing the data according to treatment requirement.

The size of each E-field data for a single port in one normal plane is 485 Megabytes (MB). Since we have 8-ports and 3-normal planes, the size of time domain E-field data becomes 485*8*3 (MB). This data is transformed to frequency domain using 256-points FFT. Since the excitation signal is a Gaussian pulse, the frequency domain data is normalized to excitation signal.

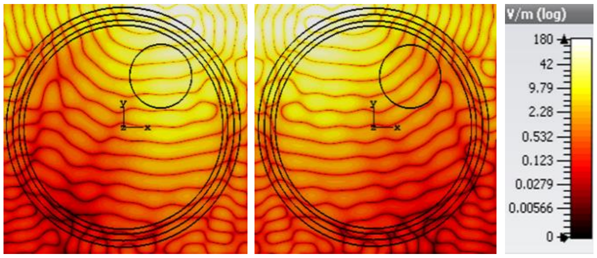
With this data available both in time and frequency domain, it was necessitated to structure the data to be able to use for a variety of requirements. It can be used for narrowband and wideband hyperthermia treatment, for a single and multiple port heating, for superficial and deep-seated HT. The structuring of data in this step allows us to choose any frequency from the available excitation signal (0.3-2.5 GHz) capable to heat the tumor from any port location. As an example, the E-field maps of ports excited with different subcarriers are exhibited in Fig. 6 to Fig. 9.

Figure 6 depicts propagation of E-field when ports 1 and 8 are excited with 2 GHz. Figure 7 shows E-field, when ports 2 and 3 are excited with 1.5 GHz. Results corresponding to exciting ports 4 and 5with frequency of 1 GHz is depicted in Fig. 8. Since ports 6 and 7 are distant apart from tumor location, we choose low excitation frequency such as 0.5 GHz for improved penetration depth. E-field inside the phantom is illustrated in Fig. 9.



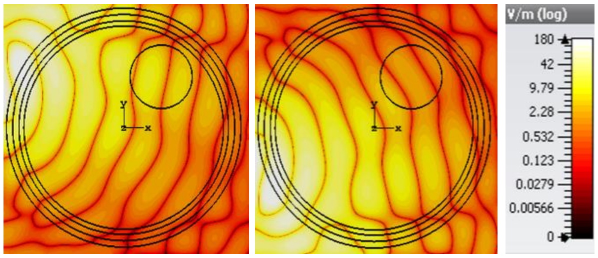
(a) (b)

Fig. 6. E-field map of: (a) port 1 and (b) port 8 excited with 2 GHz.



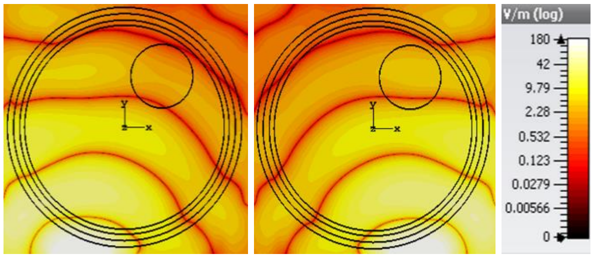
(a) (b)

Fig. 7. E-field map of: (a) port 2 and (b) port 3 excited with 1.5 GHz.



(a) (b)

Fig. 8. E-field map of: (a) port 4 and (b) port 5 excited with 1 GHz.

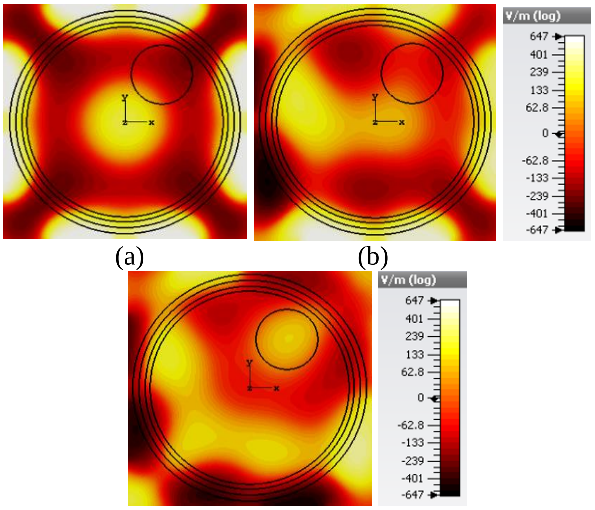


(a) (b)

Fig. 9. E-field map of: (a) port 6 and (b) port 7 excited with 0.5 GHz.

C. Validating E-field data

With the objective to guide energy from ports to tumor location, a time delay tool is developed based on coherent phased-array concept [22]. This time delay tool takes all factors into consideration, which makes accumulation of EM energy at tumor region possible with minimum damage to the surrounding healthy tissue. These factors include the phase, amplitude, number of subcarriers, and the propagated distances of subcarriers from the ports to the target region. This time delay tool when implemented in the model-based configuration with different phase shifts gives energy localization at different phantom locations. For a zero phase shift, energy is concentrated at phantom center presenting an ideal case for a tumor to be located at phantom center. In case of an undesirable phase shift, energy is distributed at some off-target region. For a desirable phase shift, maximum energy is focused at tumor location. This is shown by Ez-field map of Figs. 10 (a), (b) and (c) respectively.

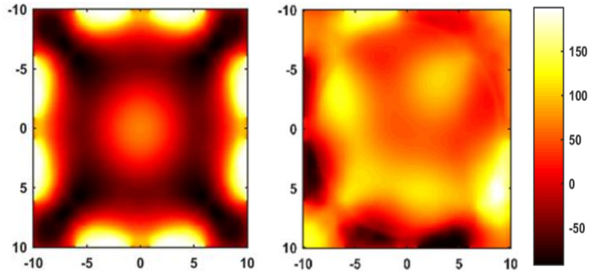


(a) (b)

(c)

Fig. 10. The capability of time delay tool to localize energy. E-field map with: (a) zero phase shift, (b) inappropriate phase shift, and (c) appropriate phase shift when implemented in model-based format.

The time delay tool is incorporated in model-to-data transformed format. The results are show in Figs. 11 (a) and (b), illustrating the potential of this tool in energy localization.



(a) (b)

Fig. 11. The capability of time delay tool to localize energy. E-field map with: (a) zero phase shift and (b) appropriate phase shift when implemented in model-to-data transformed format.

D. Optimization of E-field data

The proposed simplified optimization tool in this research is based on transmission line theory. Transmission line with different shapes and types find a variety of applications in imaging, communication, and high frequency circuit designs. Recently, it is used in real time through wall imaging and for addressing cochlear nonlinearity in acoustic signal processing by utilizing its length and characteristic impedance [23, 24]. For high frequency applications, the length of transmission line (TL) becomes crucial. The phase delays and reflections in transmission line becomes essential to analyze if its length becomes greater than a significant portion of wavelength of the transmitted frequency. The length and characteristics impedance of a transmission line can be related as [25]:

$$V_1(t) - I_1(t) \times Z_0 = V_2(t-\tau) + I_2(t-\tau) \times Z_0, \quad (12)$$

and

$$V_2(t) - I_2(t) \times Z_0 = V_1(t-\tau) + I_1(t-\tau) \times Z_0, \quad (13)$$

where, V_1 is the voltage and I_1 is the current at the source end of the TL, while V_2 and I_2 is the time delayed voltage and current respectively at the load end and τ is the time delay offered by TL. The characteristics impedance of the TL is given by Z_0. The proposed TL model is developed in SIMULINK[21] environment .

The source of TL is connected to carrier and load end is terminated with 50 ohms resistance. A set of four subcarriers with frequency values of 0.5, 1, 1.5, and 2 GHz are considered at each of the eight ports. A tool is developed to adjust the delay of the TL corresponding to each subcarrier to achieve constructive interference at target region and destructive interference away from the target thereby enhancing energy localization. The average power deposition is calculated at each pixel of the phantom, where the phantom is discretized to 80x80 pixels.

With the tumor located at the top-right the results of optimization tool for narrow and wideband hyperthermia are shown in Figs. 12 and 13 respectively. These results demonstrate that the proposed optimization tool accomplishes two objectives. It can be invoked to enhance energy localization at target region and can be used to compare the performance of wideband HT system to the conventional narrowband HT system in terms of energy localization.

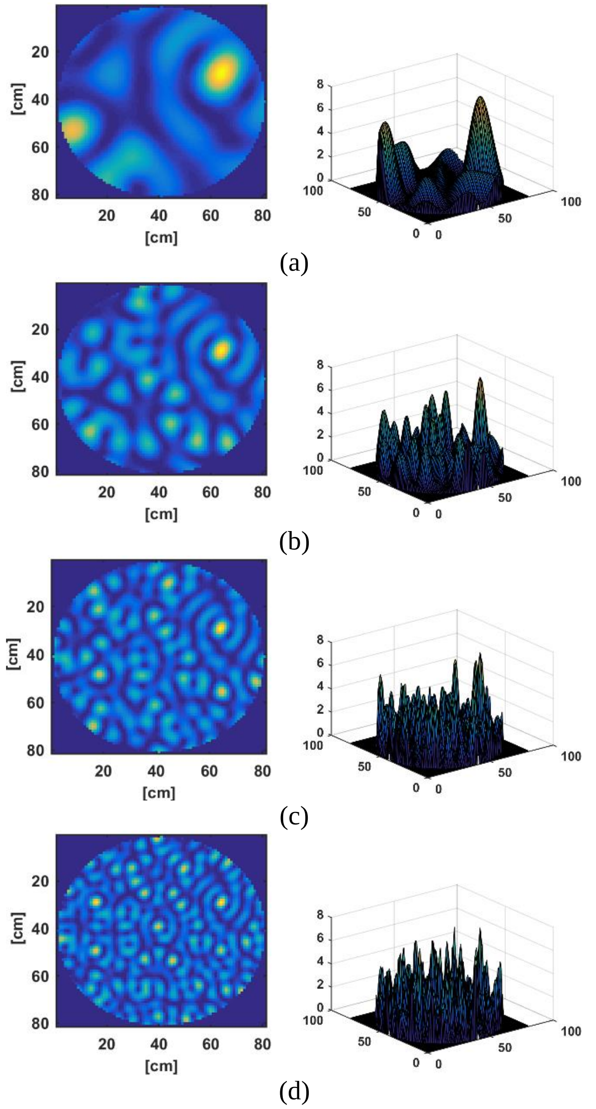
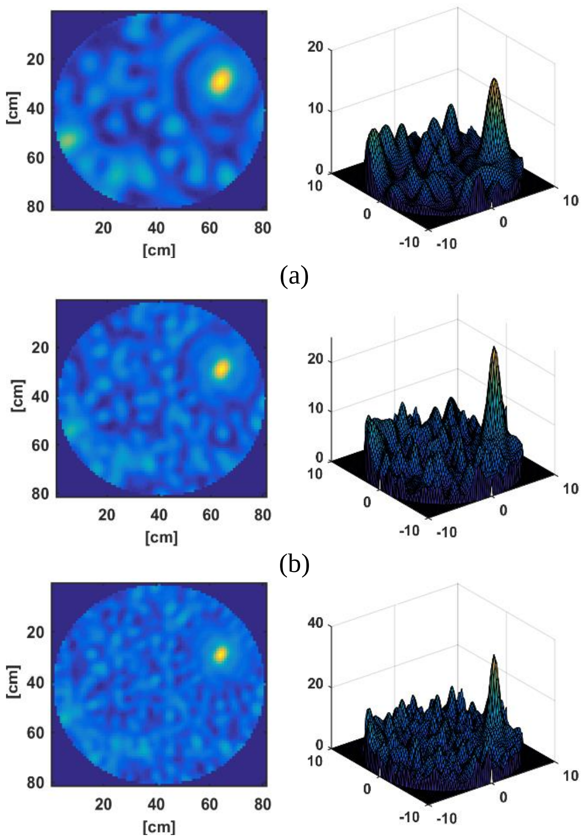


Fig. 12. Energy localization (left column) and corresponding heating pattern (right column) in crucial region for narrowband case: (a) 0.5 GHz, (b) 1 GHz, (c) 1.5 GHz, and (d) 2 GHz.



(a)

(b)

(c)

Fig. 13. Energy localization (left column) and corresponding heating pattern (right column) in crucial region for wideband case when: (a) two frequencies (0.5 and 1 GHz) are combined, (b) three frequencies (0.5, 1 and 1.5 GHz) are combined, and (c) four frequencies (0.5, 1, 1.5 and 2 GHz).

V. DISUSSION AND CONCLUSIONS

The potential of wideband system in enhancing hyperthermia treatment is investigated. Mathematical formulation of optimization process is developed to allow energy focus to tumor location, while persevering healthy tissue. Time delay tool is developed and deployed to validate the potential of wideband techniques in energy localization.

The proposed system provides increase in degrees of freedom that allows performance enhancement. Wideband energy can be shaped to target tumors at different depth. Multichannel configuration allows enhancement in energy focus. Model based optimization allows adopting patient-specific model in localizing the energy.

These features however are associated with the need to deal with big data. Clinical adopting of such system will require the development of real time optimization systems that can accommodate various parameters of the system. The tissue properties will also vary within the treatment session in accordance with temperature maps. This requires the update of patient model progressively during the treatment session.

These challenges can be addressed using two approaches. First, the computational capability can be enhanced by adopting cluster techniques. Hardware acceleration has also been used in this research and is found to have real impact on reducing computational analysis time. On the other hand, more robust optimization techniques can be adopted to allow real time use of the optimization tools

This research addressed the use of transformation from time to frequency domain, which allows the efficient handling of big data associated with this problem. Further investigation will consider various signal decomposition techniques. Time reversal techniques should simplify the optimization process by providing an initial solution that can approach close to the global optimum parameter values.

ACKNOWLEDGMENT

This research is supported by a research grant from King Abdul-Aziz City for Science and Technology (KACST), General Administration for Research Grants, Research Project AT 35-210.

REFERENCES

[1] S. Fodeh and Q. Zeng, “Mining big data in biomedicine and health care,” J*ournal of Biomedical Informatics*, vol. 63, p. 400, 2016.

[2] K. J. Archer, K. Dobbin, S. Biswas, R. S. Day, D. C. Wheeler, and H. Wu, “Computer simulation, bioinformatics, and statistical analysis of cancer data and processes,” *Cancer Informatics*, vol. 14, p. 247, 2015.

[3] J. Roski, G. W. Bo-Linn, and T. A. Andrews, “Creating value in health care through big data: opportunities and policy implications,” *Health Affairs*, vol. 33, pp. 1115-1122, 2014.

[4] P. Groves, B. Kayyali, D. Knott, and S. Van Kuiken, “The ‘big data’revolution in healthcare,” *McKinsey Quarterly*, vol. 2, 2013.

[5] A. Shitzer and R. C. Eberhart, *Heat Transfer in Medicine and Biology: Analysis and Applications*. vol. 1, Plenum Pub. Corp., 1985.

[6] H. Trefna and M. Persson, “Heating of deep seated tumours using microwaves radiation,” *ACES 2007*, Verona, Italy, March 19-23, 2007.

[7] X. Wu, B. Liu, and B. Xu, “Theoretical evaluation of high frequency microwave ablation applied in cancer therapy,” *Applied Thermal Engineering*, vol. 107, pp. 501-507, 2016.

[8] J. Mallorqui, A. Broquetas, L. Jofre, and A. Cardama, “Non-invasive active thermometry with a microwave tomographic scanner in hyperthermia treatments,” *Applied Computational Electromag-*

netics Society Journal, vol. 7, pp. 121-127, 1992.

[9] M. M. Paulides, P. R. Stauffer, E. Neufeld, P. F. Maccarini, A. Kyriakou, R. A. Canters, C. J. Diederich, J. F. Bakker, and G. C. Van Rhoon, “Simulation techniques in hyperthermia treatment planning,” *International Journal of Hyperthermia*, vol. 29, pp. 346-357, 2013.

[10] M. Converse, E. J. Bond, B. Veen, and S. C. Hagness, “A computational study of ultra-wideband versus narrowband microwave hyperthermia for breast cancer treatment,” *Microwave Theory and Techniques, IEEE Transactions on*, vol. 54, pp. 2169-2180, 2006.

[11] M. N. Tabassum, I. Elshafiey, and M. Alam, “Enhanced noninvasive imaging system for dispersive highly coherent space,” in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 912-916, 2015.

[12] M. N. Tabassum, I. Elshafiey, and M. Alam, “Efficient techniques to enhance nearfield imaging of human head for anomaly detection,” in *IEEE International Symposium Medical Measurements and Applications (MeMeA)*, pp. 565-569, 2015.

[13] M. N. Tabassum, I. Elshafiey, and M. Alam, “Compressed sensing based nearfield electromagnetic imaging,” in *IEEE International Conference on Control System, Computing and Engineering (ICCSCE)*, pp. 571-575, 2014.

[14] P. Takook, H. D. Trefná, A. Fhager, and M. Persson, “Evaluation of the 3D time reversal method for hyperthermia treatment planning in head and neck tumors,” in *9th European Conference on Antennas and Propagation (EuCAP)*, pp. 1-5. 2015.

[15] Y. Tang, Y. D. Zhang, M. G. Amin, and W. Sheng, “Wideband multiple-input multiple-output radar waveform design with low peak-to-average ratio constraint,” *IET Radar, Sonar & Navigation*, vol. 10, pp. 325-332, 2016.

[16] F. Uysal, M. Yeary, N. Goodman, R. F. Rincon, and B. Osmanoglu, “Waveform design for wideband beampattern and beamforming,” in *IEEE Radar Conference (RadarCon)*, pp. 1062-1066, 2015.

[17] C. Chi and Z. Li, “Fast computation of wideband beam pattern for designing large-scale 2-D arrays,” *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 63, pp. 803-816, 2016.

[18] R. Du, Y. Fan, and J. Wang, “Wideband beamforming based on compressive sensing,” in *5th International Conference on Information Science and Technology (ICIST)*, pp. 387-391, 2015.

[19] M. B. Hawes and W. Liu, “Sparse array design for wideband beamforming with reduced complexity in tapped delay-lines,” *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 22, pp. 1236-1247, 2014.

[20] R. M. C. Mestrom, J. P. van Engelen, M. C. van Beurden, M. M. Paulides, W. C. M. Numan, and A. G. Tijhuis, “A refined eigenvalue-based optimization technique for hyperthermia treatment planning,” in *Antennas and Propagation (EuCAP), 2014 8th European Conference on*, pp. 2010-2013, 2014.

[21] *MathWorks. Simulink.* Available: http://www mathworks.com/

[22] C. M. Furse, “A survey of phased arrays for medical applications,” *Applied Computational Electromagnetics Society Journal*, vol. 21, pp. 365-379, 2006.

[23] M. Donelli, “A real time through the wall imaging method based on a simple transmission line model,” in *IEEE Conference on Antenna Measurements & Applications (CAMA), 2014*, pp. 1-3, 2014.

[24] S. Pan, S. J. Elliott, and D. Vignali, “Comparison of the nonlinear responses of a transmission-line and a filter cascade model of the human cochlea,” in *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA), 2015*, pp. 1-5, 2015.

[25] S. E. Sussman-Fort and J. C. Hantgan, “SPICE implementation of lossy transmission line and Schottky diode models,” *IEEE Transactions on Microwave Theory and Techniques*, vol. 36, pp. 153-155, 1988.

N. Nizam-Uddin received his B.S. degree in Electrical Engineering from NWFP University of Engineering & Technology Peshawar, Pakistan in 2003 and M.S. from Edinburgh Napier University, UK in 2005, majoring Communication. He is currently pursuing his Ph.D. in Electrical Engineering at King Saud University. His research interest includes antenna theory, bio-electromagnetics and microwave engineering.

Ibrahim Elshafiey received his B.S. degree in Communications and Electronics Engineering from Cairo University in 1985. He obtained his M.S. and Ph.D. degrees from Iowa State University in 1992 and 1994, respectively. He is currently a Professor in the Electrical Engineering Department at King Saud University. His research interests include computational electromagnetics, biomedical imaging, communication systems and non-destructive evaluation.

Design of Tilted-Beam Fabry-Perot Antenna with Aperiodic Partially Reflective Surface

You-Feng Cheng, Wei Shao, Xiao Ding, and Meng-Xia Yu

School of Physical Electronics
University of Electronic Science and Technology of China, Chengdu, 610054, China
juvencheng1377@gmail.com, weishao@uestc.edu.cn, xding@uestc.edu.cn, yumengxia@263.net

Abstract **─** A fast synthesis model for designing high-gain tilted-beam antennas with the aperiodic partially reflective surface (PRS) is presented in this paper. In this model, concepts of array synthesis and thinning are introduced to realize the tilted beam and control the sidelobe level (SLL). Since a large number of data need to be processed during the optimization process, an efficient hybrid real-binary bat algorithm (HRBBA) is proposed and embedded in the model to optimize the loop widths and the absence of the square loops on the PRS. As an example, an antenna with an expected tilted beam in the elevation plane (θ = 40°) is manufactured and measured, and both the simulated and measured results show that the antenna can tilt its beam toward the expected direction with good performance.

Index Terms **─** Big data, Fabry-Perot (FP) antenna, Hybrid Real-Binary Bat Algorithm (HRBBA), Partially Reflective Surface (PRS), tilted-beam antenna.

I. INTRODUCTION

Tilted-beam antennas are widely used in modern communication systems, such as in satellite communication systems, wireless LAN (WLAN) systems and mobile communication systems [1]. A common approach to achieve high gain and titled beam is the phased array. However, due to the complex feed system and limited space, the phased array is not suitable for some applications. In the literature, several single element antennas have been proposed with good performance. In [2-5], spiral antennas with tilted beams are analyzed and investigated. These antennas possess asymmetric off-broadside beams in the elevation plane. In [6], a printed star antenna which can generate both doughnut-shaped and tilted beam radiation patterns is presented, and its tilted beam depends on the star flare angle. [7] and [8] present two circularly polarized tilted-beam antennas whose radiation patterns possess a beam pointing angle at 30° and 50°, respectively.

A partially reflective surface (PRS) placed in front of a patch antenna can increase its gain due to the multiple reflections of electromagnetic waves between the two planes [9-11]. Furthermore, this structure possesses a potential to tilt the radiation beam. [12] proposes an antenna system composed of a patch antenna and arrayed loops. The loops are periodically placed on the parasitic substrate which makes the antenna system radiate a tilted high-gain beam. However, there is a lack of fast synthesis model for forming the tilted beam. In [13], a reconfigurable PRS is used for beam steering with the help of RF micro electro-mechanical system (MEMS) switches. The high-gain antenna can control the beam shape among broadside, symmetric conical, and an asymmetric single beam which is either frequency-scanned or steered at the fixed frequency.

This paper presents an efficient design method for the high-gain tilted-beam antenna from the perspective of array synthesis and thinning. Initially, 7 × 7 uniform square strip loops are applied to obtain good resonant characteristics at a fixed frequency of 5.2 GHz. Then, by means of a fast synthesis model, the phase distribution of the electric field over the PRS is changed by optimizing the loop widths, and the radiation pattern is tilted toward an arbitrary expected direction in the elevation plane. This operation not only provides the tilted-beam design, but also increases the design freedom. At last, the concept of array thinning is introduced into the design. Several loops are removed for the purpose of sidelobe level (SLL) reduction. An improved hybrid real-binary bat algorithm (HRBBA) is applied to the optimization of the aperiodic PRS due to the generated big data. As an example, an antenna with an expected tilted beam (θ = 40°) is manufactured and measured, and both the simulated and measured results show that the proposed antenna can tilt the maximum beam in the elevation plane with good performance.

II. FAST SYNTHESIS MODEL

A fast synthesis model is proposed for forming the tilted beam. The flowchart of the fast synthesis model is shown in Fig. 1. There are two interrelated modules in this model, namely the optimization module and the

Submitted On: September 27, 2016
Accepted On: January 13, 2017

simulation module, respectively.

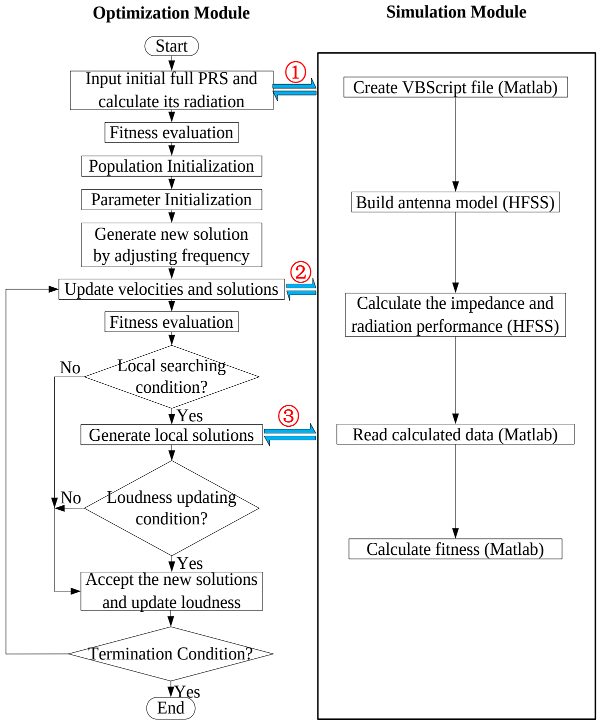


Fig. 1. Flowchart of the fast synthesis model for forming the tilted beam.

Initially, a full PRS (as shown in Fig. 2 (c)) with uniform loop widths, which radiates a high-gain broadside beam, is adopted for the antenna design. Then for the purpose of steering the radiation beam toward any expected directions in the elevation plane, the loop widths and the absence of the square loops should be optimized in the optimization module. As an efficient meta-heuristic method, here the bat algorithm (BA) is selected as a useful tool for the optimization problem. In each generation, for fast and accurate numerical analysis of each newly generated PRS, the BA programed in the Matlab software is linked with a full-wave solver based on the ANSYS High Frequency Structure Simulator (HFSS) in the simulation module. The link between Matlab and HFSS is realized through the macro programing. Matlab inputs the optimization parameters and calls HFSS to simulate the reflection coefficient and far-field pattern of the proposed antenna. Next, the simulated results from HFSS are returned to Matlab and used to calculate the cost functions. After a few generations, a thinned PRS with optimized loop widths is obtained and used for the final tilted-beam antenna design. Note that only the three steps marked with arrows and circled numbers in the Optimization Module are linked to the Simulation Module.

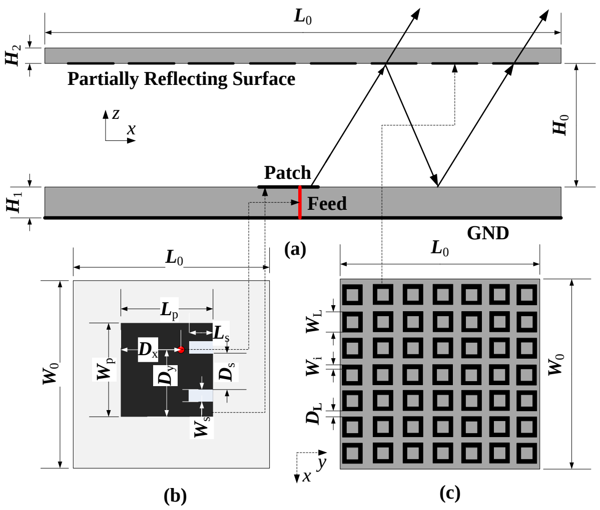


Fig. 2. Initial structure of the proposed antenna. (a) Side view, (b) source antenna, and (c) PRS.

Besides the beam tilt, resonant frequency and gain, the sidelobe level (SLL), which maybe deteriorates in the optimization, is another important objective for the antenna design. Therefore, the objective functions are given as below:

$$objv_1 = \begin{cases} 0 & \text{if } S_{11}(f_0) < -10\text{ dB} \\ 1 & \text{if } S_{11}(f_0) \geq -10\text{ dB} \end{cases}, \tag{1}$$

$$objv_2 = Gain,\ objv_3 = SLL,\ objv_4 = |\theta_t - \theta_0|,$$

where f_0 is the expected resonant frequency, namely 5.2 GHz in this paper, *Gain* is the realized gain, *SLL* is the peak sidelobe level, θ_0 is the expected tilted direction and θ_t is the direction corresponding to the maximum value of the pattern in the *xoz* plane. Here $Objv_1$ represents the resonance characteristics of the antenna, $Objv_2$ and $Objv_3$ yield a high gain and a low SLL, and $Objv_4$ makes the maximum beam tilt to an expected direction. Thus, the fitness function of the optimization problem can be written as:

$$Fitness = \omega_1 \cdot Objv_1 + \omega_2 \cdot Objv_2 + \omega_3 \cdot Objv_3 + \omega_4 \cdot Objv_4, \tag{2}$$

where ω_1, ω_2, ω_3 and ω_4 are weight values, and they are selected as 0.2, 0.2, 0.2 and 0.4, respectively. It is worth noting that θ_0 determines the beam tilting of the design, which increases the design freedom greatly.

III. ANTENNA DESIGN

A. Initial structure and its analysis

The initial structure with uniform square strip loops placed on a parasitic substrate is shown in Fig. 2. An E-shaped patch with increased bandwidth is fabricated on the bottom substrate as the radiation source [14]. The grounded FR4 substrate has a thickness of 3.2 mm and a relative permittivity of 4.4. A coaxial probe penetrates the bottom substrate to provide the excitation for the square patch. On the parasitic substrate, the 7 × 7 uniform square strip loops are periodically arranged to constitute the PRS. The parasitic substrate is Rogers 5880 with a relative permittivity of 2.2 and a thickness of 0.508 mm. The distance H_0 between the source antenna and the PRS is about half wavelength. The antenna operates at 5.2 GHz with a total size of 154 mm × 154 mm × 30.4 mm.

The detail physical parameters of the proposed antenna are: $L_0 = W_0 = 35$mm, $L_\mathrm{p} = 15.2$ mm, $L_\mathrm{s} = 3$ mm, $W_\mathrm{s} = 1.7$ mm, $W_\mathrm{L} = 18.5$ mm, $D_\mathrm{L} = 3.5$ mm, $D_\mathrm{x} = 10.2$ mm, $D_\mathrm{y} = 10.4$ mm, $D_\mathrm{s} = 4$ mm, $H_0 = 28.9$ mm, $H_1 = 3.2$ mm and $H_2 = 0.508$ mm. All these parameters are fixed throughout this paper except the width of each loop W_i ($i = 1, 2, \ldots, 7$) which is set as 2.5 mm initially and will be optimized in the fast synthesis model.

A Fabry-Perot (FP) cavity is formed by the PEC ground plane and the PRS, and the multiple transmission and reflection of the electromagnetic waves generated by the radiation source occur inside the cavity. Thus, the initial antenna would radiate a high-gain broadside beam. Here a scalar-based analysis for the antenna is formulated. Note that the weak mutual coupling effects among the loops are ignored in the analysis.

Assuming the far-field electric field of the source antenna is $E_\mathrm{source}(r,\theta,\phi)$, the total radiated power of the source antenna in the upper half space can be calculated as:

$$P_\mathrm{source} = \frac{1}{2\eta_0}\int_0^{2\pi}\int_0^{\pi/2} |E_\mathrm{source}(r,\theta,\phi)|^2\, r^2 \sin\theta \mathrm{d}\theta \mathrm{d}\phi\,, \tag{3}$$

where η_0 is the free-space impedance.

In terms of the loop m of the PRS, the power flow density of the incident plane wave from the source antenna and towards the direction of the loop is:

$$S(r^m,\theta^m,\phi^m) = \frac{|E_\mathrm{source}(r^m,\theta^m,\phi^m)|^2}{2\eta_0}\,, \tag{4}$$

where (r^m, θ^m, ϕ^m) is the loop position. Then the available power of each loop can be written in terms of the power flow density and the effective area (A_e) of the loop, and that is:

$$P_c^m = S(r^m,\theta^m,\phi^m) A_e \cos\delta^m\,, \tag{5}$$

where δ^m is the angle between the radial vector of the incident power flow density and the unit vector normal to the loop m. According to the transmission coefficient of the loop, the transmitted power of the loop can be given by:

$$P_t^m = P_c^m \left|S_{21}^m\right|^2. \tag{6}$$

Therefore, the transmitted electric field (E^m) of loop m can be obtained:

$$E^m(r,\theta,\phi) = \sqrt{\frac{2\eta_0 P_t^m D_\mathrm{Loop}(\theta,\phi)}{4\pi r^2}}\, e^{j(-kr^m + k\hat{r}\cdot\mathbf{r}^m + \gamma^m)}\,, \tag{7}$$

where $D_\mathrm{Loop}(\theta,\phi) = G_\mathrm{Loop}(\theta,\phi)/\varepsilon_\mathrm{rad}$, ε_rad is the radiation efficiency, $G_\mathrm{Loop}(\theta,\phi) = (\frac{4\pi}{\lambda_0^2})A_e$, λ_0 is the free-space wavelength and γ^m is a phase term determined by the loop. It can be seen from (5)-(7) that the far-field electric field of the PRS is dependent on A_e, $|S_{21}|$, γ^m and the loop position. Moreover, A_e and $|S_{21}|$ determine the far-field amplitude, and γ^m and the loop position determine the phase.

At last, the far-field electric field of the PRS in the upper half space can be calculated as the superposition of the fields of all loops, and it can be expressed as:

$$E_{PRS} = \sum_{m=1}^{M} E_m(r,\theta,\phi) e^{-jkr}\,. \tag{8}$$

B. Tilted-beam realization

Equation (8) represents the far-field electric field of an FP antenna relating to the PRS elements. Note that (8) is similar to the formula of the array factor of a phased array in form. In addition, when the loop width is changed, A_e, $|S_{21}|$ and γ^m of the loop would all be changed but the loop position is fixed. It means that the loop widths determine the amplitudes and phases of the PRS elements. That is to say, by optimizing the loop widths, array synthesis techniques can be used for forming the tilted beam in the design. Unfortunately, accurate calculation of (8) is hard to be realized due to the uncertainty of E_source. Thus, the far-field and input characteristics with different PRS are obtained by the HFSS simulations.

However, one of the problems that may arise in the optimization is the deterioration of the SLL. In order to realize the low sidelobe, the concept of array thinning is introduced into the optimization since one advantage of array thinning is that lower sidelobes can be obtained [15-17]. A thinned PRS with some square loops removed has the capability to reduce the SLL of the optimized antenna. In addition, the thinned PRS increases the gain due to the sidelobe suppression.

The PRS characterized by rectangular unit cells with serial numbers is illustrated in Fig. 3. In order to generate expected tilted beams in the *xoz* plane with good performance, loop widths in each row are fixed to be the same, and those in each column ($W_1 - W_7$) are optimized in a range from 0.25 mm to 6.25 mm. Furthermore, PRS loops, which are listed from 0 to 49 in Fig. 3, are determined to be retained or removed in the optimization. Thus in the whole synthesis, there are 56 optimization variables (7 real variables and 49 binary ones), and the optimization objectives are expressed in (1).

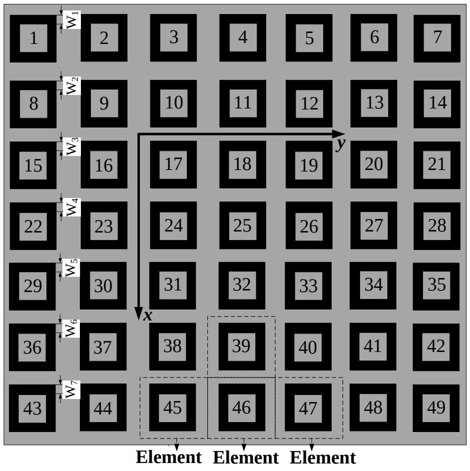


Fig. 3. PRS characterized by rectangular unit cells with serial numbers.

IV. IMPROVED HRBBA

In order to optimize the loop widths and the absence of the square loops, the multi-objective BA [18-20], which can efficiently and reliably process big-data optimization problems, is selected as the synthesis algorithm. Inspired from the echolocation of micro bats, BA has shown that it is very promising and could outperform most existing algorithms.

The basic BA can be briefly described as follows. There are five characteristic parameters in the algorithm: the frequency f, the velocity v, the position x, the loudness A and the rate of pulse emission r. For each bat (i), its frequency f_i, position x_i and velocity v_i are updated by:

$$f_i = f_{\min} + \beta(f_{\max} - f_{\min}), \tag{9}$$

$$v_i^t = v_i^{t-1} + (x_i^{t-1} - x^*) f_i, \tag{10}$$

$$x_i^t = x_i^{t-1} + v_i^t, \tag{11}$$

where t represents the time step. β is a random vector drawn from a uniform distribution and it is in the range of [0, 1]. [$f_{\min}$, $f_{\max}$] is the range of the frequency, and it is set as [0, 100] in this problem. Initially, each bat is randomly assigned a frequency that is drawn uniformly from [$f_{\min}$, $f_{\max}$]. x^* is the current global best solution. In terms of the local search part, when a solution is selected among the current best solutions, a local random walk is used for each bat to generate a new solution, and it can be calculated by:

$$x_{\text{new}} = x_{\text{old}} + \varepsilon A^t, \tag{12}$$

where ε is a random number drawn from [-1, 1]. And the loudness is updated by:

$$A_i^{t+1} = \alpha A_i^t, \; r_i^{t+1} = r_i^0 (1 - e^{-\gamma t}), \tag{13}$$

where α and γ are constants, and they are set as $\alpha = \gamma = 0.9$ in this problem. It can be seen that, compared with other algorithms, the algorithm uses the frequency-based tuning and pulse emission rate to mimic bat behaviors, and this leads to a good convergence and simple implementation. Furthermore, the authors in [14] use a weighted sum to combine multi objectives into a single objective.

In the antenna design, the parameter vector of each solution consists of 56 variables: $x = [W_1, W_2, \ldots, W_7, LA_1, LA_2, \ldots, LA_{49}]$. Here LA_n represents the presence or absence of the nth loop shown in Fig. 3. When there is a square loop in the position, LA_n is set as 1, and it is set as 0 when there is no loop in this position. It is worth noting that W_n and LA_n are of different types. W_n is a continuous variable while LA_n is a discrete one which only can be 0 or 1. The traditional real-number BA is incapable of solving this problem. On the other hand, since the binary coding of the real variables has more dimensions than the real one, the binary algorithms, such as genetic algorithm, are inefficient for the optimization compared to a hybrid-parameter algorithm [21]. An improved HRBBA is proposed here for this case.

The improved HRBBA is divided into two parts, namely the real part and the binary one. Each individual is represented by an ($N_\text{r} + N_\text{b}$)-dimensional vector:

$$\vec{x} = \{x_r^1, x_r^2, \cdots, x_r^{N_r}, x_b^1, x_b^2, \cdots, x_b^{N_b}\}. \tag{14}$$

At each generation, the real and binary parts share the same characteristic parameters and global best solutions. In the real part, the algorithm process is the same as the basic BA. While in the binary part, updated parameters, such as v and x, may be fallen outside their range [0, 1]. To make the parameters update in an alternative manner, the sigmoid limiting transformation is implemented in the algorithm by defining an intermediate variable [22]:

$$S(v^m) = \frac{1}{1 + e^{-v^m}}. \tag{15}$$

The binary BA can be described as follows. Firstly, parameters v and x are initialized as binary-valued ones. Then parameters f and v are also updated by the formulations in (9) and (11). Note that v is limited in a range of [$-V_{\max}$, $V_{\max}$] and $V_{\max}$ is selected as 6 in the proposed design. $S(x)$ can be calculated by (15) and its range is [$1/(1+e^{V\max})$, $1/(1+e^{V\max})$]. The value of $S(v^m)$ represents a probability threshold, and then x can be updated by:

$$x_b^m = \begin{cases} 1, & \text{if } r^m < S(v^m) \\ 0, & \text{if } r^m \geq S(v^m) \end{cases}, \tag{16}$$

where r^m is a random number with a uniform distribution in the range of [0, 1]. From (16), it can be indicated that the probability that the mth bit equals to 1 is $S(v^m)$.

In each generation of the improved HRBBA, after the independent completion of the real and binary parts, the fitness functions can be calculated based on the newly-generated individuals.

Finally, in this design, the HRBBA is embedded into the synthesis as shown in Fig. 1.

In order to verify the optimization performance of the proposed HRBBA, the classic functional testbeds are shown here by using the Rastrigin fitness function:

$$f(x)=\sum_{i=1}^{3}[x_i^2-10\cos(2\pi x_i)+10], \tag{17}$$

where x_1 and x_2 are real variables in a range of [-5,5], x_3 is related to the 16-bit binary part via a binary-to-real mapping:

$$x_3=\frac{10}{2^{16}-1}\sum_{N=1}^{16}2^{N-1}\cdot B_N-5, \tag{18}$$

and B_N is 0 or 1. Thus the total 18-dimentional solution space is composed of two real variables and sixteen binary bits.

The traditional GA, binary particle swarm optimization (BPSO) and hybrid real-binary PSO (HPSO) algorithms presented in [22] are used here to illustrate the superiority of the proposed HRBBA. For GA, the recombination probability is chosen as 0.7. For BPSO and HPSO, the acceleration coefficients c_1 and c_2 are both chosen as 2. The velocities in the real part of HPSO and HRBBA are distributed in a range of [-0.5, 0.5], and the maximum velocities in the binary part of BPSO, HPSO and HRBBA are all chosen as 6. In addition, in HRBBA, the loudness and pulse emission control parameters (α and γ) are selected as 0.95 and 0.95, respectively.

In the optimization, a population with 10 individuals is used for 200 iterations. In Table 1, the minimum, maximum and mean values of the fitness from the 200 executions of GA, BPSO, HPSO and HRBBA are presented. It can be seen that hybrid real-binary algorithms (HPSO and HRBBA) is able to find lower fitness values than the binary algorithms (GA and BPSO), which means the hybrid real-binary algorithms have a more powerful optimizing ability. Moreover, the obtained maximum and mean fitness values of HRBBA are lower than those of HPSO, which means HRBBA is better than HPSO for optimizing such a single-objective function, although the minimum fitness value of HRBBA is a litter higher than that of HPSO. To reflect the convergence characteristics of the algorithms, Fig. 4 shows the best fitness values at each iteration over 200 trials. Obviously, HRBBA has the fastest convergence speed.

Table 1: Final optimized results derived from GA, BPSO, HPSO and HRBBA after 200 executions

	Maximum	Minimum	Mean
GA	11.94	1.06×10^{-5}	3.72
BPSO	8.95	3.46×10^{-6}	3.65
HPSO	0.39	8.52×10^{-6}	1.77×10^{-2}
HRBBA	4.97×10^{-2}	4.75×10^{-5}	1.31×10^{-2}

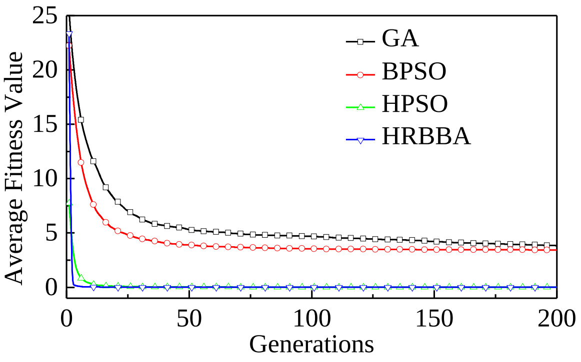


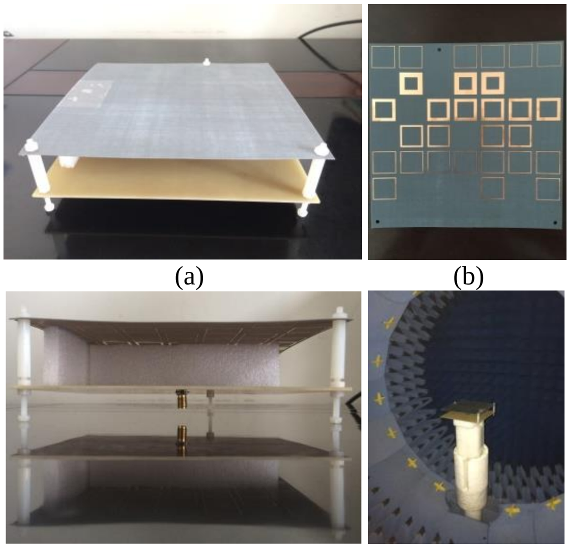
Fig. 4. Comparison of best values averaged over 200 trials for GA, BPSO, HPSO and HRBBA.

V. EXPERIMENT VALIDATION

Here, a design of an antenna with radiation beam toward $\theta = 40°$ is shown to validate the proposed design method. The optimized PRS of the antenna is shown in Table 2, where "1" indicates that the square loop arranged with serial number in Fig. 3 is retained while "0" indicates that the square loop is removed. The optimized FP antenna with aperiodic PRS is fabricated and measured. The prototype of the antenna and the PRS are shown in Fig. 5.

Table 2: Optimized aperiodic PRS (unit: mm)

Widths	W_1	W_2	W_3	W_4	W_5	W_6	W_7
	0.40	4.10	2.60	1.45	0.55	0.70	0
Distribution	1111011001101011111010110110111111 110100010000000						


(a) (b)

(c) (d)

Fig. 5. Photographs of the prototype of the proposed FP antenna. (a) 3D view, (b) aperiodic PRS, (c) front view, and (d) masurement in a Satimo Starlab nearfield measurement system.

The simulated and measured input performance in terms of S_{11} is depicted in Fig. 6. It can be seen that the proposed antenna possesses a bandwidth of 5.09-5.57 GHz with S_{11} lower than -10 dB from the measured results. The relative bandwidth of the proposed antenna is about 10%. Such a wide bandwidth is mainly due to the contribution of the structure of the source antenna. It is noted that the effect of the thinned PRS on the resonance characteristics is very small.

Figure 7 shows the simulated and measured radiation patterns. It can be seen that good agreement is found between the simulated and the measured results. The main lobe points to $\theta = 40°$ in the elevation plane at 5.2 GHz. The measured SLL is -10.33 dB which is a litter higher than the simulated one. This mainly because the upper substrate is very soft and its surface is uneven.

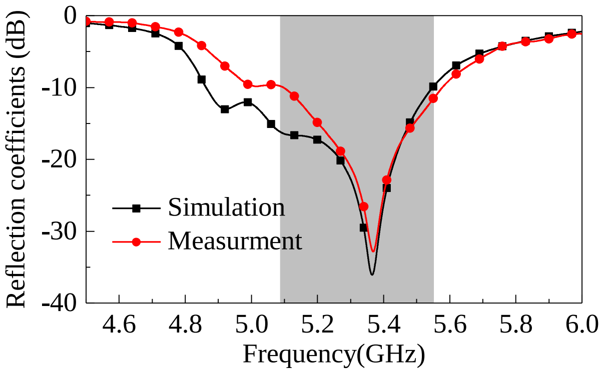


Fig. 6. Measured and simulated reflection coefficients of the optimized FP antenna.

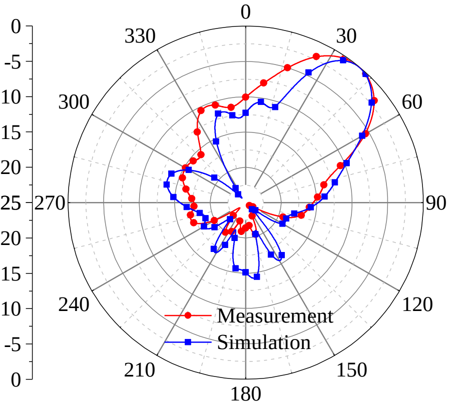


Fig. 7. Measured and simulated radiation patterns of the optimized antenna at 5.2 GHz in the *xoz* plane.

Figure 8 shows the measured peak gain and efficiency of the antenna. The peak gain at 5.2 GHz is 12.50 dBi, which is lower than the simulated one about 1 dB. This is mainly due to the loss of the measurement setup. The peak gain is about 11.12-13.45 dBi in the bandwidth and the maximum gain appears at 5.45 GHz. It is worth noting that the PRS with more square loops can lead to a higher gain. The efficiency of the antenna mainly varies between 50% and 67%. At 5.2 GHz, the efficiency is 62%. The thinned PRS maybe decreases the efficiency of the antenna.

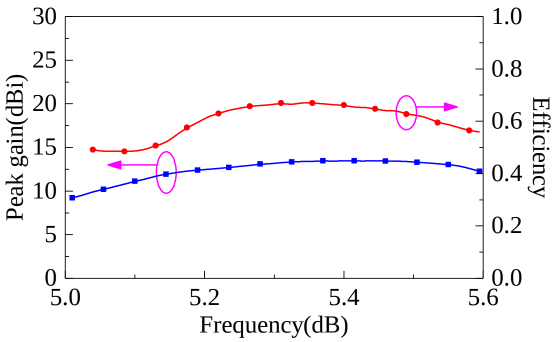


Fig. 8. Measured peak gain and efficiency of the optimized antenna.

At last, a comparison between the proposed design and that in [12] is given in Table 3. The antenna design method, PRS structure, design freedom and antenna performance are listed. An efficient fast synthesis model, which is lacked in [12], is proposed to design the antenna in our design. Based on the fast synthesis model, any required tilted radiation beams can be obtained by optimizing the PRS. In [12], only the antennas with some designated tilted beams can be designed with good performance. With the help of the synthesis model and the improved HRBBA, the aperiodic PRS structures, which afford benefit for the antenna radiation performance, are introduced in the proposed design. The periodic PRS structure restricts the tilted angle in [12.] In addition, the proposed design leads to a wider bandwidth due to the used source antenna, and the antennas designed in [12] have higher realized gains since PRSs with larger size are utilized. Overall, compared with the design method in [12], the proposed design possesses a higher design freedom and design convenience with the powerful synthesis model and the aperiodic PRS structure.

Table 3: Comparison of the antenna design between the proposed method and reference [12]

	Proposed Method	Ref. [12]
Fast synthesis model	Yes	No
Design freedom	Any required tilted beams	Some designated tilted beams
PRS structure	Aperiodic	Periodic
Antenna performance	Wider bandwidth	Higher gain

VI. CONCLUSION

In this paper, an efficient fast synthesis model is proposed to design a high-gain tilted-beam FP antenna with the aperiodic PRS. The FP antenna is analyzed and the concepts of array synthesis and thinning are introduced into the presented model. In addition, an improved HRBBA is proposed to optimize the loop widths and the absence of the square loops on the PRS. Experiment validation with an expected tilted beam (θ = 40°) shows the good performance and powerful design flexibility of the method.

ACKNOWLEDGMENT

This work was supported by the National Natural Science Foundation of China (No. 61471105 and 61401065) and the 973 Project of China (No. 613273).

REFERENCES

[1] M. Arrebola, E. Carrasco, and J. A. Encinar, "Beam scanning antenna using a reflectarray as sub-reflector," *Applied Computational Electromagnetics Society Journal*, vol. 26, pp. 473-483, 2011.

[2] H. Nakano, Y. Shinma, and J. Yamauchi, "A monofilar spiral antenna and its array above a ground plane—Formation of a circularly polarized tilted fan beam," *IEEE Trans. Antennas Propagat.*, vol. 45, pp. 1506-1511, Oct. 1997.

[3] H. Nakano, J. Eto, Y. Okabe, and J. Yamauchi, "Tilted- and axial-beam formation by a single-arm rectangular spiral antenna with compact dielectric substrate and conducting plane," *IEEE Trans. Antennas Propagat.*, vol. 50, no. 1, pp. 17-24, Jan. 2002.

[4] H. Nakano, Y. Okabe, H. Mimaki, and J. Yamauchi, "Monofilar spiral antenna excited through a helical wire," *IEEE Trans. Antennas Propagat.*, vol. 51, no. 3, pp. 661-664, Mar. 2003.

[5] H. Nakano, N. Aso, N. Mizobe, and J. Yamauchi, "Low-profile composite helical-spiral antenna for a circularly-polarized tilted beam," *IEEE Trans. Antennas Propagat.*, vol. 59, pp. 2710-2713, July 2011.

[6] A. Pal, A. Mehta, D. Mirshekar, and P. J. Messy, "Doughnut and tilted beam generation using a single star patch antenna," *IEEE Trans. Antennas Propagat.*, vol. 57, no. 10, pp. 3413-34180, Oct. 2009.

[7] P. Ravindranath, S. Raghavan, and P. H. Rao, "Circularly polarized tilted beam microstrip antenna," *Computing Communication & Networking Technologies (ICCCNT), 2012 Third International Conference*, pp. 26-28, July 2012.

[8] M. Lakshmi, A. Ambika, and P. H. Rao, "Circularly polarized tilted beam monopole antenna," *Communications and Signal Processing (ICCSP), 2014 International Conference*, pp. 973-976, Apr. 2014.

[9] A. P. Feresidis and J. C. Vardaxoglou, "High gain planar antenna using optimized partially reflective surfaces," *Proc. Inst. Elect. Eng. Microwave Antennas Propagat.*, vol. 148, pp. 345-350, Dec. 2001.

[10] F. Costa, A. Monorchio, and G. Manara "An overview of equivalent circuit modeling techniques of frequency selective surfaces and metasurfaces," *Applied Computational Electromagnetics Society Journal*, vol. 29, no. 12, pp. 960-976, Dec. 2014.

[11] A. Pirhadi, H. Bahrami, and A. Mallahzadeh, "Electromagnetic band gap (EBG) superstrate resonator antenna design for monopulse radiation pattern," *Applied Computational Electromagnetics Society Journal*, vol. 27, no. 11, pp. 908-917, Nov. 2014.

[12] H. Nakano, S. Mitsui, and J. Yamauchi, "Tilted-beam high gain antenna system composed of a patch antenna and periodically arrayed loops," *IEEE Trans. Antennas Propagat.*, vol. 62, no. 6, pp. 2917-2925, June 2014.

[13] H. Moghadas, M. Daneshmand, and P. Mousavi, "MEMS-tunable half phase gradient partially reflective surface for beam-shaping," *IEEE Trans. Antennas Propagat.*, vol. 63, no. 1, pp. 369-373, Jan. 2015.

[14] A. A. Razzaqi, M. Mustaqim, and B. A. Khawaja, "Wideband E-shaped antenna design for WLA applications," *2013 IEEE 9th Int. Conf. on Emerging Technologies (ICET)*, pp. 1-6, Dec. 2013.

[15] R. L. Haupt, "Adaptive arrays," *Applied Computational Electromagnetics Society Journal*, 1983.

[16] R. L. Haupt, "Thinned interleaved linear arrays," *ACES International Conference on Wireless Communications and Applied Computational Electromagnetics*, vol. 53, pp. 2858-2864, Sep. 2005.

[17] R. L. Haupt, "Genetic algorithm applications for phased arrays," *Applied Computational Electromagnetics Society Journal*, vol. 21, no. 3, pp. 325-336, Nov. 2006.

[18] X. S. Yang, *A New Metaheuristic Bat-Inspired Algorithm*. In: J. R. Gonzalez, et al. (eds), Nature Inspired Cooperative Strategies for Optimization (NISCO 2010), Springer, Berlin, pp. 65-74, 2010.

[19] A. H. Gandomi, X. S. Yang, A. H. Alavi, and S. Talatahari, "Bat algorithm for constrained optimization tasks," *Neural Computing and Applications*, pp. 1239-1255, 2012.

[20] X. S. Yang, "Bat algorithm for multi-objective optimization," *International Journal of Bio-Inspired Computation*, pp. 267-274, 2013.

[21] E. BouDaher and A. Hoorfar, "Electromagnetic optimization using mixed-parameter and multi-

objective covariance matrix adaptation evolution strategy," *IEEE Trans. Antennas Propag.*, vol. 63, no. 4, pp. 1712-1724, Apr. 2015.

[22] N. Jin and Y. Rahmat-Samii, "A particle swam optimization algorithm with hybridized real and binary parameters," *ACES Conference on Optimization Techniques for Electromagnetic Applications*, 2008.

You-Feng Cheng was born in Anhui, China, in 1989. He received his B. E. degree in Electronic Information Science and Technology from the Chengdu University of Information Technology, Chengdu, China, in 2012. He is currently working toward the Ph.D. degree in Radio Physics at the University of Electronic Science and Technology of China (UESTC), Chengdu, China. His main research interests include phased arrays, evolutionary algorithms, and reconfigurable antennas.

Wei Shao received the M. Sc. and Ph. D. degrees in Radio Physics from UESTC, Chengdu, China, in 2004 and 2006, respectively. He joined the UESTC and is now a Professor. He has been a Visiting Scholar in the Electromagnetic Communication Laboratory, Pennsylvania State University in 2010. His research interests include computational electromagnetics and antenna design.

Xiao Ding received the B.S. and M.S. degrees in Electronic Engineering, from Guilin University of Electronic Science and Technology (GUET), Guilin, China, and the Ph.D. degree in Radio Physics from University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2014. He joined the UESTC in 2014, where he is currently an Associate Professor. In 2013, he was a Visiting Scholar at the South Dakota School of Mines and Technology, SD, USA. From June 2016 to June 2017, he was a Visiting Scholar at the University of Houston, TX, USA. His research interests include reconfigurable antennas and its' applications and phased arrays.

Meng-Xia Yu received the M.Sc. degree in Electronic Engineering and Ph.D. degree in Radio Physics from UESTC, Chengdu, China, in 1997 and 2006, respectively. She joined UESTC in 1997 and is now an Associate Professor. She has been a Visiting Scholar in the University of California, San Francisco, from 2011 to 2012. Her research interests include microwave circuits and antenna design.

Distributed Markov Chain Monte Carlo Method on Big-Data Platform for Large-Scale Geosteering Inversion Using Directional Electromagnetic Well Logging Measurements

Qiuyang Shen [1], Xuqing Wu [2], Jiefu Chen [1], and Zhu Han [1]

[1] Department of Electrical and Computer Engineering
University of Houston, Houston, Texas 77204, USA
qshen4@uh.edu, zhan2@uh.edu, jchen82@central.uh.edu

[2] Department of Information and Logistics Technology
University of Houston, Houston, Texas 77204, USA
xwu8@central.uh.edu

Abstract **─** Inversion problems arises in many fields of science focusing on the process that explores the causal factors from which a set of measurements are observed. Statistical inversion is an alternative approach compared to deterministic methods with better capability to find optimal inverse values. Due to the increasing volume of data collections in the oil and gas industry, statistical approaches show its advantage on the implementation of large-scale inverse problems. In this paper, we address on the solution of big-data-scale inverse problems. After examining both conventional deterministic and statistical methods, we propose a statistical approach based on the Markov Chain Monte Carlo (MCMC) method and its implementation with the scalable dataset on the big data platform. The feasibility and methods to apply statistical inversion on the big data platform is evaluated by examining the use of parallelization and MapReduce technique. Numerical evidence from the simulation on our synthetic dataset suggests a significant improvement on the performance of inversion work.

Index Terms **─** Big data, geosteering, MapReduce, MCMC, multiple chains, well logging.

I. INTRODUCTION

Obtaining a reliable and detailed information about the earth's subsurface is of great challenging and requirement for the scientists and engineers to figure out the interior structures and then to be served for economic exploitation or geological prediction [1]. Inversion process is an organized set of mathematical techniques for projecting data to obtain knowledge about the physical world on the basis of inference drawn from observations [2]. The observations consist of a set of measurements from the real world. In this article, we are focusing on the background of geosteering inverse problems. Geosteering is a technique to actively adjust the direction of drilling, often in horizontal wells, based on real-time formation evaluation by using directional electromagnetic (EM) logging measurements [3]. This process enables drillers to efficiently reach the target zone and actively respond while drilling to geological changes in the formation so they can maintain the maximal reservoir contact. Among different types of logging techniques, azimuthal resistivity logging-while-drilling (LWD) tools are widely used in geosteering to provide electromagnetic measurements [3, 4]. A schematic model diagram of an azimuthal resistivity LWD tool is shown in Fig. 1.

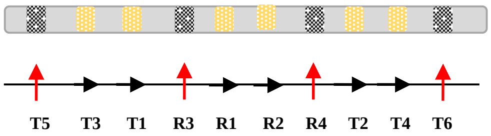


Fig. 1. The structure and schematic of an azimuthal resistivity LWD tool. T1, T2 T3 and T4 are the transmitters whose moments are with the tool axis, while T5 and T6 are transverse antennas that perpendicular to the tool axis. Similarly, R1 and R2 are the receivers directing along the tool axis. R3 and R4 are the vertical receiver antennas with directional sensitivity.

A set of antennas with different polarizations and working frequencies are installed on the azimuthal resistivity tool. They play the roles as transmitters and receivers that always in pairs and symmetric to the center of the tool. Each pair of symmetric antennas can provide a group of electromagnetic signals while the tool is rotating and drilling forwards. The signals which have a sensitivity to layer interfaces and medium resistivity are collected as measurements, and serve for inverse work.

Submitted On: December 24, 2016
Accepted On: January 28, 2017

The formation of an earth model is usually reconstructed through inverted parameters such as depth-to-boundary and layer resistivity. However, a huge uncertainty of parameters is introduced due to constraint propagation of EM signals when the tool is far from detecting boundary [5]. In recent years, a new generation of azimuthal resistivity logging-while-drilling tools, which have longer spacing between antennas and lower working frequency, with a much larger depth of investigation have emerged on the market.

Though more advanced techniques are applied on logging tools, the geosteering inverse problems are always challenging since given measurements are generally finite and imprecise and infinite number of solutions to the inverse problem exists due to the sparsity and uncertainty of the data, and incomplete knowledge of operating circumstance [6]. In the view of Bayes, the optimal solution of inverse parameters can be extracted from the solution set by the statistical characteristics of model structures, whereas the analytical solution for the posterior distribution is not always available anymore [7, 8]. In practice, the Markov Chain Monte Carlo (MCMC) method is one of the most popular sampling approaches to draw the samples from an unknown distribution. The MCMC method guarantees asymptotically exact solution for recovering the posterior distribution [9], though the computational cost is inevitably high while the most MCMC algorithms suffer from low acceptance rate and slow convergence with long burn-in periods. Launching multiple Markov chains is a choice to improve the sampling quality since that the samples from an individual chain have serial independence. A multiple chain MCMC method weakens the correlation of each sample, and thus improves the possibility of convergence [10]. However, the computational cost will not be reduced while traditionally each Markov chain runs in sequence. Fortunately, the strategy of multiple chains is well suited to the platform with support of parallel computing. It comes to be one of our interests to deploy the multiple chains MCMC inversion in an efficient way.

Meanwhile, more accurate and deeper investigation can produce a larger amount of measuring data. Though these data play a critical role in a better reconstruction of earth model, the sheer volume of data and the high dimensional parameter spaces involved in the inversion process means that the statistical methods can scarcely keep up with the demand to deliver in-time information for the decision-making. The current situation conforms to the trend of the age of big data, in which data's volume, variety and velocity are defined as three primary criterions [11]. Hence, the objective of this article is to find out and examine an appropriate strategy to apply the big data techniques on the scalable statistical MCMC method of multiple chains. Moreover, we will show how this scheme can be fit into the MapReduce programming model to take advantage of the potential speed up.

II. INVERSION ALGORITHMS

The adjusting of real time geosteering process is determined by the current working condition of the position and attitude of the tool to minimize the gas or oil breakthrough and maximize the economic production. The operations are relying much on the output from geosteering inversion, from which a group of parameters are generated to draw a reconstructed earth model with the constraint of the measurements. In the real job of geosteering, modeling and inversion are applied to 1D model, in which the layer interfaces are assumed infinitely extended and parallel to each other. A group of results containing distance-to-boundaries and resistivity of each layers are collected to represent the earth model at the current tool position. Along with the trajectory of the tool, inversion is conducted at a fixed interval of distance. Thus the inversed model parameters are grouped together to draw a whole subsurface profile.

A. The deterministic inversion methods

The deterministic inversion method in geosteering applications fit the model function to measured data by minimizing an error term between the forward model responses and observed measuring data. Assume a geosteering tool provides N measurements denoted by $m \in R^N$, while $x \in R^M$ represents M earth model parameters inverted from the measured data. A computation model function or so-called forward function $S: R^M \to R^N$ is designed to synthesize N responses from M model parameters. The forward transformation from model parameters to the responses, which is denoted as $S(x) \in R^N$, is a result of the interaction of a physical system defined by those model parameters. The objective of the inverse problem is to infer the model parameters through observed measurement. A good agreement between the response of the forward model and measured data will be reached if the inversed parameters of the physical model are accurate. The difference between the forward response and measurements is defined as data misfit $F(x)$, which is defined as:

$$F(x) = S(x) - m. \tag{1}$$

Since both forward responses and measurements are vectors, a cost function is defined as the square of L2 norm of the misfit function $F(x)$ as follows:

$$f(x) = \sum_{i=1}^{N} F_i^2(x) = \|F(x)\|_2^2, \tag{2}$$

where $f(x)$ is the cost function representing the magnitude of the data misfit. Hence, the inverse problem is to find the optimal model parameters, x, which minimize the cost function given a forward model function and measurements. Mathematically, this problem is presented as:

$$\underset{x \in R^M}{\text{minimize}} \;\; f(x). \tag{3}$$

This is an unconstraint nonlinear least-square minimization problem. Many iterative numerical algorithms, such as gradient decent method, Gauss-

Newton method, and the Levenberg-Marquardt algorithm (LMA), have been well established to solve this least-square problem [12, 13].

Though most of the aforementioned optimization algorithms are robust and have satisfying performance for convergence, most inverse problems are ill-posed with infinite solutions given finite measurements. In many nonlinear cases, the solution of gradient-based methods is highly dependent on the initial guess or getting trapped in a local minimum. Improvement can be made by using regularizer as an additional constraint. For example, the objective function can be regularized subjecting to the minimization of neighboring inverse points. For example, total variation (TV) regularized inverse problems can be written as:

$$\underset{x\in R^M}{\text{minimize}}\ \ \|S(x)-m\|_2^2$$
$$\text{s.t}\ \ \left|m-m_{pre}\right|<\lambda, \tag{4}$$

where the constraint λ regularizes the maximum change between neighboring inversed parameters m_{pre} and current model parameters, m [14]. However, this kind of methods rely much on the assumption that the prior underground information has been acquired fully and the ill inverse only happens in some of inverse points. Hence, the limitation of deterministic inversion methods becomes more prominent especially with the increasing scale of the problem.

B. The statistical inversion methods

Statistical methods arise as an alternative approach to deal with many ill-posed scientific inverse problems. This sort of methods based on the Bayes' theorem has attracted many attentions nowadays. It can be concluded as a method to obtain the posterior distribution from which the solution is deduced after combining the likelihood and the prior. The assumptions made by the forward model $y=f(x)$ (y is data and x denotes the earth model parameter) may not include all factors that affect measurements. Suppose the noise is additive and comes from external sources, the relationship between observed outputs $\tilde{y}$ and corresponding model parameters can be represented as:

$$\tilde{y}=f(x)+\varepsilon, \tag{5}$$

where ε denotes an additive noise. The experiments empirically suggest the additive noise usually follows a zero-mean Gaussian random distribution: $\varepsilon\sim\mathcal{N}(0,\sigma^2 I)$. With the given hypothesis to model parameters x and observed data $\tilde{y}$, the likelihood can be deduced as:

$$p(\tilde{y}|x)\sim\mathcal{N}(\tilde{y}-f(x),\sigma^2 I). \tag{6}$$

Suppose the prior distribution of x is governed by a zero-mean isotropic Gaussian distribution such that $p(x)\sim\mathcal{N}(0,\beta^2 I)$. By virtue of the Bayes' formula, the posterior distribution of x is given by:

$$p(x|\tilde{y})\sim\mathcal{N}(\tilde{y}-f(x),\sigma^2 I)\mathcal{N}(0,\beta^2 I). \tag{7}$$

It suggests that the posterior distribution of model parameters x given observations $\tilde{y}$ can be obtained by calculating the product of two Gaussian distributions. The solution of x can be sampled and estimated according to the probability distribution function $p(x|\tilde{y})$. It is an effective way to overcome the shortcomings of deterministic inversion especially when the problems are underdetermined (ill-posed) because of the large parameter space and the sparsity of the measurements. The earth model parameters are determined by sampling from posterior distribution $p(x|\tilde{y})$ while the measurements $\tilde{y}$ have been acquired.

Drawing samples from posterior distribution $p(x|\tilde{y})$ is challenging when f indicates a non-linear mapping relationship between x and $\tilde{y}$ since the analytical solution for the posterior distribution is not always available. In practice, the MCMC method is one of the most popular sampling approaches to draw the samples from an unknown distribution while the state of the chain after a number of steps can reach its equilibrium. Then the samples are selected to draw a desired distribution.

In the next section, we will first elaborate the implementation of the MCMC sampling method by the Metropolis-Hastings algorithm, and then explore the applicability pf the parallel computing scheme that suits MCMC statistical inversion and is scalable to large inverse problems. Finally, an application of the MapReduce model, a prevailing scheme on the big data platform, and its implementation on the statistical inverse problems, will be introduced.

III. LARGE-SCALE GEOSTEERING INVERSE PROBLEMS

The Metropolis-Hastings (MH) algorithm [15] is a popular MCMC method. In brief description, a MH step of invariant distribution $p(x)$ and proposal distribution or jumping function $q(x^*|x)$ involves sampling a candidate value x^* given the current value x according to $q(x^*|x)$. The Markov chain then moves towards x^* with the following acceptance probability:

$$\mathcal{A}(x,x^*)=\min\left\{1,\frac{p(x^*)q(x|x^*)}{p(x)q(x^*|x)}\right\}, \tag{8}$$

otherwise it remains at x. In the random walk MH algorithm, a zero-mean normal distribution is a popular choice of $q(x^*|x)$ as a symmetric candidate-generating function, which helps reduce the moving probability to $p(x^*)/p(x)$. Algorithm 1 presents the MH algorithm for sampling from the posterior distribution:

Algorithm 1: The Metropolis-Hastings algorithm for sampling from $p(x|\tilde{y})$

input: initial value $x^{(0)}$, jumping function $q\left(x^{(i)}\middle|x^{(j)}\right)$

output: $x^{(k)}$ where $k\le K$

begin

```
    Initialize with arbitrary value $x^{(0)}$
    while length of MCMC chain < pre-defined length $K$ do
        Generate $x^{(k)}$ from $q\left(x^{(k)} \middle| x^{(k-1)}\right)$
        $\mathcal{A}\left(x^{(k)}, x^{(k-1)}\right) = \min\left\{1, \frac{p\left(x^{(k)}|\tilde{y}\right)}{p\left(x^{(k-1)}|\tilde{y}\right)}\right\}$
        Generate $\mathcal{A}_0$ from uniform distribution $\mathcal{U}(0,1)$
        if $\mathcal{A}_0 < \mathcal{A}\left(x^{(k)}, x^{(k-1)}\right)$ then
            keep $x^{(k)}$
        else
            $x^{(k)} = x^{(k-1)}$
        end
        save $x^{(k)}$ in the chain
    end
end
```

A. Distributed multiple chains MCMC methods

Although the MCMC method guarantees asymptotically exact recovery of the posterior distribution as the number of posterior samples grows, it may suffer from a large number of “burn-in” steps to reach the equilibrium and slow convergence [10]. Noted that the time cost may be prohibitively high for the inverse problem when the forward model computation is required by every sample drawn from the MH, which may take at least $O(\mathrm{N})$ operations to draw one sample [9]. Meanwhile, with the increasing space of model parameters, burn-in period may reach over thousands of steps. As both data volume and the space of model parameters increase explosively, a dispersion manner of the MCMC algorithms, which distributes computation to multi-processor and multi-machine, is urgently required. It is in our interests to investigate and provide answers to these questions on how to extend the proposed framework to serve for our statistical inverse problems that involve in big volume of data measurements, variety of measurement types and require high speed on data processing.

The essence of the distributed implementation is enlightened by the application of multiple chains (or multiple sequence) of the MCMC method proposed by Gelman and Rubin, 1992. It is based on the idea that the samples from a sequence of chain has a tendency to be unduly influenced by slow-moving realization of the iterative simulation. Whereas a multiple starting points are needed to avoid inferences being influenced [10]. The evidence apparently shows in geosteering statistical inversions and it is not generally possible to reach convergence by running a single chain MH algorithm. The main difficulty is that the random walk can remain for many iterations in a region heavily influenced by the starting distribution, especially in the case of sampling multidimensional random variables with a strong correlation. Hence, a multiple chains MCMC method has been proposed by starting multiple independent chains of the iterative sampling at multiple starting points. The target distribution and the estimation of model parameters can be obtained more quickly by the samples using between-sequence as well as within-sequence information [10]. Once all sequences reach the maximum chain length, an easy estimation of model parameters can be accomplish by sampling results from each chain instead of an inference from the time-series structure of samples from one single chain.

However, efficiency cannot be achieved while multiple chains are running on a single thread since all sequences are queued up and executed one by one. Due to a rapid expansion of computer science, the parallel computing technique is well developed serving for the parallelized tasks [16]. The parallelism can be supported not only within the level of a single machine with multiple-cores or multiple-processors, but also on the scale of clusters, grids and clouds. The multiple chains MCMC method is able to take full advantage of data and task parallelism. A simple strategy of a distributed MCMC method is built on the parallelization of multiple chains, which distributes the data and the task of the MCMC sampling method to multiple processing units.

Nevertheless, hundreds of measured data are collected for one measuring point by the geosteering tool and tons of measurements will be yield while the operating region may extend to thousand feet long. Therefore, the computation structure of the geosteering inverse problem needs to be able to accommodate multiple level of parallelism. The parallelization scheme of geosteering inversion is depicted in Fig. 2. The point-wise inversion tasks with respective measured data are distributed to multiple cluster nodes by the master node. Within each task, multiple chains are launched and run in parallel on the multiple processors or cores. Saturated with huge volumes of collected data, the computation power is well utilized by distributing the task and data to every computing node and thread. A facile solution for solving the large-scale statistical inverse problem in parallel is to take the advantage of the big data platform.

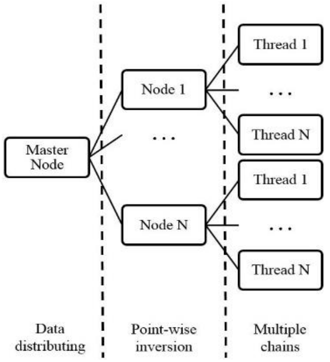


Fig. 2. Parallelization of large-scale statistical geosteering inverse problems on distributed clusters.

B. Large-scale geosteering inversion on MapReduce

To keep up with the requirement for large-scale parallel computing, we need a platform that allows easy access to data and is transparent to its scheduling process. Moreover, the platform should be able to dynamically allocate it computing resources, such as the number of cluster nodes or processing units, and achieve high scalability with low implementation costs. MapReduce is an idea model to meet these requirement.

MapReduce is a programming model and an associated implementation for processing and generating large data sets with a parallel, distributed algorithm on a cluster [17]. It serves for a purpose that allows programmers without any experience with parallel and distributed systems to easily utilize resources in a distributed computing environment. The model is composed of a map function which performs a preprocessing to the data collection, and a reduce function that performs a summary operation. A simple diagram of MapReduce is shown in Fig. 3. A master node is responsible for generating initial starting value set and assigning each starting value to a map function. The mappers filter and sort the allocated data and generate associative value. These intermediate outputs are sent to the reducers which aggregate the data and make calculation for the final outputs.

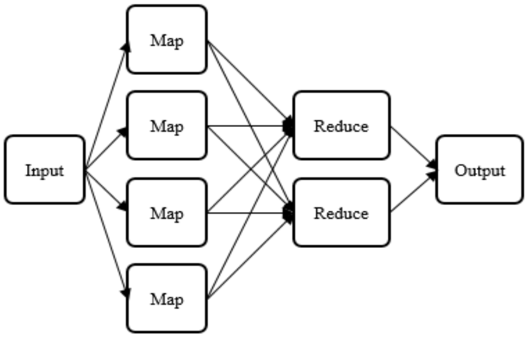


Fig. 3. Diagram of MapReduce programming model.

In the case of geosteering inverse problems, the MapReduce programming model can be suitably applied on the multiple chains MCMC method. Initially, multiple sets of initial values of model parameters are generated by a master node as the starting value for multiple chains. Each set is sent to a mapper where MCMC are running simultaneously for all sets. Samples generated by each chain are cached locally on map nodes for subsequent map-reduce invocations. After intermediate data is collected, the master node coordinates the mappers and reducers and in turn invokes the reducers to process samples from Markov chains. The reducer will calculate the variance within and between chains, evaluate convergence and statistical characteristics of all samples and output them as the final result.

IV. EXAMPLES AND DISCUSSIONS

In this section, we will evaluate the computation and convergence performance of the parallel multi-chain statistical geosteering inversion. We will take a synthetic earth model as an instance to examine the scheme. It is noted that the synthetic data is generated based on analytical EM wave solutions for the 1D multi-layer model with dipole transmitters and receivers. The speed of inversions depends largely on the forward program. In the experiment, the forward model is treated as a black box that returns a group of measurements given earth model parameters. The test ran on a DELL PowerEdge T630, with dual Intel Xeon E5-2667 V3 3.2GHz 20M Cache 8C/16T, and 8x16GB RDIMM 2133MT/s.

To verify the computational efficiency of the parallel computing, we firstly examine the time cost by launching multiple chains at fixed length on different numbers of processing units. The inversion is conducted on a single point of a three-layer model with five parameters: the resistivities of the upper, middle and lower layers are 10 Ω·m, 50 Ω·m and 1 Ω·m, respectively. The distance to the upper and lower boundary is 7 ft and 10 ft. Figure 4 shows the relationship between the time cost and the number of processing units. We archived significant performance gain when increasing processing unit from one to four. And the improvement is observed for all chain lengths. We also noted that the performance gain became less when adding more processing units. We think the execution latency variation is due to other shared resource conflicts when running the test on a shared memory platform.

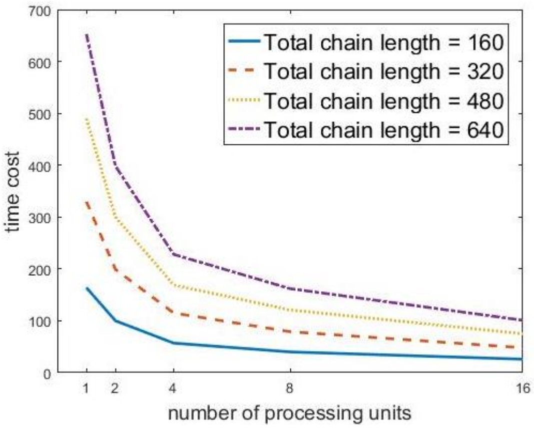


Fig. 4. Comparison of time cost by different total length of chain on different number of processing units.

To examine the convergence of inversion, we run two instances separately. First, we tested a single chain with 640 iterations. Second, we run eight chains simultaneously with the same number of iterations. We compare the accuracy of the inversed results through the unnormalized parameters misfit, a L2 norm of the difference between the inverted model parameters and the true model parameters. In Fig. 5, the result shows that both instances have reached the equilibrium state after a burn-in period. The parameter misfit of multiple chains is smaller and inversed results are more close to the ground truth in comparison with the single chain result. The test demonstrates the advantage of using multiple chains to reduce negative effects of parameter correlation and improve convergence accuracy. It should be noted that the time cost for running eight chains simultaneously is nearly the same as running a single chain on our parallel testing bed.

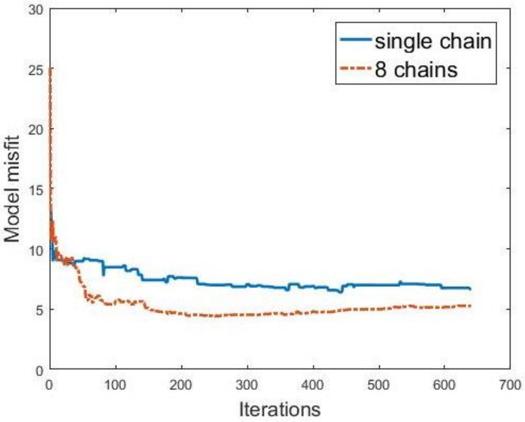


Fig. 5. Comparison of model misfit convergence.

Our next test is to reconstruct a subsurface profile by applying the parallel multi-chain MCMC statistical inversion to a three-layer model. The earth model in Fig. 6 (a) is a three-layer model with the resistivities of 10 Ω·m, 50 Ω·m and 1 Ω·m in the upper, middle and lower layer respectively. The central green line indicates the tool's navigation trajectory. In this case, we assume the tool dip angle is fixed as 90 degrees. The inversion is conducted at each 1ft while the total working region extends to 80 ft. The depth to either the upper or lower boundary is varying on different position. The largest depth to boundary (D2B) is around 18 ft and the smallest is around 2 ft. Both operations, point-wise inversion and multi-chain MCMC sampling, are parallelized. We validate the feasibility of the implementation for large-scale inverse problems on the big data scheme. Figure 6 (b) shows the inverse result which provides a successful and satisfying reconstruction of the earth model compared with the Fig. 6 (a).

Figure 7 also shows the quantified interpretation for the uncertainty of inversed results. The variance at each inversed point denotes the uncertainty of the estimation of the parameter. One can see from Fig. 7 that when the tool is far from the top boundary (the beginning part), the earth model cannot be well reconstructed. In other words, the uncertainty involved in solving the inverse problem is high. The reason is that the sensitivity of the measurements is relatively poor when the tool is far from the layer interface. As the tool moves forward and the top boundary bends downward, the inversion can clearly resolve both the upper and lower boundaries.

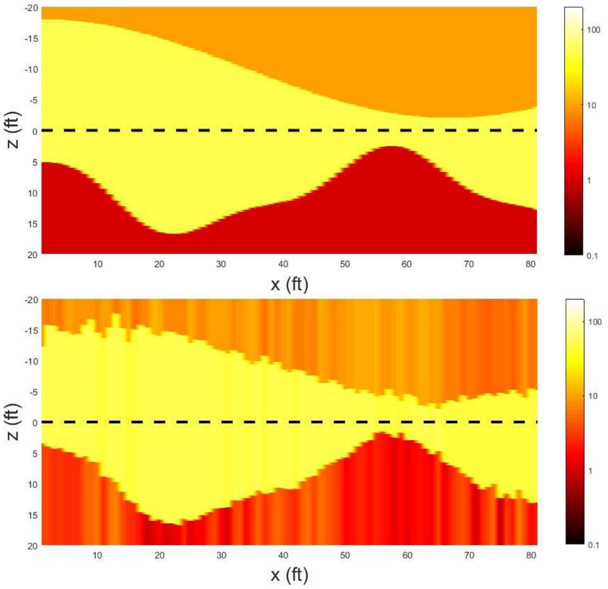


Fig. 6. (a) A synthetic three-layer earth model in 80 ft horizontal region, and (b) the inverse earth model.

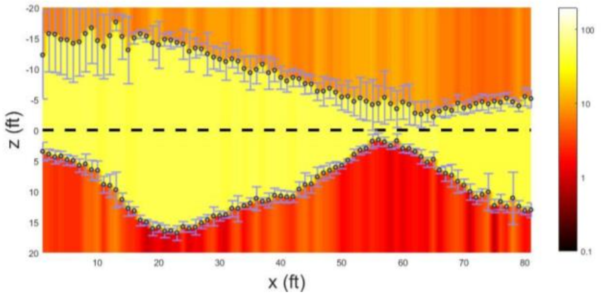


Fig. 7. The inverse earth model with uncertainty.

V. CONCLUSION

Since more advanced measuring technologies have been commercialized and pushed into the market around these years, the increasing volume of measured data undoubtedly will require a more efficient solution for solving large-scale inverse problems. The proposed work in this articles takes full advantage of the latest technology in big data in exploration of computationally efficient methods for solving statistical inverse problems. The multiple chains MCMC method is well suited for using the MapReduce framework in a distributed or cloud environment. Our simulation verifies the feasibility of solving large-scale inverse problem on a big data platform. The techniques developed in this research could have a large positive impact in exploring computationally efficient method to solved real-time statistical inverse problem.

REFERENCES

[1] M. Moorkamp, *Integrated Imaging of the Earth: Theory and Applications*. John Wiley & Sons, 2016.

[2] M. S. Zhdanov, *Inverse Theory and Applications in Geophysics*. Elsevier, 2015.

[3] Q. Li, D. Omeragic, L. Chou, L. Yang, and K. Duong, "New directional electromagnetic tool for proactive geosteering and accurate formation evaluation while drilling," *SPWLA 46th Annual Logging Symposium*, *Society of Petrophysicists and Well-Log Analysts*, 2005.

[4] M. S. Bittar, "Electromagnetic wave resistivity tool having a tilted antenna for determining the horizontal and vertical resistivities and relative dip angle in anisotropic earth formations," U.S. Patent No. 6,163,155, 19 Dec. 2000.

[5] W. G. Lesso Jr. and S. V. Kashikar, "The principles and procedures of geosteering," *SPE/IADC Drilling Conference*, *Society of Petroleum Engineers*, 1996.

[6] K. Key, "1D inversion of multicomponent, multifrequency marine CSEM data: Methodology and synthetic studies for resolving thin resistive layers," *Geophysics*, vol. 74, no. 2, F9-F20, 2009.

[7] A. M. Stuart, "Inverse problems: A Bayesian perspective," *Acta Numerica*, vol. 19, pp. 451-559, 2010.

[8] B. T. Knapik, A. W. van Der Vaart, and J. H. Van Zanten, "Bayesian inverse problems with Gaussian priors," *The Annals of Statistics*, pp. 2626-2657, 2011.

[9] W. Neiswanger, C. Wang, and E. Xing, "Asymptotically exact, embarrassingly parallel MCMC," *arXiv preprint arXiv: 1311.4780*, 2013.

[10] A. Gelman and D. B. Rubin, "Inference from iterative simulation using multiple sequences," *Statistical Science*, pp. 457-472, 1992.

[11] J. Manyika, M. Chui, B. Brown, J. Bughin, R. Dobbs, C. Roxburgh, and A. H. Byers, "Big data: The next frontier for innovation, competition, and productivity," 2011.

[12] K. Levenberg, "A method for the solution of certain problems in least squares," *Quart. Appl. Math.*, vol. 2, pp. 164-168, 1944.

[13] D. Marquardt, "An algorithm for least-squares estimation of nonlinear parameters," *SIAM J. Appl. Math.*, vol. 11, pp. 431-441, 1963.

[14] Q. Shen, J. Chen, and Z. Han, "Total variation denoising of azimuthal resistivity inversion in geosteering," *SEG Technical Program Expanded Abstracts 2016*, *Society of Exploration Geophysicists*, pp. 719-723, 2016.

[15] W. K. Hastings, "Monte Carlo sampling methods using Markov chains and their applications," *Biometrika*, vol. 57, no. 1, pp. 97-109, 1970.

[16] L. Murray, "Distributed Markov chain Monte Carlo," *Proceedings of Neural Information Processing Systems Workshop on Learning on Cores, Clusters and Clouds*, vol. 11, 2010.

[17] J. Dean and S. Ghemawat, "MapReduce: simplified data processing on large clusters," *Communications of the ACM*, vol. 51, no. 1, pp. 107-113, 2008.

Qiuyang Shen received his B.S. degree in Telecommunication Engineering from Huazhong University of Science and Technology in 2015. Now he is a second year Ph.D. student at the Department of Electrical and Computer Engineering, University of Houston, co-advised by Dr. Jiefu Chen and Dr. Zhu Han. His research interests include geophysical inversions, data optimization and machine learning.

Xuqing Wu received the B.S. degree in Electronic Engineering from the University of Science and Technology Beijing in 1995, the M.S. degree in Computer Science from the Carleton University in 2001, and the Ph.D. degree in Computer Science from University of Houston in 2011.

From 2001 to 2007, he was a Software Engineer for EMC/INTEL/HP. He was a Data Scientist of Schlumberger before joining the University of Houston in 2015. Currently, he is an Assistant Professor in the Information and Logistics Department at the University of Houston, Texas. His research interests include multi-physics interpretation, machine learning, computer vision, high performance, mobile and cloud computing.

Jiefu Chen is an Assistant Professor at the Department of Electrical and Computer Engineering, University of Houston. Previously, he worked for 4.5 years as a Staff Scientist for Weatherford International. His current research interests include electromagnetic modeling and inversion, subsurface wireless communication, and well logging. Chen holds a B.S. degree in Engineering Mechanics from Dalian University of Technology and a Ph.D. degree in Electrical Engineering from Duke University.

Zhu Han received the B.S. degree in Electronic Engineering from Tsinghua University, in 1997, and the M.S. and Ph.D. degrees in Electrical and Computer Engineering from the University of Maryland, College Park, in 1999 and 2003, respectively.

From 2000 to 2002, he was an R&D Engineer of JDSU, Germantown, Maryland. From 2003 to 2006, he was a Research Associate at the University of Maryland. From 2006 to 2008, he was an Assistant Professor at Boise State University, Idaho. Currently, he is a Professor in the Electrical and Computer Engineering Department as well as in the Computer Science Department at the University of Houston, Texas. His research interests include wireless resource allocation and management, wireless communications and networking, game theory, big data analysis, security, and smart grid. Han received an NSF Career Award in 2010, The Fred W. Ellersick Prize of the IEEE Communication Society in 2011, the EURASIP Best Paper Award for the Journal on Advances in Signal Processing in 2015, IEEE Leonard G. Abraham Prize in the field of Communications Systems (Best Paper Award in IEEE JSAC) in 2016, and several Best Paper Awards in IEEE conferences. Currently, Han is an IEEE Communications Society Distinguished Lecturer.

A Double Band-Notched UWB Antenna for Flexible RF Electronics

Chuicai Rong [1,2], Wei Xiao [1], Yuehang Xu [1*], and Mingyao Xia [1]

[1] School of Electrical Engineering
University of Electronic Science and Technology, Chengdu, 61173, P.R. China
*yuehangxu@uestc.edu.cn

[2] School of Physics and Electronic Information
Gannan Normal University, Ganzhou, Jiangxi, 341000, China

Abstract **─** A double band-notched UWB compact antenna based on ultrathin liquid crystal polymer (LCP) substrate is designed and experimented for flexible electronics in this paper. The antenna is constituted by swallow tail radiation patch and trapezoid ground. Two elliptic single complementary split-ring resonators (ESCSRRs) are used to realize double band-notched characteristic in UWB band. The measurement results show that the VSWRs of antennas are lower than 2 in the whole UWB band (3.1 GHz-10.6 GHz) except the two band-notched frequency-bands in 3.7 GHz-4.2 GHz and 5.15 GHz-5.825 GHz. The bending performance of band-notched UWB antenna are measured. The measured radiation patterns indicate that the antenna is an omni-directional antenna both at flat and bent circumstances. The overall dimension of the antenna is 27 mm × 21 mm × 0.05 mm. These results suggest that this compact mechanical flexible antenna would be useful for flexible UWB wireless system.

Index Terms **─** Band-notched, ESCSRR, flexible electronics, Liquid Crystal Polymer (LCP), UWB antenna.

I. INTRODUCTION

Since the ultra-wideband (UWB) spectrum (3.1 to 10.6 GHz) is allocated by the US Federal Communications Commission (FCC) to use for commercial communication purposes in 2002, high performance UWB antenna are widely concerned [1-2]. There are some narrow band interferences in UWB, such as C-band satellite communications system (3.7 GHz-4.2 GHz) and Wireless Local Area Network (5.15 GHz-5.825 GHz). Based on this situation, the best solution to reduce these interferences is introducing an antenna with band-notched performance instead of using narrow band filters [3-4]. In addition, many UWB applications need more than one band-notches in complex electromagnetic environment, so it is essential to design multiple band-notched UWB antennas [5-8].

Recently, as the increasing demands of wireless communication in wireless wearable and implanted electronics, compact microwave antenna with mechanical flexibility is become popular [9-10]. Considering the superiority of UWB wireless communication systems (i.e., high transmission data rate, low transmission power, etc.), design of compact band-notched UWB antenna with mechanical flexibility would be meaningful to wireless communication in wearable and implanted electronics [2].

In this paper, in the aim of developing mechanical flexible UWB antenna, a double band-notched UWB antenna is designed. To realize compact size, ultra-thin LCP with copper-clad laminated, which has been proved excellent mechanical flexibility [11], is used as the substrate. The antenna is constituted by swallow tail radiation patch and trapezoid ground. And the band-notched characteristics are realized by etching two elliptic single complementary split-ring resonators (ESCSRRs) on radiation patch. The performances are measured both at flat and bent circumstances.

II. ANTENNA DESIGN

The geometric structure and size of the antenna are shown in Fig. 1. The antenna is designed with 50μm-thick LCP substrate with 18μm-thick copper laminated on both sides [11]. The dielectric constant and loss tangent of the LCP used in the design are 2.9 and 0.0025, respectively. The feeding line of the antenna is microstrip line with 50 ohm characteristic impedance. The overall dimension of UWB antenna is 27 mm × 21 mm × 0.05 mm, which is more than 50% smaller area compare with Ref. [8].

The simulation of antenna is realized by using Ansys HFSS. Both of bending and flat condition can be well predicted. Without ESCSRR, the VSWRs can be achieved lower than 2 in the whole UWB band (3.1 GHz-10.6 GHz) by optimizing the swallow tail radiation patch radian and the gradient of trapezoid ground. When adding an ESCSRR, a notched band in its corresponding frequency can be produced, because of the ESCSRR

Submitted On: March 5, 2017
Accepted On: March 23, 2017

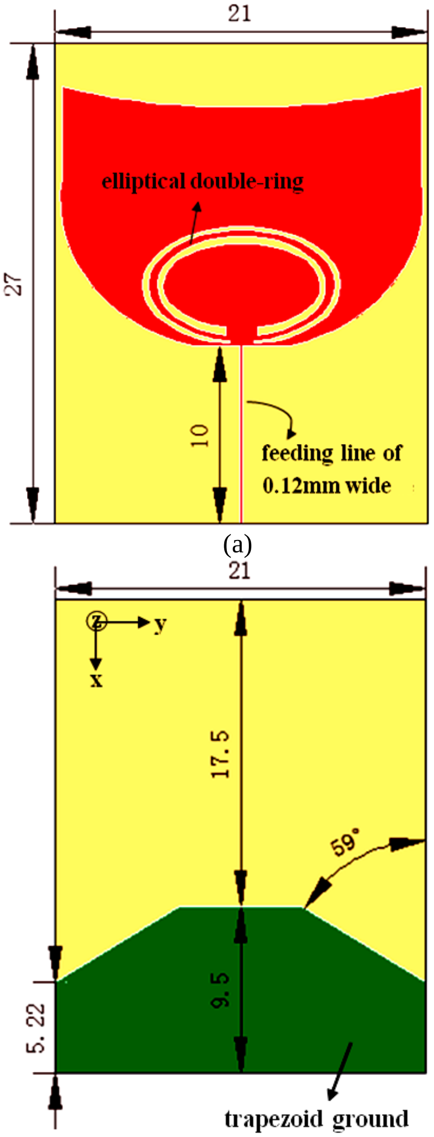
cutting and changing the radiation currents at its notched band frequency. The relationship between the dimensions of ESCSRR and the notched band frequency (f_{notch}) is shown in the following equations:

$$S_e = K_e\pi(R_2 - w_2) = \frac{\lambda_g}{2} = \frac{c}{2f_{\text{notch}}\sqrt{\varepsilon_{\text{eff}}}}, \tag{1}$$

$$K_e = 3(1 + k) - \sqrt{(3 + k)(1 + 3k)}, \tag{2}$$

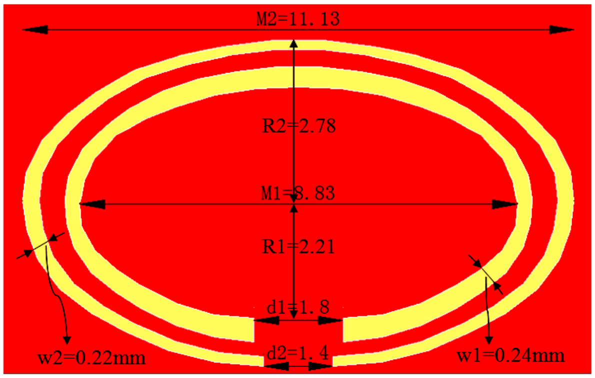
$$\varepsilon_{\text{eff}} = \frac{\varepsilon_r+1}{2} + \frac{\varepsilon_r-1}{2}\left(1 + \frac{12h}{\omega_f}\right)^{-0.5}, \tag{3}$$

where S_e means the circumference of the ESCSRR, which is approximately equal to half of the guided wavelength λ_g at the desired notch frequency. K_e is a factor related to k ($k=M_2/2R_2$). The values of w_2, M_2, and R_2 are shown in Fig. 1 (c). The effective dielectric constant is calculated through Equation (3), where h, ω_f, and ε_r are the substrate height, width of the microstrip feed and relative permittivity of substrate, respectively.



(a)

(b)

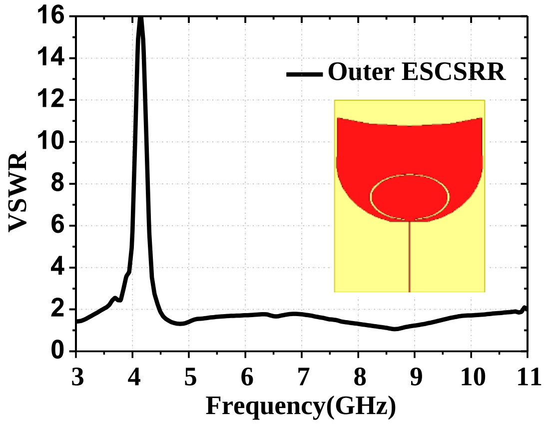


(c)

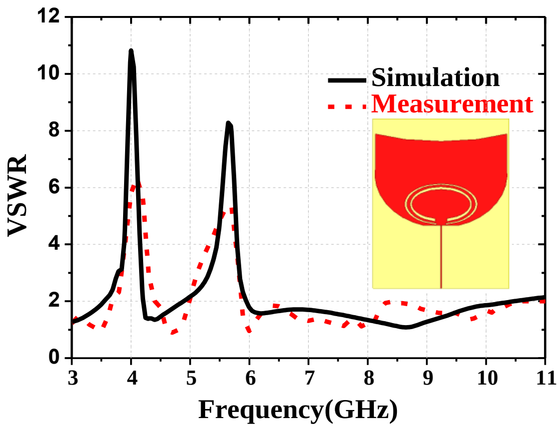
Fig. 1. Geometric structure and size of the antenna (unit: mm): (a) top view, (b) bottom view, and (c) two elliptical single complementary split-ring resonators.

By calculating the w_2, M_2, and R_2 of ESCSRR and optimizing the geometric parameters d_2, which is the key factor in effect the peak of VSWR as shown in Fig. 1 (c), the notched band between 3.7 GHz-4.2 GHz can be produced as shown in Fig. 2 (a). The second notched band between 5.15 GHz-5.825 GHz can be produced in the same way. When the antenna is operating at the center of upper notched band, the inside ESCSRR behaves as a separator, which has no effects on the other band-notches [6], as shown in Fig. 2 (b). At the same time, Fig. 2 (b) indicates the good agreement between simulated and measured results. The existing discrepancy may be attributed to fabrication tolerance.

The photography of fabricated antenna is shown in Fig. 3. The SMA connector is soldered onto the feeding microstrip. The temperature of the soldering iron tip should be kept below 200℃ in the soldering period to prevent melting the flexible substrate, meanwhile, the soldering time should be as short as possible. Figure 3 (b) is the photo of antenna bending over a polystyrene foam with a diameter of 25 mm in order to facilitate the measurement of the antenna under the bending condition.

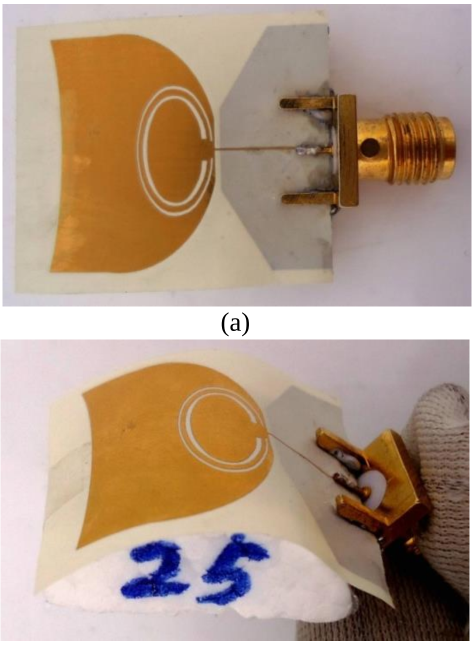


(a)



(b)

Fig. 2. Comparison of simulation and measurement results of VSWRs: (a) antenna with only outer ESCSRR, and (b) antenna with two ESCSRRs.



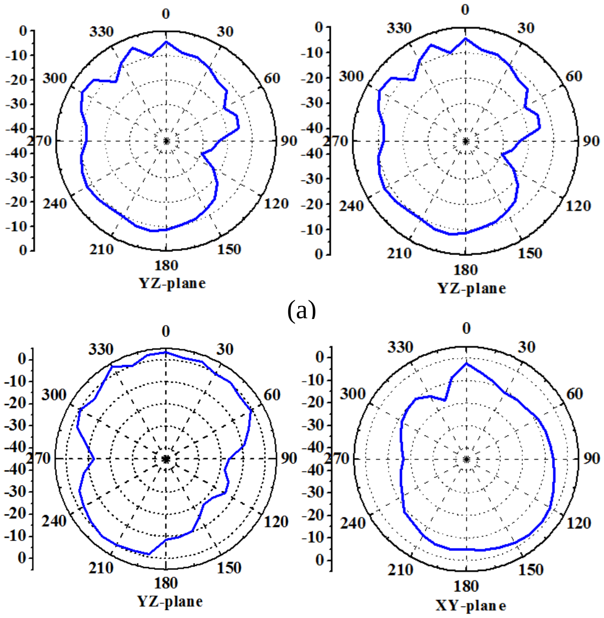
(a)

(b)

Fig. 3. Photograph of the fabricated antenna with SMA connector: (a) flat and (b) bent (diameter = 25 mm).

III. RESULTS AND DISCUSSION

The VSWRs and radiation patterns of the antenna in flat and bent circumstances are measured in the Satimo spherical near field measurement system. Good agreement between simulated and measured VSWR results has been shown in Fig. 2 (b). The measured radiation patterns of 4.5 GHz and 6 GHz are shown in Fig. 4. The measured maximum radiation gains of the flexible antenna at 4.5 GHz and 6 GHz are -4.4 dBi and 3.54 dBi, respectively. The antenna is omni-directional in XY-plane, while has directionality in YZ-plane. These results are consistent with the monopole antenna radiation pattern.



(a)

(b)

Fig. 4. Measured radiation patterns of the antenna at: (a) f = 4.5 GHz and (b) f = 6 GHz.

The measured VSWRs of the antenna under the bending and flatting conditions are shown in Fig. 5. Because of the dielectric constant of the polystyrene foam is only 1.06, which has little effects to antenna [12], it is reliable to measure the characteristics in this way. It is shown that there is little difference between the antenna VSWRs under the flat and bent conditions.

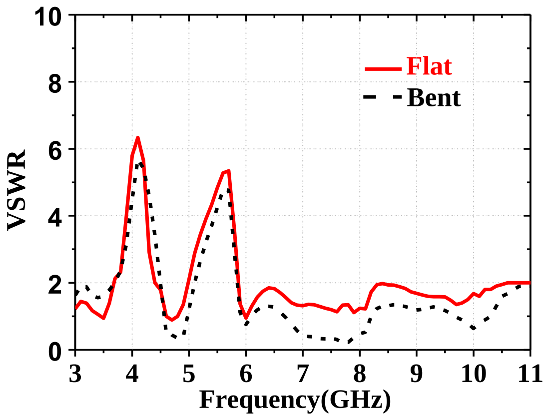


Fig. 5. Comparison of the measured VSWR results at flat and bent (diameter = 25 mm) circumstances.

Figure 6 shows the comparison of measured antenna radiation patterns at the flat and bent circumstances at 6 GHz. The monopole antenna shows good performance in bending condition with diameter of 25 mm and 33 mm, which proves that it is a good candidate for flexible systems [13].

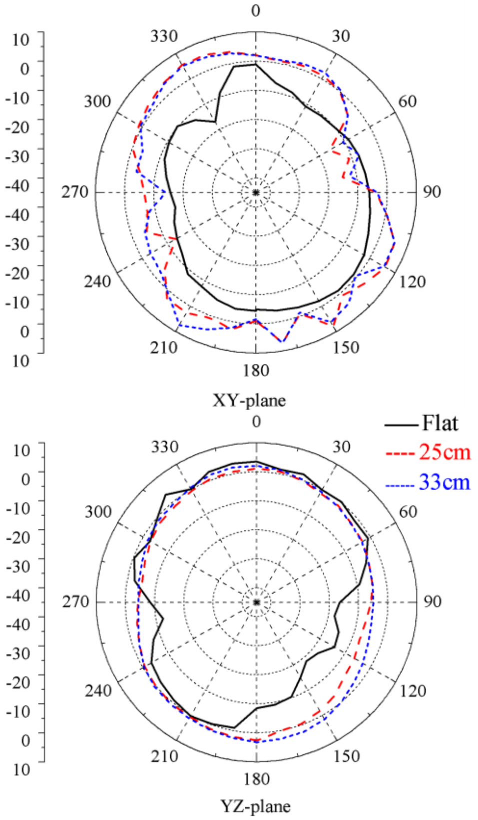


Fig. 6. Comparison of the radiation pattern at flat and bent (diameter = 25 mm and 33 mm) circumstances at 6 GHz.

IV. CONCLUSION

A two band-notched swallow tail antenna with mechanical flexibility is designed. The band-notched characteristics are realized by etching two elliptic single complementary split-ring resonators (ESCRRs) on radiation patch. The antenna is designed with more than 50% smaller size by using ultra-thin LCP substrate compared with FR-4 substrate. And good performance is shown both at flat and bent circumstances. This antenna can be used in flexible UWB system.

ACKNOWLEDGMENT

This work was supported by National Natural Science Foundation of China (Grant No. 61474020).

REFERENCES

[1] Federal Communications Commission, Washington, DC, USA, Federal Communications Commission revision of Part 15 of the commission's rules regarding ultra-wideband transmission system from 3.1 to10.6 GHz, 2002.

[2] D. Jiang, Y. Xu, and R. Xu, "A microstrip printed band-notched UWB antenna using modified CSRR structure, *The Applied Computational Electromagnetics Society (ACES) Journal*, vol. 28, no. 6, 2013.

[3] K. Chung, S. Hong, and J. Choi, "Ultrawide-band printed monopole antenna with band-notch filter," *IET Microw. Antennas Propag.*, vol. 1, no. 2, pp. 518-522, 2007.

[4] H. G. Schantz, G. Wolenec, and E. M. Myszka, "Frequency notched UWB antennas," *IEEE Conference on Ultra Wideband Systems and Technologies*, Virginia, USA, pp. 214-218, Nov. 2003.

[5] W. Xiao, T. Mei, Y. Lan, R. Xu, and Y. Xu, "Triple band-notched UWB monopole antenna on ultra-thin liquid crystal polymer based on ESCSRR," *Electronics Letters*, pp. 57-58, 2017.

[6] D. Jiang, Y. Xu, R. Xu, and W. Lin, "A compact ultra-wideband antenna with improved triple band-notched characteristics," *The Applied Computational Electromagnetics Society (ACES) Journal*, vol. 28, no. 2, 2013.

[7] Y. Cai, H. Yang, and L. Cai, "Wide band monopole antenna with three band-notched characteristics," *IEEE Antennas Wirel. Propag. Lett.*, vol. 13, pp. 607-610, 2014.

[8] D. Sarkar, K. V. Srivastava, and K. Saurav, "A compact microstrip-fed triple band-notched UWB monopole antenna," *IEEE Antennas Wirel. Propag. Lett.*, vol. 13, pp. 396-399, 2014.

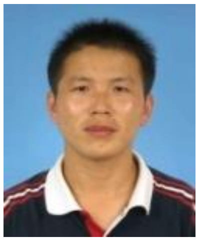
[9] C. Liu, Y. Guo, and S. Xiao, "Capacitively loaded circularly polarized implantable patch antenna for ISM band biomedical applications," *IEEE Trans. Antennas Propag.*, vol. 62, pp. 2407-2417, 2014.

[10] Y. Qiu, Y. H. Jung, S. Lee, T. Y. Shih, J. Lee, Y. Xu, R. Xu, W. Lin, N. Behdad, and Z. Ma, "Compact parylene-c-coated flexible antenna for WLAN and upper-band UWB applications," *Electron. Lett.*, vol. 50, no. 24, pp. 1782-1784, 2014.

[11] Y. Lan, Y. Xu, C. Wang, Z. Wen, Y. Qiu, T. Mei, Y. Wu, and R. Xu, "X-band flexible bandpass filter based on ultra-thin liquid crystal polymer substrate," *Electron. Lett.*, vol. 51, no. 4, pp. 345-347, 2015.

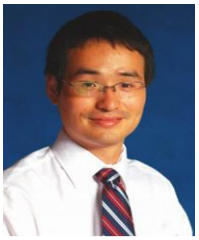
[12] C. R. Rowell and R. D. Murch, "A capacitively loaded PIFA for compact mobile telephone handsets," *IEEE Trans. Antennas Propag.*, vol. 45, pp. 837-842, 1997.

[13] Y. Xu, Y. Lan, Y. Qiu, and R. Xu, "Compact implantable rectenna with light-emitting diode for implantable wireless optogenetics, *The Applied Computational Electromagnetics Society (ACES) Journal*, vol. 31, no. 6, pp. 712-716, June 2016.



Chuicai Rong received the Bachelor of Science in Physics from Liaocheng University, Liaocheng, China, in 2000. Master of Science in Theoretical Physics from Qufu Normal University, Qufu, China, in 2006, and is currently working toward the Ph.D. degree at University of Electronic Science and Technology of China (UESTC). His research interests include microwave PA design and RF integrated circuit (RFIC) design.

Xiaowei received the B.S. degree from the University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2014. She is currently pursuing the M.S. degree at the University of Electronic Science and Technology of China (UESTC), Chengdu, China. Her research interests include microwave antennas and MMIC amplifiers.



Yuehang Xu was born in Zhejiang Province, China, in 1981. He received the B.S. and M.S. degrees in Electromagnetic Field and Microwave Techniques from the University of Electronic Science and Technology of China (UESTC), Chengdu, in 2004 and 2007, respectively, and the Ph.D. from the UESTC jointed with Columbia University in the City of New York in 2010.

He joined the Department of Electronic Engineering, UESTC, in Dec. 2010 and was promoted to Associate Professor in Aug. 2012, exceptionally. His current research interests are on modeling and characterization of micro/nano-scale electronic devices for RF applications, especially on GaN HEMTs and graphene electronics. He is an author or a co-author of more than 30 scientific papers in international journals and conference proceedings.

Fast Prediction of Coupling Coefficient between Monopole Antennas on Electrically Large Cylindrical Platforms Using a Linear Parametric Model

Huapeng Zhao [1], Si-Ping Gao [2], Wei-Jiang Zhao [2], Ying Zhang [1], Jun Hu [1], and Zhizhang Chen [3,1]

[1] School of Electronic Engineering
University of Electronic Science and Technology of China, Chengdu, 611731, P. R. China
huapengzhao@uestc.edu.cn

[2] Department of Electronics and Photonics
A*STAR Institute of High Performance Computing, 1 Fusionopolis Way, 138632, Singapore

[3] Department of Electrical and Computer Engineering
Dalhousie University, Halifax, Nova Scotia, Canada B3J 1Z1

Abstract **─** Antenna placement on electrically large cylindrical platforms requires a number of time consuming computer simulations that involve analysis of big electromagnetic data. In order to save simulation time, a linear parametric model is proposed to predict the coupling coefficient between monopole antennas on electrically large cylindrical platforms. The two parameters of the model are determined by the values of coupling coefficients at two different distances. The parametric model is then used to predict the coupling coefficient between monopole antennas at various distances. The proposed parametric model avoids repeated time consuming analysis of big electromagnetic data for predicting the coupling coefficients at various distances, and it thus significantly reduces simulation time. Given the desired level of isolation, the proposed parametric model can be used to find the required distance without a trial-and-error process, which further saves the time required by the analysis of big electromagnetic data.

Index Terms **─** Antenna placement, electromagnetic interference, parametric model.

I. INTRODUCTION

Electrically large cylindrical platforms represent an important class of platforms (e.g., commercial aircraft, rocket, etc.). A number of antennas may be installed on a platform for various onboard wireless systems. These antennas may result in interference between different electronic systems. In order to reduce the level of interference, a suitable position should be chosen for every antenna, which is called antenna placement.

An important step in antenna placement is to predict the level of coupling between antennas. Due to the large size of a platform, full wave simulation of antennas on a platform involves the analysis of big electromagnetic data, and it is very time consuming. Besides, the multi-scale feature of antennas on a platform results in ill-conditioned matrix, which further increases the simulation time. Moreover, if an antenna is provided by an external supplier, its detailed design information is usually unavailable to the system engineer, which prohibits the direct modeling of antennas on a platform. Last but not least, a trial-and-error process is usually performed in order to find suitable positions of antennas, which further increases the size of electromagnetic data.

In order to deal with the aforementioned challenges, various new methods were proposed. Recent advances include hybrid methods [1, 2], domain decomposition methods [3-5] and equivalent model methods [6, 7]. These methods have significantly reduced the CPU time and memory cost required by the simulation of antennas on large platforms. However, the trial-and-error process is still needed for antenna placement. Hence, it is highly desirable to develop new models that can avoid the trial-and-error process.

This paper presents a parametric model for the fast prediction of coupling coefficient between monopole antennas on electrically large cylindrical platforms. Ray analysis is performed to derive an analytical formula of the radiated electromagnetic fields by a transmitting monopole antenna. The Poynting vector is then obtained, and a linear parametric model of the coupling coefficient is proposed based on the expression of the Poynting vector. There are two parameters in the proposed parametric model, and they can be determined by performing two electromagnetic simulations. Once the two parameters are found, the parametric model can be

Submitted On: December 29, 2016
Accepted On: April 8, 2017

used to predict the coupling coefficients at various distances between monopole antennas.

Once determined, the proposed parametric model no longer requires electromagnetic simulations to predict the coupling coefficients at various distances. It can also be used to calculate the required distance to achieve the desired isolation. Therefore, the trial-and-error process is avoided and the time required by antenna placement can be significantly reduced.

The rest of this paper is organized as follows. Section II presents derivation of the proposed parametric model. Section III shows numerical results and Section IV concludes this paper.

II. THE PROPOSED PARAMETRIC MODEL

Consider the electrically large cylindrical platform shown in Fig. 1. Assume a transmitting monopole antenna and a receiving monopole antenna are located at the source point Q' and the observation point Q, respectively. According to [8], the electromagnetic filed radiated by a uniform short monopole current I at point Q' can be expressed by Equation (1) at the bottom of this page. In Equation (1), k is the wavenumber, η_0 is the free space wave impedance, and U and V denote the Fock functions whose definition can be found in the Appendix of [9]. As show in Fig. 1, $\hat{n}$ is the unit vector normal to the cylindrical surface at point Q, $\hat{t}$ is the unit vector tangential to the geodesic between Q and Q', and $\hat{b} = \hat{t} \times \hat{n}$. The other symbols in Equation (1) are defined as follows $\xi = t(k/2)^{1/3} a^{-2/3} (\sin\delta)^{4/3}$, $T_0 = \cot\delta$, and $G_0(kt) = e^{-jkt}/t$, where t is the length of the geodesic between Q and Q', a is the radius of the cylindrical surface, and δ is the angle between the axial direction and the geodesic at Q. For antenna placement, the worst case of electromagnetic interference is the most interesting. For two monopole antennas on a cylindrical surface, the worst case occurs when Q and Q' have the same φ coordinate. In this case, $\delta = 0$ and $\xi = 0$. We further assume that $(kt)^{-2} << 1$, Equation (1) is then simplified and the radiated electromagnetic field is rewritten as:

$$\vec{E}(Q|Q')\big|_{\delta=0} = \eta_0 I F(t) G_0(kt)\hat{n}, \tag{2.1}$$

$$\vec{H}(Q|Q')\big|_{\delta=0} = I F(t) G_0(kt)\hat{b}, \tag{2.2}$$

where

$$F(t) = \frac{-jk}{2\pi}\left\{1 - j\left(\frac{\sqrt{\pi}e^{j\pi/4}}{4\sqrt{2}}\frac{(kt)^{\frac{1}{2}}}{ka} + \frac{1}{kt}\right)\right\}. \tag{3}$$

The Poynting vector $\vec{P}$ is then calculated as follows:

$$\vec{P}(Q|Q') = \frac{k^2\eta_0 I^2}{8\pi^2 t^2}\left[1 + X(t)\right]\hat{t}, \tag{4}$$

where

$$X(t) = \frac{\pi}{32(ka)^2}kt + \frac{\sqrt{\pi}}{4ka}\sqrt{kt} + \frac{\sqrt{\pi}}{4ka}\frac{1}{\sqrt{kt}}. \tag{5}$$

It can be shown that the first and third terms of $X(t)$ are negligible as long as the following two conditions are satisfied:

$$\frac{312.5}{(ka)^2} < \frac{t}{\lambda} < \frac{0.16(ka)^2}{\pi^2}, \tag{6}$$

where λ is the wavelength. For electrically large cylindrical platforms, it can be assumed that $ka > 20\pi$. When deriving Equation (2), it is assumed that $(kt)^{-2} << 1$. Combining that assumption with the condition in (6), the following range of t is derived:

$$1.6 < \frac{t}{\lambda} < 64. \tag{7}$$

Equation (7) defines the range of t that allows us to neglect the first and third terms in (5). For electrically large cylindrical platforms, the wavelength is small and the lower bound of t is usually satisfied. Beyond the upper bound of t, the electromagnetic interference is negligible. Therefore, the condition in (7) is assumed to be satisfied in this work. The Poynting vector $\vec{P}$ is then simplified as:

$$\vec{P}(Q|Q') = \frac{k^2\eta_0 I^2}{8\pi^2}\left[t^{-2} + \frac{\sqrt{k\pi}}{4ka}t^{-1.5}\right]\hat{t}. \tag{8}$$

Equation (8) indicates that the decaying rate of radiated power is between -2 and -1.5. Therefore, the power received by a receiving monopole antenna at Q

$$\vec{E}(Q|Q') = \frac{-jk}{2\pi} I\left[\eta_0\hat{n}\left\{\left(1-\frac{j}{kt}\right)V(\xi) + \left(\frac{j}{kt}\right)^2 U(\xi) + T_0^2\frac{j}{kt}\left[U(\xi)-V(\xi)\right]\right\}\right]G_0(kt), \tag{1.1}$$

$$\vec{H}(Q|Q') = \frac{-jk}{2\pi} I\left[\hat{b}\left\{\left(1-\frac{j}{kt}\right)V(\xi) + T_0^2\frac{j}{kt}\left[U(\xi)-V(\xi)\right]\right\} + \hat{t}\left\{T_0\frac{j}{kt}\left[U(\xi)-V(\xi)\right]\right\}\right]G_0(kt), \tag{1.2}$$

is proportional to t^{β}, where $-2<\beta<-1.5$, and the magnitude of coupling coefficient between transmitting and receiving antennas can be described by the following parametric model:

$$\left|S_{21}\right| = Ct^{\beta}. \tag{9}$$

By expressing the coupling coefficient in dB scale, we have:

$$\left|S_{21}\right|_{dB} = \beta 10\log t + \alpha, \tag{10}$$

where $\alpha = 10\log C$. In Equation (10), α and β are two unknown parameters. With two groups of data $\left(\log t_1, \left|S_{21}\right|_{dB,1}\right)$ and $\left(\log t_2, \left|S_{21}\right|_{dB,2}\right)$, α and β can be calculated as follows:

$$\begin{cases} \alpha = \dfrac{\log t_1 \left|S_{21}\right|_{dB,2} - \log t_2 \left|S_{21}\right|_{dB,1}}{\log t_1 - \log\log t_2} \\ \beta = \dfrac{\left|S_{21}\right|_{dB,1} - \left|S_{21}\right|_{dB,2}}{10\left(\log t_1 - \log\log t_2\right)} \end{cases}. \tag{11}$$

Therefore, the two parameters in (10) can be determined by two groups of simulation or measurement data. Once α and β are solved, the coupling coefficient can be computed from (8) without running the electromagnetic simulation. The simulation cost is then saved.

Moreover, from Equation (10), the required distance between transmitting and receiving antennas can be calculated directly from the desired isolation by the following formula:

$$t_{\min} = 10^{-\frac{IS_{dB}+\alpha}{10\beta}}, \tag{12}$$

where IS_{dB} is the desired isolation and $t_{\min}$ is the minimum distance between transmitting and receiving antennas to satisfy the isolation requirement. Equation (12) avoids the tedious trial-and-error process in the conventional antenna placement method, and provides a fast way to determine the distance between transmitting and receiving antennas.

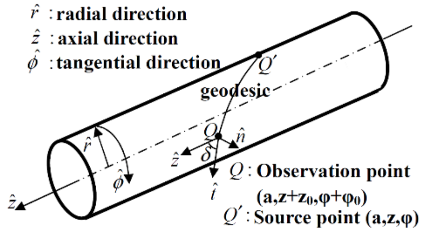


Fig. 1. Placement of monopole antennas on an electrically large cylindrical structure.

III. SIMULATION RESULTS

This section presents simulation results to show the advantages of the proposed method. Three examples are considered. The first example is a monopole antenna on a large planar ground plane, for which $ka \to \infty$. The second example is two monopole antennas on a half-cylinder. The third example is two monopole antennas on a commercial aircraft.

A. A monopole antenna on a large planar platform

In order to verify the theoretical analysis in Section II, a planar platform is first considered. In this case, $ka \to \infty$ and the second term in Equation (8) becomes zero. Consider a quarter-wavelength monopole antenna and assume the current distribution along the monopole antenna is $I(z) = I_0 \cos(kz)$. By taking integral of the current along the monopole antenna, it is found that the Poynting vector of the monopole antenna on a large planar platform is:

$$\vec{P}_m\left(Q|Q'\right) = \frac{15 I_0^2 \lambda^2}{\pi} t^{-2}\hat{t}. \tag{13}$$

Equation (13) is used to calculate the magnitude of Poynting vector of a quarter-wavelength monopole antenna on a rectangular ground plane. As shown in the inset of Fig. 2, the size of the ground plane is $15\lambda \times 20\lambda$. The monopole is 4λ away from the left edge of the ground plane. The magnitude of Poynting vector is calculated at a series of points on the right of the monopole antenna, and the results are shown in Fig. 2. It is seen that the proposed model agrees well with FEKO simulation. Note that discrepancies occur when $t > 11\lambda$, because the observation point is no longer on the platform.

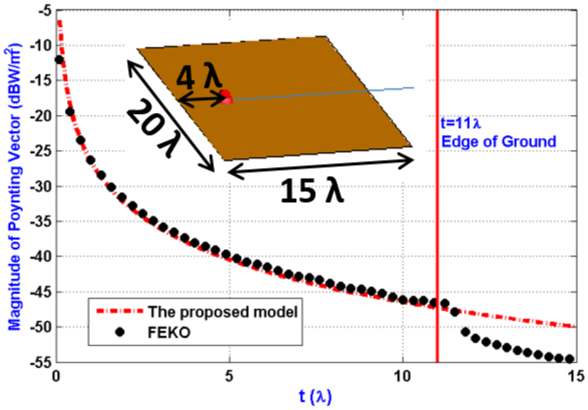


Fig. 2. The magnitude of Poynting vector of the quarter-wavelength monopole antenna on a large planar platform.

B. Two monopole antennas on a half-cylinder

The second example is two monopole antennas on a half-cylinder. The half-cylinder is a one tenth scaled model of the curved surface of a commercial aircraft fuselage. As shown in Fig. 3, the radius of the half-cylinder is 18.8 cm, and its length is 150 cm. Eleven holes of equidistance are drilled on the half-cylinder to mount monopole antennas. Ten measurements of coupling coefficient between transmitting and receiving antennas are performed. In these measurements, the transmitter is fixed at the first hole H1, and the receiver moves from the second hole H2 to the eleventh hole H11. To make the surface smooth, unused holes are covered by aluminum foils in every measurement.

Using the measurement data at H2 and H11, α and β are calculated by Equation (11). The coupling coefficient for different values of t is then calculated by Equation (10). The calculation results by the proposed model are shown in Fig. 4. Also shown in Fig. 4 are the measurement results. It is seen that the proposed model agrees well with measurement. In this example, the proposed model only requires two measurements to predict the coupling coefficients for different distances between transmitting and receiving antennas, which significantly saves measurement cost. If the desired isolation between transmitting and receiving antennas is given, the required distance can be calculated directly by Equation (12).

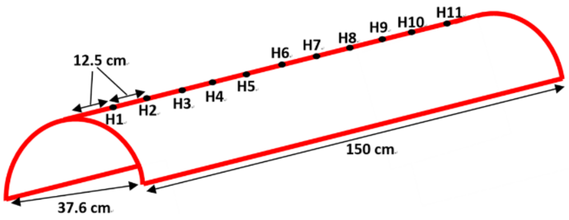


Fig. 3. Geometry of a half-cylinder with eleven holes.

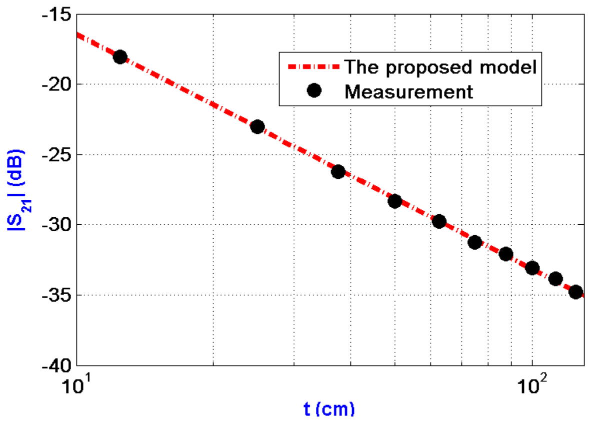


Fig. 4. Coupling coefficient between two monopole antennas on a half-cylinder.

C. Two monopole antennas on a commercial aircraft

In the third example, two monopole antennas mounted on a commercial aircraft are considered. Figure 5 illustrates the geometry of the aircraft. It is assumed that two monopole antennas are mounted on the fuselage of the aircraft. FEKO simulation is performed to calculate the coupling coefficient between the two antennas. Two simulated values of the coupling coefficient are used to compute the two parameters the proposed model. Equation (10) is then used to predict the coupling coefficient versus the distance between the two antennas, which is plotted in Fig. 6. Also shown in Fig. 6 are the FEKO simulation results. It is observed that the proposed model agrees well with simulation by FEKO.

Regarding computational efficiency, the proposed model only requires two electromagnetic simulations, while FEKO requires repeated simulation for every distance of interest. Furthermore, to find the minimum distance that satisfies the isolation requirement, a trial-and-error process is needed by FEKO simulation. However, the proposed model can find the minimum distance directly by using Equation (12), which will significantly save simulation time.

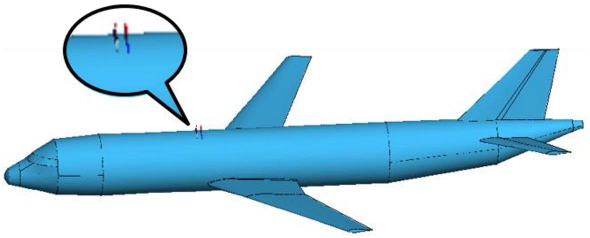


Fig. 5. Two monopole antennas on a commercial aircraft.

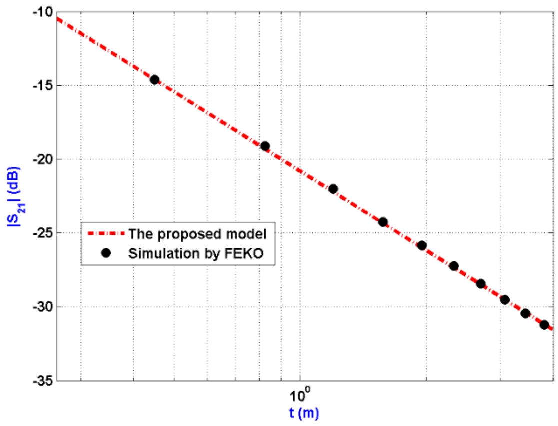


Fig. 6. Coupling coefficient between two monopole antennas on a commercial aircraft.

IV. CONCLUSION

This paper has presented a parametric model for fast prediction of coupling coefficient between monopole antennas on electrically large cylindrical platforms. Theoretical analysis has been performed to find the Poynting vector of a monopole antenna on electrically large cylindrical platforms. It has been found that the decaying rate of the magnitude of Poynting vector against distance is between -2 and -1.5, based on which a parametric model with two parameters has been proposed. The proposed model has been validated by three examples. It has been shown that prediction by the proposed model agrees well with both measurement and simulation results. The proposed model allows us to find the decaying rate of the coupling coefficient against distance from two group of simulation data, which significantly saves the time required by the analysis of big electromagnetic data.

In the present work, only monopole antennas are considered. It is interesting to extend the proposed model to other antennas, which is under investigation and will be reported in the future.

ACKNOWLEDGMENT

The work of H. Zhao, J. Hu, and Z. Chen are partially supported by the National Science Foundation of China under Grants No. 61471107, No. 61425010, No. 61271033, No. 61671121, the UESTC Research Start-up funding under Grant No. ZYGX2016KYQD106, the Fundamental Research Funds for the Central Universities under Grant No. ZYGX2016J032, and the China Postdoctoral Science Foundation under Grant No. 2016M602673.

REFERENCES

[1] J. Ma, S. X. Gong, X. Wang, et al., "Efficient IE-FFT and PO hybrid analysis of antennas around electrically large platforms," *IEEE Antennas and Wireless Propagation Letters*, vol. 10, pp. 611-614, 2011.

[2] Z. L. Liu and C. F. Wang, "Efficient iterative method of moments—physical optics hybrid technique for electrically large objects," *IEEE Trans. Antennas Propag.*, vol. 60, no. 7, pp. 3520-3525, July 2012.

[3] X. Wang, Z. Peng, K.-H. Lim, and J.-F. Lee, "Multisolver domain decomposition method for modeling EMC effects of multiple antennas on a large air platform," *IEEE Trans. Electromagn. Compat.*, vol. 54, no. 2, pp. 375-388, Apr. 2012.

[4] H. Wang, V. Khilkevich, Y.-J. Zhang, and J. Fan, "Estimating radio frequency interference to an antenna due to near-field coupling using decomposition method based on reciprocity," *IEEE Trans. Electromagn. Compat.*, vol. 55, no. 6, pp. 1125-1131, Dec. 2013.

[5] Y. Li, X. Zhao, and H. Zhang, "Out-of-core solver based DDM for solving large airborne array," *Applied Computational Electromagnetics Society Journal*, vol. 31, no. 5, pp. 509-515, May 2016.

[6] S.-P. Gao, B. Wang, H. Zhao, W. Zhao, and C. Png, "Installed radiation pattern of patch antennas: Prediction based on a novel equivalent model," *IEEE Antennas Propag. Mag.*, vol. 57, no. 3, pp. 81-94, June 2015.

[7] S.-P. Gao, H. Zhao, H.-W. Deng, B. Wang, and W. Zhao, "Estimating interference to airborne patch antenna with limited information," *IEEE Trans. Electromagn. Compat.*, vol. 58, no. 2, pp. 631-634, Apr. 2016.

[8] P. H. Pathak and N. Wang, "Ray analysis of mutual coupling between antennas on a convex surface," *IEEE Trans. Antennas Propag.*, vol. AP-29, no. 6, pp. 911-922, 1981.

[9] S. W. Lee and S. Naini, "Approximate asymptotic solution of surface field due to a magnetic dipole on a cylinder," *IEEE Trans. Antennas Propag.*, vol. AP-26, no. 4, pp. 593-597, 1978.

Huapeng Zhao received the Ph.D. degree from the Nanyang Technological University, Singapore, in June 2012. He was a Scientist with the A*STAR Institute of High Performance Computing from August 2011 to December 2015. Since December 2015, he has been an Associate Professor with the School of Electronic Engineering at the University of Electronic Science and Technology of China, Chengdu, P. R. China.

Si-Ping Gao received the B.Eng., the M.Eng. and the Ph.D. degrees all from Nanjing University of Aeronautics and Astronautics, P. R. China. He is currently a Scientist with the Institute of High Performance Computing, A*STAR, Singapore. His research interests include applied electromagnetics, discontinuous Galerkin modeling of high-speed electronics, electromagnetic compatibility of antennas on large platforms, and lightning protection for aerospace applications.

Gao has published over 30 refereed papers. He is a recipient of APEMC 2016 Best Symposium Paper Award and SPI 2017 Young Investigator Training Program. He served as a Session Chair of NEMO 2016, EPTC 2016, APEMC 2017, Finance Chair of 2018 Joint IEEE EMC & APEMC, and has been the Secretary of IEEE EMC Singapore Chapter since 2016.

Wei-Jiang Zhao received the B.Sc. degree in Mathematics from Nankai University, Tianjin, China, and the M.Eng. and Ph.D. degrees in Electrical Engineering from Xidian University, Xi'an, China.

In 1992, he joined Xidian University, where he was promoted to Associate Professor in 1999. From 2001 to 2010, he had been with the National University of Singapore, Singapore, as a Research Scientist. In 2010, he joined the Institute of High Performance Computing (IHPC), A*STAR, Singapore, where he is currently a Senior Scientist. His research interests include computational electromagnetics, antenna analysis and design, electromagnetic/acoustic wave propagation and scattering, EMC/EMI analysis, and numerical optimization.

Ying Zhang received the Ph.D. degree from the Nanyang Technological University (NTU), Singapore, in June 2011. She was a Postdoctoral Researcher with the School of Electronic Engineering at NTU from May 2010 to May 2011. Since July 2011, she has been an Associate Professor with the School of Electronic Engineering at the University of Electronic Science and Technology of China, Chengdu, P. R. China.

Jun Hu received the B.S., M.S., and Ph.D. degrees in Electromagnetic Field and Microwave Technique from the University of Electronic Science and Technology of China (UESTC), Chengdu, P. R. China, in 1995, 1998, and 2000, respectively.

Currently, he is a Full Professor with the School of Electronic Engineering, UESTC. During 2001, he was with the Center of Wireless Communication in the City University of Hong Kong, Kowloon, Hong Kong, as a Research Assistant. From March to August of 2010, he was a Visiting Scholar in the ESL of the Electrical and Computer Engineering Department, Ohio State University. He was a Visiting Professor of the City University of Hong Kong from February to March in 2011. His current research interests include integral equation methods in computational electromagnetics, electromagnetic scattering and radiation.

Hu is a Member of the Applied Computational Electromagnetics Society. He served as the Chairman for the IEEE Chengdu Joint AP/EMC Chapter during 2014-2016. He was a recipient of the 2004 Best Young Scholar Paper Award from the Chinese Radio Propagation Society and was selected into 'Who's Who in the World' (Marquis, 31 Edition).

Zhizhang Chen received the M.S. degree in Radio Engineering from Southeast University, Nanjing, China, and the Ph.D. degree in Electrical Engineering from the University of Ottawa, Ottawa, ON, Canada. He was an NSERC Postdoctoral Fellow with McGill University, Montreal, QC, Canada. He is currently a Professor at Dalhousie University, Halifax, NS, Canada, where he has served as the Head of the Department of Electrical and Computer Engineering. He has been an Adjunct or Visiting Professor at the University of Nottingham, U.K., Ecole Nationale Superieure des Telecommunications de Bretagne of France, Shanghai Jiaotong University, Beihang University, Fuzhou University, and the University of Electronic Science and Technology of China. He has authored and co-authored more than 300 journal and conference papers in computational electromagnetics, RF/microwave electronics, antennas, and wireless technologies. His current research interests are in unconditionally stable electromagnetic modeling techniques, ultrawideband wireless communication systems and wireless power transfer.

Chen is a Fellow of the IEEE, the Canadian Academy of Engineering and the Engineering Institute of Canada. He was one of the originators of the unconditionally stable methods that have been highly cited and used. He received the 2005 Nova Scotia Engineering Award, a 2006 Dalhousie Graduate Teaching Award, the 2007 & 2015 Dalhousie Faculty of Engineering Research Award, and the 2013 IEEE Canada Fessenden Medal.

A Compact Triple Band Antenna for Bluetooth, WLAN and WiMAX Applications

Kai Yu [1], Yingsong Li [1,*], and Wenhua Yu [2]

[1] College of Information and Communications Engineering
Harbin Engineering University, Harbin 150001, China
*liyingsong@ieee.org

[2] 2COMU, Inc.,
4031 University Drive, Suite 100, Fairfax, VA 22030, USA

Abstract ─ In this paper, we propose a compact triple band microstrip antenna, whose performance is discussed and investigated by using full-wave simulation and measurements. To achieve expected triple-band characteristics, three nested loop radiation elements are presented and a rectangular stub connected to the feed line is also introduced. Our proposed antenna has a size of 28 mm×41 mm and a thickness of 1.5 mm, which is fed by a coplanar waveguide (CPW). The simulated and measured results show that the proposed antenna provides a tri-band characteristic that covers 2.1-2.8 GHz, 3.3-4.0 GHz and 5.5-5.8 GHz. The center frequencies of the designed triple bands can be controlled by adjusting the dimensions of the nested loop elements. As a result, the designed tri-band antenna has a good impedance matching characteristic and omnidirectional radiation patterns, which make the proposed antenna could be a favorable candidate for Bluetooth, WiMAX and WLAN applications.

Index Terms ─ Bluetooth; CPW feeding; nested loop radiation elements; tri-band antenna; WiMAX and WLAN applications

I. INTRODUCTION

In recent years, wireless communication technology has been extensively studied and used for many portable devices like smartphone, laptop and other personal terminals. As a wireless equipment, an antenna plays an important role in transmitting and receiving electromagnetic wave signals. Furthermore, microstrip antennas have gotten more and more attentions in recent decades due to the rapid development of wireless communications and their advantages of low cost, easy of fabrication and conformability [1-6]. Wireless local area network (WLAN) and worldwide interoperability for microwave access (WiMAX) based on IEEE 802.11 and 802.16 standards are two most widely studied wireless technologies. Also, WLAN and WiMAX have been used in our daily life, and operate at 2.4-2.484 GHz, 5.15-5.35 GHz, 5.725-5.825 GHz and 2.5-2.7 GHz, 3.3-3.69 GHz and 5.25-5.85 GHz. After that, many research articles have been published to make an effort to design WLAN and WiMAX antennas [7-10]. Meanwhile, Bluetooth, a short distance transmission technology, has also been widely integrated into a smart phone, a smart watch, a wireless headset and an on-ear headphone [11]. The Bluetooth is also operating in the band from 2.4 GHz to 2.483 GHz. Thus, multi-band design is needed for developing a wireless device by sharing only one antenna, which has an advantage of high efficiency and low cost [12]. Then, many methods have been presented to design an antenna with good multi-band characteristics, including etching slots on the patch or ground [13-14], loading the shorted pins and walls [15], adding stubs to the ground and patch [16], using metamaterials [17] and EBG [18], and so on. However, designing an antenna to cover Bluetooth, WiMAX, and WLAN bands is still a challenging work today. Most of the mentioned antennas cannot be integrated into portable devices because of their large size or complex structure.

We present, in this paper, a compact CPW feed triple band microstrip antenna for Bluetooth, WLAN and WiMAX applications. By using three nested loops and a rectangular stub connected to the feed line, three resonance bands can be achieved to operate at 2.1-2.8 GHz, 3.3-4.0 GHz and 5.5-5.8 GHz, which can be used for Bluetooth, WiMAX at 2.5/3.5/5.5 GHz and WLAN at 2.4/5.8 GHz. By adjusting the dimensions of the antenna, the center frequencies of these three operating bands are tunable, which makes it flexible for practical fabrication. The full-wave simulation and measurement results have verified that our proposed antenna has not only a triple-band achievement but also a good impedance matching characteristic and omnidirectional radiation patterns at three operating bands, which make it suitable for Bluetooth, WLAN and WiMAX communication applications. The numerical

Submitted On: December 14, 2016
Accepted On: March 15, 2017

full-wave simulation results and optimized dimensions of the proposed tri-band microwave antenna are obtained by using the commercial software package high frequency structure simulator (HFSS).

II. ANTENNA CONFIGURATION

The configuration of the proposed triple-band antenna is depicted in Fig. 1. The proposed antenna is fabricated on a substrate with a relative permittivity of 3.5, a loss tangent of 0.002 and a thickness of 1.5 mm. The antenna has a total size of 28×41 mm^2. Our proposed antenna is comprised of a CPW-fed structure, which includes a CPW ground plane and a 50-Ohm feed signal line whose width is W0=4 mm, three nested loop radiation elements and a rectangular stub connecting to the feeding feed line. The three operating bands are generated by the presented three nested loops so that the center frequencies of three operating bands can be adjusted by selecting the dimension of the loops. The distance M1, M2 and M3 between the centers of the three loops and the end of the feeding point are 22 mm, 25.5 mm and 28.5 mm, respectively. The end of the feeding point is the connecting point of the loops and the feeding line. The rectangular stub is used to improve the efficiency. The structure of the antenna is very simple in which the CPW fed structure has the advantage of easy to integration with the radio-frequency front-end. The proposed antenna is investigated and optimized by using the Ansoft High Frequency Structure Simulator (HFSS) Version 13, and the optimized parameters of the antenna are listed in Table 1.

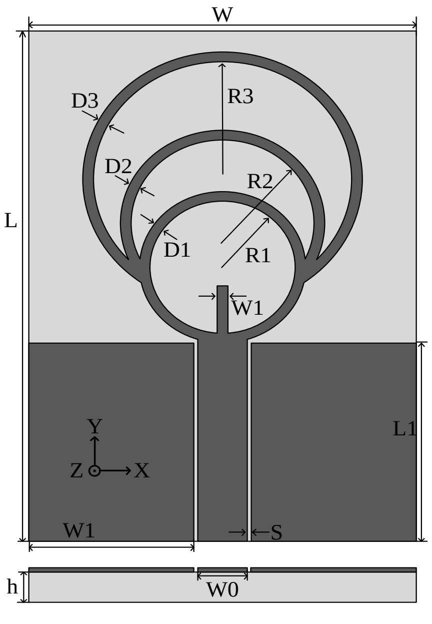


Fig. 1. Geometry structure of the proposed antenna.

Table 1: Optimized parameters of the antenna

L	41 mm	R1	5.5 mm	D3	1 mm
L1	15.5mm	R2	7 mm	S	0.4 mm
W	28 mm	R3	10 mm	M1	22 mm
W0	4 mm	D1	1 mm	M2	25.5 mm
W1	1 mm	D2	1 mm	M3	28.5 mm

III. PERFORMANCE ANALYSIS

For the design of the antenna, three nested loops are used to generate the expected resonance frequencies. Each nested loop is initially set to be about quarter wave length corresponding to the resonance frequency. Then, the antenna is modeled and optimized by means of the HFSS to obtain the optimal dimensions. To sufficiently investigate the performance of the proposed triple-band antenna, the key parameters are selected to investigate the antenna performance. Here, R2, R3 and D1 are chosen to discuss their effects on the antenna impedance. The effects of R2 on the reflection coefficient of the proposed antenna are shown in Fig. 2. It is observed from Fig. 2 that, R2 has an important effect on the highest operating band while the center frequency of the highest band moves to low frequency with an increment of the R2 from 7 mm to 8 mm. And the bandwidth of the middle band is slightly narrowed while the center frequency and bandwidth of the lowest band are almost invariable.

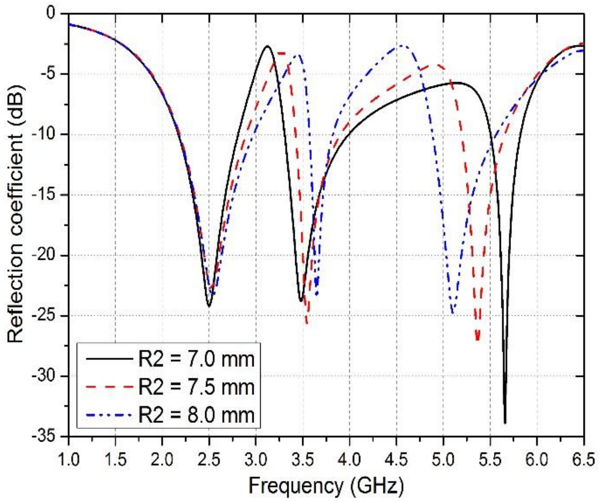


Fig. 2. Effects of the parameter R2 on the reflection coefficient.

Figure 3 depicts the parameter effects on the reflection coefficient of the proposed antenna with variation of R3. It can be found from Fig. 3 that the resonance frequency at 3.5 GHz for WiMAX moves towards the high frequency with the reduction of R3 from 10 mm to 9 mm. At the same time, the bandwidth of the lowest band is improved, while the bandwidth and the center frequency keep constant.

In Fig. 4, the effect of parameter D1 is presented. We can find from the figure that the center frequency of the highest band moves to high frequency with the increment of D1, while the middle band and the lowest band are almost invariable. Thus, we can conclude that the center frequencies of the three operating bands can be controlled by adjusting the dimensions of the loops.

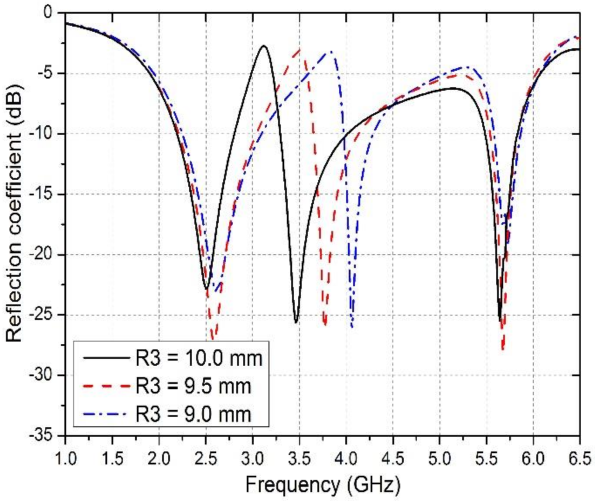


Fig. 3. Effects of the parameter R3 on the reflection coefficient.

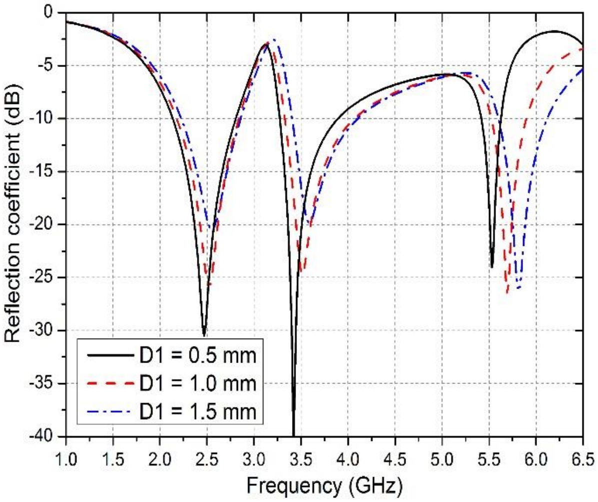


Fig. 4. Effects of the parameter D1 on the reflection coefficient.

From discussions above, we can see the radius and the width of the designed nested loop radiation element have important effects on the impedance bandwidth. The effect of the rectangular stub on the reflection coefficient is depicted in the Fig. 5 from which it can be found that the proposed antenna without the stub can operating at 5.5 GHz to 6.3 GHz. By using this stub, we remove the unwanted bandwidth that is out of the WLAN band. Thus, we can conclude that the proposed antenna with stub can effectively suppress the out-of-band radiation, and help the proposed antenna to improve the efficiency.

The impedance characteristic of the proposed triple bands is illustrated in Fig. 6, in which the black line is the real and the red line is the imaginary. It is found that the real part of the antenna impedance at the three operating bands is around 50 Ohm and the imaginary part is about zero, which means that the proposed triple-band antenna has a good impedance matching characteristic at the three operation bands.

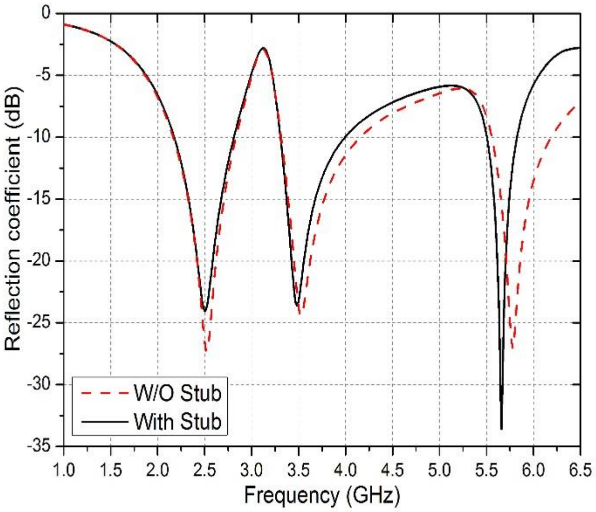


Fig. 5. Reflection coefficient of proposed antenna with and without stub.

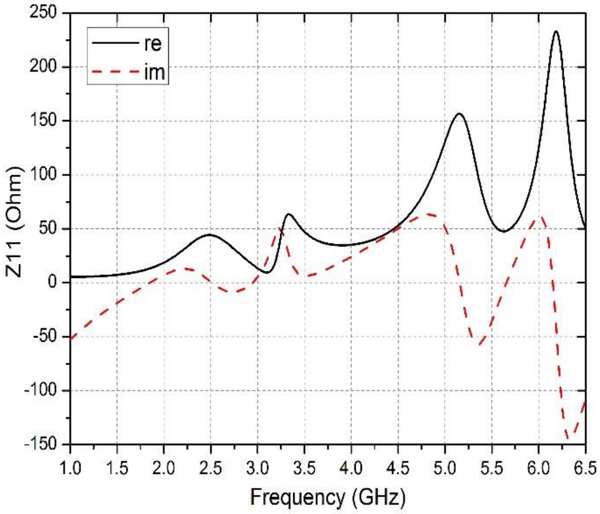


Fig. 6. Impedance characteristic of proposed tri-band antenna.

In order to verify the performance of the proposed CPW-fed triple-band antenna, the optimized antenna is fabricated and antenna performance is measured. The phototype of the fabricated antenna is shown in Fig. 7. The measured and simulated reflection coefficients of the designed antenna are shown in Fig. 8, where the measured result is obtained by using Agilent N5224A vector network analyzer. We can observe that the measurement result in a good agreement with the simulated one, which can help verify the effectiveness of the simulation. As is noted, there are some discrepancies between the measurement and the simulation, which may be caused by the fabrication tolerances and experimental conditions. From Fig. 8, we can find that our proposed antenna has three operation bands at 2.5 GHz, 3.5 GHz and 5.6 GHz in terms of the reflection coefficient less than -10 dB. The bandwidth of the lowest band, middle band and highest band are 0.7 GHz, 0.7 GHz and 0.3 GHz, respectively. The three resonance frequencies can be used for Bluetooth, WLAN 2.4GHz, WiMAX 3.5 GHz and WLAN 5.8 GHz.

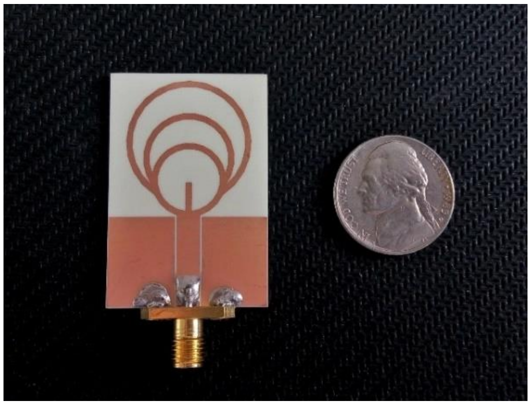


Fig. 7. Fabricated tri-band antenna.

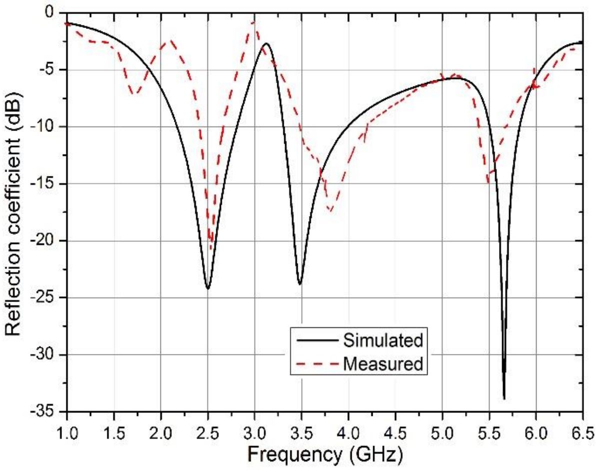
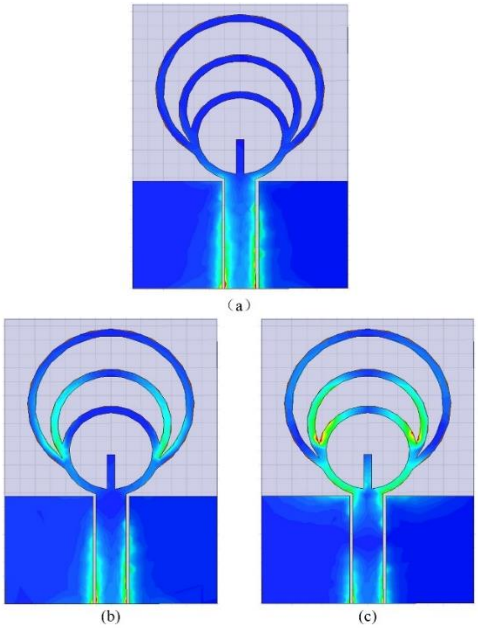


Fig. 8. The measured and simulated reflection coefficients of antenna.

Current distribution of the proposed tri-band antenna in Fig. 9 at 2.5 GHz, 3.5 GHz and 5.65 GHz are also studied to get a further understanding of the principle. From the figure, we can see that the current distribution at 2.5 GHz mainly focuses on the feed line. The current distributions concentrate on the both feed line and second loop at 3.5 GHz. Similarly, a large current appears on the both first and second loop at 5.65 GHz.

Figure 10 shows the gain and the efficiency of proposed antenna. It is found that the gain of the antenna are 1.8 dBi, 1.8dBi and 5dBi at 2.5 GHz, 3.5 GHz and 5.6 GHz, respectively, which are higher than zero in the operating bands. The efficiency is higher than 90% in most of the operating bands.



(a)

(b)

(c)

Fig. 9. Current distribution of antenna: (a) 2.5 GHz, (b) 3.5 GHz, and (c) 5.65 GHz.

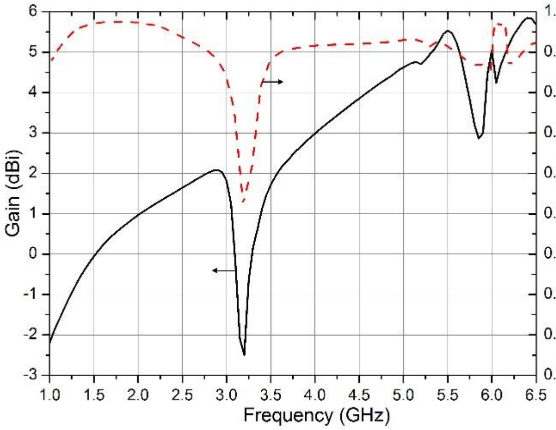


Fig. 10. Gain and radiation efficiency of the proposed antenna.

The radiation patterns at 2.5 GHz, 3.5 GHz and 5.65 GHz of proposed antenna are obtained in a chamber, which are shown in Fig. 11. It can be found from the figure that our designed tri-band antenna has a nearly omnidirectional radiation patterns in the H-plane, and a figure-of-eight radiation patterns in E-plane, which can meet the requirement for the above designated wireless communication applications.

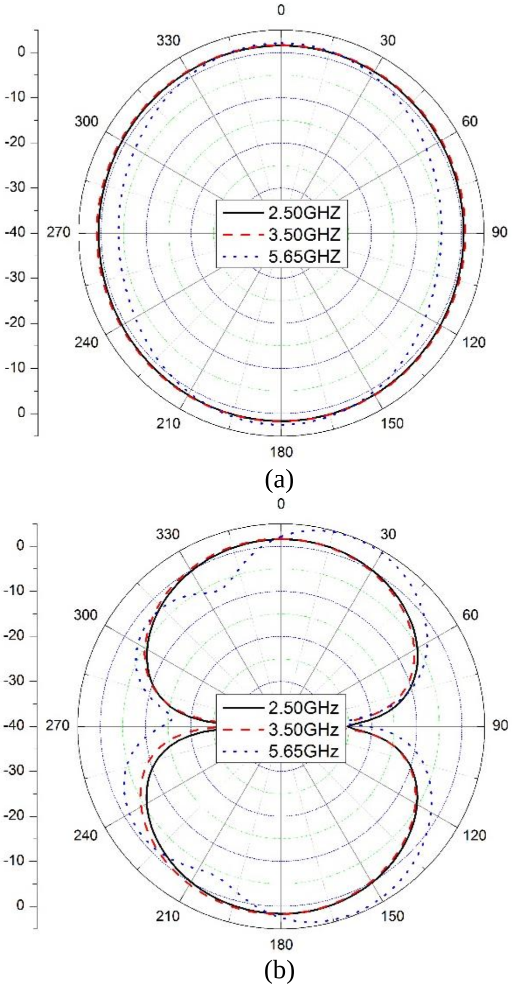


Fig. 11. Radiation patterns of the proposed antenna: (a) H-plane and (b) E-plane.

IV. CONCLUSION

A compact triple-band microstrip antenna has been proposed and its performance has been investigated by the simulation and measurement. Three nested loops and a rectangular stub have been employed to realize the desired triple-band characteristics at 2.5 GHz, 3.5 GHz and 5.6 GHz, which are used for Bluetooth, WLAN and WiMAX wireless communication applications. The proposed antenna has a size of 28 mm × 41 mm, a simple structure. Furthermore, the simulated and experimental results demonstrated that our designed triple-band antenna has a tunable frequency characteristics and good omni-directional radiation patterns, rendering it suitable for multi-band wireless communication system applications.

ACKNOWLEDGMENT

This work was also partially supported by the National Key Research and Development Program of China-Government Corporation Special Program (2016YFE0111100), the Science and Technology innovative Talents Foundation of Harbin (2016RAX XJ044), Projects for the Selected Returned Overseas Chinese Scholars of Heilongjiang Province and MOHRSS of China, and the Foundational Research Funds for the Central Universities (HEUCFD1433, HEUCF160815). This work is also supported in part by the National Science Foundation of China (61671223).

REFERENCES

[1] Y. Li, W. Li, and W. Yu, “A multi-band/UWB MIMO/diversity antenna with an enhance isolation using radial stub loaded resonator,” *Applied Computational Electromagnetics Society Journal*, vol. 28, no. 1, pp. 8-20, 2013.

[2] Y. Li, W. Li, and W. Yu, “A switchable UWB slot antenna using SIS-HSIR and SIS-SIR for multi-mode wireless communications applications,” *Applied Computational Electromagnetics Society Journal*, vol. 27, no. 4, pp. 340-351, 2012.

[3] P. Zibadoost, J. Nourinia, C. Ghobadi, S. Mohammadi, A. Mousazadeh, and B. Mohammadi, “Full band MIMO monopole antenna for LTE systems,” *Applied Computational Electromagnetics Society Journal*, vol. 29, no. 1, pp. 54-61, 2014.

[4] Y. Li, W. Li, and R. Mittra, “Miniaturized CPW-fed UWB antenna with dual frequency rejection bands using stepped impedance stub and arc-shaped parasitic element,” *Microwave and Optical Technology Letters*, vol. 56, pp. 783-787, 2014.

[5] M. N. Shakib, M. Moghavvemi, and W. N. L. Mahadi, “Design of a compact planar antenna for ultra-wideband operation,” *Applied Computational Electromagnetics Society Journal*, vol. 30, no. 2, pp. 222-229, 2015.

[6] Y. Li, W. Li, and Q. Ye, “A compact asymmetric coplanar strip-fed dual-band antenna for 2.4/5.8 GHz WLAN applications,” *Microwave and Optical Technology Letters*, vol. 55, no. 9, pp. 2066-2070, 2013.

[7] X. Sun, L. Liu, S. Cheung, and T. Yuk, “Dual-band antenna with compact radiator for 2.4/5.2/5.8 GHz WLAN applications,” *IEEE Transactions on Antennas and Propagation*, vol. 60, no. 12, pp. 5924-5931, 2012.

[8] A. Mehdipour, A. Sebak, C. Trueman, and T. Denidni, “Compact multiband planar antenna for

2.4/3.5/5.2/5.8-Ghz wireless applications," *IEEE Antennas and Wireless Propagation Letters*, vol. 11, pp. 144-147, 2012.

[9] H. Zhai, Z. Ma, Y. Han, and C. Liang, "A compact printed antenna for triple-band WLAN/WiMAX applications," *IEEE Antennas and Wireless Propagation Letters*, vol. 12, pp. 65-68, 2013.

[10] K. Thomas and M. Sreenivasan, "Compact triple band antenna for WLAN/WiMAX applications," *Electronics Letters*, vol. 45, no. 16, pp. 811-813, 2009.

[11] T. Li, H. Zhai, X. Wang, L. Li, and C. Liang, "Frequency-reconfigurable bow-tie antenna for bluetooth, WiMAX, and WLAN applications," *IEEE Antennas and Wireless Propagation Letters*, vol. 14, pp. 171-174, 2014.

[12] T. Chang and J. Kiang, "Compact multi-band H-shaped slot antenna," *IEEE Transactions on Antennas and Propagation*, vol. 61, pp. 4345-4349, 2013.

[13] C. Hsu and S. Chung, "Compact antenna with U-shaped open-end slot structure for multi-band handset applications," *IEEE Transactions on Antennas and Propagation*, vol. 62, pp. 929-932, 2014.

[14] Y. Cao, S. Cheung, and T. Yuk, "A multiband slot antenna for GPS/WiMAX/WLAN systems," *IEEE Transactions on Antennas and Propagation*, vol. 63, pp. 952-958, 2015.

[15] S. Pan and K. Wong, "Dual-frequency triangular microstrip antenna with a shorting pin," *IEEE Antennas and Wireless Propagation Letters*, vol. 45, pp. 1889-1891, 1997.

[16] Y. Li, T. Jiang, and R. Mittra "A miniaturized dual-band antenna with toothbrush-shaped patch and meander line for WLAN applications," *Wireless Personal Communications*, vol. 91, pp. 595-602, 2016.

[17] S. Pan and K. Wong, "Compact EBG for multi-band applications," *IEEE Transactions on Antennas and Propagation*, vol. 60, pp. 4440-4444, 2012.

[18] M. Booket, A. Jafargholi, M. Kamyab, H. Eskandari, M. Veysi, and S. Mousavi, "Compact multi-band printed dipole antenna loaded with single-cell metamaterial," *IET Microwaves, Antennas & Propagation*, vol. 6, pp. 17-23, 2012.

Multiplication Theory for Prediction of the Scattering Grating-lobe of Array Antenna

Shuai Zhang, Xing Wang, Lixin Guo, and Ji Ma

National Key Laboratory of Antennas and Microwave Technology
Xidian University, Xi'an, 710071, China
zhangshuai@mail.xidian.edu.cn

Abstract **─** Experimental results reported in many literature have demonstrated that an antenna array should have two or more grating lobes in its scattering pattern; however, it is not clear how these grating lobes occur and vary. In this paper, we propose a new method to predict the number and location of scattering grating lobes generated by an array antenna. In the proposed method, the radar cross section (RCS) of array antenna can be decomposed into multiplication of the array RCS factor and the element RCS factor. The array RCS factor bears universal applicability so that it can be used to directly determine scattering properties of array antenna. Compared with the full-wave simulation techniques, the new method requires much less computation time and memory storage so that it is more suitable to be employed as the basis of a synthesis method to predict the desired RCS pattern of a large array antenna. As examples, the scattering properties of the dipole antenna arrays are investigated to validate the new method, numerical results demonstrate that the number and location of scattering grating lobes predicted by using the proposed method coincide with those simulated by using a MoM based commercial; software FEKO.

Index Terms **─** Array antenna, array factor, decomposition, Radar Cross Section (RCS), scattering grating lobe.

I. INTRODUCTION

As the radar cross section (RCS) of military platform itself has been reduced to realize a low observation level along with the development of stealth techniques, there is still an increasing interest to reduce the RCS of array antenna mounted on the platform. To this end, the principal task is how to efficiently predict the scattering feature of an antenna array; therefore, many efforts have been given to the full-wave simulation techniques to simulate scattering patterns for different types of array antennas such as microstrip patch antenna array [1, 2], waveguide slot antenna array [3, 4], dipole antenna array [5, 6], and so on. Both experimental data and numerical simulation results have shown that the scattering grating lobes arise if the inter-element spacing is larger than a half wavelength considering the two way transit of the radar signal, which may be significant in terms of the increased detectability. Hence, an important work is to suppress the grating lobes in the scattering patterns to improve the stealth property of an array antenna.

Since existing of the grating lobes is determined by the spatial arrangement of array elements [7], they cannot be independently suppressed using the traditional stealth techniques such as geometrical shaping optimization [8, 9], applications of radar absorbing materials [10, 11] and other electromagnetic materials [12, 13]. One practicable way is to properly adjust the element positions to break the periodicity. As shown in [14], Haupt applied the hybrid genetic algorithm (GA) to optimize the position of each perfectly conducting strip to generate prescribed backscattering pattern, in which the method of moments (MoM) is employed to evaluate the objective function. This kind of synthesis concept for scattering pattern optimization is also employed by people else to tune other parameters. For example, to obtain the desired RCS pattern of resistive strips, Choi [15] proposed to optimize the resistance distribution varying transversely on the strip to generate corresponding distribution of the induced surface current density on the strip. In [16], Coe successfully realized the desired null in the RCS pattern of a two dipoles array by optimizing the terminal impedance loads. The similar concept was used by Thors and Josefsson [17] to design the low RCS and high radiation performance conformal array antenna.

In the aforementioned RCS synthesis methods, the scattering patterns of various arrays are calculated using full-wave numerical methods, therefore these synthesis methods are time consuming for large arrays or even infeasible when the structure of antenna element is complex. Hence, it is significant to look for an approximation technique that is more efficient to be employed as the basis of a synthesis method to generate the desired RCS pattern of a large and complex array antenna.

Submitted On: October 8, 2016
Accepted On: December 26, 2016

For this purpose, approximation closed-form expressions based on the multiplication theory are derived to calculate the scattering patterns of an array antenna. Extra computational cost for the proposed method increases slightly with an increase of the number of array elements, and it is worth to mention that there is no relation of the simulation time with the antenna type. Moreover, the number and location of scattering grating lobes can be directly determined by using this analytical method, therefore, it can help set up the RCS optimization goal in advance. Utilizing the proposed new method, the relationship between the scattering grating lobes and the inter-element spacing of dipole antenna arrays are discussed.

II. THEORY AND FORMULAS FOR THE SCATTERING FROM ARRAY ANTENNA

Traditionally, the scattered field of an antenna can be decomposed into two components, which are called the antenna and structural modes [18]. The multiplication theory derived in this paper is extracted from the structural mode and antenna model RCS of the array antenna, respectively, which can be used to clarify the physical mechanism of array scattering and to guide the stealth technique choice. At first, the theory of scattered field from single element is briefly reviewed.

A. Scattered field of single element in the array

As the references [6, 18] discussed, when the antenna load impedance Z_L is equal to the characteristic impedance Z_c of the transmission line that connects the antenna and load, the reflection seen from the transmission line to the load is $\Gamma^l = (Z_L - Z_c)/(Z_L + Z_c) = 0$. It is obvious that no energy is reflected back to the antenna and only induced current is on the antenna body. Therefore, the antenna acts as a general passive scattering object. A planar array with *M*×*N* elements is shown in Fig. 1, in which each square represents one element is composed of only metallic conductors. Applying the conclusion to the element in the array, the scattered field of single antenna (*m*, *n*) in the array is given as follows:

$$\vec{E}^s_{mn}(Z_L) = \vec{E}^s_{mn}(Z_c) + \vec{E}^a_{mn}(Z_L), \tag{1}$$

where the first term $\vec{E}^s_{mn}(Z_c)$ on the right side is called the structural mode scattering, and the second term $\vec{E}^a_{mn}(Z_L)$ on the right side is called antenna mode scattering.

B. Extracting the array scattering factor from the structural mode scattering of array antenna

$\vec{E}^s_{mn}(Z_c)$ in (1) represents the scattered field generated by the induced current on the antenna physical structure. This scattered field can be calculated by using the MoM method as follows [19]:

$$\begin{aligned}\vec{E}^s_{mn}(Z_c) = -\frac{j\omega\mu_0}{4\pi r_{mn}} e^{-j\vec{k}_0\cdot\vec{r}_{mn}} \\ \cdot\sum_{l=1}^{L} I_{l(mn)} \int_{s'_{mn}} (\hat{\theta}\hat{\theta} + \hat{\varphi}\hat{\varphi})\, \vec{f}_l(\vec{r}^{\,'}_{mn}) e^{j\vec{k}_0\cdot\vec{r}^{\,'}_{mn}} dS',\end{aligned} \tag{2}$$

where $\vec{k}_0$ is the vector wave number of the free space, ω is the angular frequency of scattered field, the permeability of free space is μ_0, $\vec{r}_{mn}$ represents the observation point with spherical coordinates (*r*, θ, φ), $\vec{r}^{\,'}_{mn}$ is the location of the induced current on the element (*m*, *n*) and S'_{mn} and $\vec{f}_l(\vec{r}^{\,'}_{mn})$ represent the source range and basis function of element (*m*, *n*), respectively. Ignoring the mutual coupling between elements, the current coefficients should have an equal value $I_{l(mn)} = I_l$ (l=1, 2, $\cdots$, *L*). Considering the far field observation and referring to Fig. 1, we have:

$$\vec{r}_{mn} = \vec{r}_{11} - \vec{d}_{mn}, \tag{3}$$

$$\vec{r}^{\,'}_{mn} = \vec{r}^{\,'}_{11} + \vec{d}_{mn}, \tag{4}$$

$$r_{mn} \approx r_{11}. \tag{5}$$

The Equation (2) reduces to:

$$\vec{E}^s_{mn}(Z_c) = \vec{E}^s_{11}(Z_c)\cdot e^{2j\vec{k}_0\cdot\vec{d}_{mn}}, \tag{6}$$

where

$$\begin{aligned}\vec{E}^s_{11}(Z_c) = -\frac{j\omega\mu_0}{4\pi r_{11}} e^{-j\vec{k}_0\cdot\vec{r}_{11}} \\ \cdot\sum_{l=1}^{L} I_l \int_{s'_{11}} (\hat{\theta}\hat{\theta} + \hat{\varphi}\hat{\varphi})\, \vec{f}_n(\vec{r}_{11}{}') e^{j\vec{k}_0\cdot\vec{r}_{11}{}'} dS'\end{aligned} \tag{7}$$

$$\vec{d}_{mn} = \hat{x}\cdot dx_m + \hat{y} dy_n,$$

where dx_m and dy_n are the distances from the element (*m*, *n*) to the element (1, 1) at the origin point in terms of the *x*-and *y*-axes, respectively. The total monostatic structural model scattering of the antenna array is obtained by summing over all the elements:

$$\vec{E}^s(Z_c) = \vec{E}^s_{11}(Z_c)\sum_{m=1}^{M}\sum_{n=1}^{N} e^{2j\vec{k}_0\cdot\vec{d}_{mn}}. \tag{8}$$

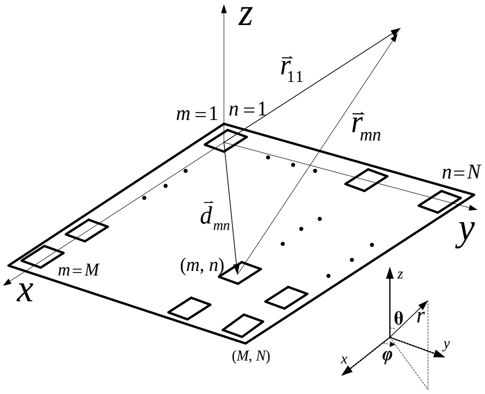


Fig. 1. Configuration of an array antenna with *M*×*N* elements.

C. Extracting the array scattering factor from antenna mode scattering of array antenna

When the incident wave frequency falls into the antenna operating band, the incident energy collected by the antenna travels through a feed network, and is radiated. The radiation field, namely, antenna mode scattering, is expressed as [20]:

$$\vec{E}^a_{mn}(Z_L) = \frac{\Gamma^l_{mn}}{1-\Gamma^A_{mn}\Gamma^l_{mn}} b_{mn}\vec{E}_r, \tag{9}$$

where $\Gamma^l_{mn} = (Z_L - Z_c)/(Z_L + Z_c)$, $\Gamma^A_{mn} = (Z_A - Z_c)/(Z_A + Z_c)$, $\vec{E}_r$ is the radiation field of the element excited by a unit amplitude source with a=1 (watt)$^{1/2}$, b_{mn} is the receiving amplitude of the element (m, n) when terminated with a match load. Then b_{mn} is given as follows [17]:

$$b_{mn} = \frac{1}{2a}\oint_s (\vec{E}^i_{mn} \times \vec{H}^r_{mn} - \vec{E}^r_{mn} \times \vec{H}^i_{mn}) \cdot d\vec{S}, \tag{10}$$

where $(\vec{E}^i_{mn}, \vec{H}^i_{mn})$ and $(\vec{E}^r_{mn}, \vec{H}^r_{mn})$ are the incident field and scattered field of the antenna element, respectively. Substituting (3)-(5) into (10), b_{mn} can be expressed in terms of b_{11} as follows:

$$b_{mn} = b_{11} \cdot e^{j2\vec{k}_0 \cdot \vec{d}_{mn}}. \tag{11}$$

Substituting (11) into (9), the antenna mode scattering of antenna element (m, n) is given by:

$$\vec{E}^a_{mn}(Z_L) = \frac{\Gamma^l_{mn}}{1-\Gamma^A_{mn}\Gamma^l_{mn}} e^{2j\vec{k}_0 \cdot \vec{d}_{mn}} b_{11}\vec{E}_r. \tag{12}$$

Then the total antenna mode scattered field of the array is obtained by summing over all the elements:

$$\vec{E}^a(Z_L) = b_{11}\vec{E}_r \sum_{m=1}^{M}\sum_{n=1}^{N} \frac{\Gamma^l_{mn}}{1-\Gamma^A_{mn}\Gamma^l_{mn}} e^{2j\vec{k}_0 \cdot \vec{d}_{mn}}. \tag{13}$$

Assuming that each element in the array is connected with the identical feed structure and ignoring the mutual coupling effects, the elements should have the equal load impedance so that reflections from them should have the same amplitude and phase. Consequently, (13) can be expressed as follows:

$$\vec{E}^a(Z_L) = \vec{E}^a_{11}(Z_L) \sum_{m=1}^{M}\sum_{n=1}^{N} e^{2j\vec{k}_0 \cdot \vec{d}_{mn}}. \tag{14}$$

D. New expression for RCS of array antenna

Based on (1), (8) and (14), the monostatic scattered field of an array antenna can be written in a multiplication form:

$$\vec{E}^s(Z_L) = E^s_a(Z_L) \cdot \vec{E}^s_e(Z_L), \tag{15}$$

where $E^s_a(Z_L) = \sum_{m=1}^{M}\sum_{n=1}^{N} e^{2j\vec{k}_0 \cdot \vec{d}_{mn}}$ is defined as array scattering factor and $\vec{E}^s_e(Z_L)$ is defined as element scattering factor.

III. SCATTERING GRATING LOBES OF ARRAY ANTENNA

It can be found in literature that an antenna array can generates two or more grating lobes in the RCS pattern [3-5], but it is not clear how these grating-lobes occur and vary. In fact, these grating lobes appear due to the certain inter-element spacing of the array, and this will be proved in the following. Hence, in this section the properties of these grating lobes are analyzed utilizing the array RCS factor. In addition this discussion can also be used to validate (15).

For the sake of mathematic simplification, an *M*-element dipole antenna with an uniform linear distribution, as shown in Fig. 2, is taken as an example. As (15) indicated, once the element is chosen the element scattering factor is determined, and the scattering pattern of array antennas depends mainly on the array scattering factor that includes the inter-element spacing information of array. The array scattering factor in (15) can be rewritten as below:

$$E^s_a(Z_L) = e^{j[(M-1)/2]\psi} \cdot \sin(\frac{M}{2}\psi) \Big/ \sin(\frac{1}{2}\psi). \tag{16}$$

Fig. 2. Configuration of a uniform linear array consists of *M* half-wave dipole antennas.

Then the array MRCS (monostatic RCS) factor can be written as:

$$\sigma_a(\theta,\varphi) = \left|\frac{\vec{E}^s_a(Z_L)}{\vec{E}^i}\right|^2 = \left|\sin(\frac{M}{2}\psi) \Big/ \sin(\frac{1}{2}\psi)\right|^2, \tag{17}$$

where $\psi = 2\vec{k}_0 \cdot \vec{d}_{mn} = 2k_0 d \sin\theta \cos\varphi$, d is the inter-element spacing. The maximum values of (17) occur in the following condition:

$$\frac{\psi}{2} = k_0 d \sin\theta \cos\varphi = \pm l\pi, (l = 0, 1\cdots), \tag{18}$$

where the integer l has a constraint condition of $l \leq |2d/\lambda|$.

Considering the scattering pattern in the plane of $\varphi = 0^{\circ}$ and the range of $-90^{\circ} \leq \theta \leq -90^{\circ}$, (18) reduces to:

$$\begin{aligned} &\frac{\psi}{2} = k_0 d \sin\theta\,|_{\theta=\theta_m} = \pm l\pi \\ &\Rightarrow \theta_m = \sin^{-1}[\pm\frac{\lambda l}{2d}], \end{aligned} \qquad (l = 0, 1\cdots). \tag{19}$$

Utilizing (19), we can simply predict the number and location of the grating lobes in the RCS pattern.

Considering a 11-dipole antenna linear array with an inter-element spacing of 0.6λ, the maximum scattering will occur when *l* equals to 0 and 1 according to (19). For *l*=0 (19) has only one solution of $\theta_m = 0^{\circ}$ and there are two solutions of $\theta_m = \pm 56.4^{\circ}$ for *l*=1. A good agreement between the calculated and FEKO full-wave simulated MRCS patterns is evident, as shown in Fig. 3. From Fig. 3, one of the maximum scattering grating lobes occurs at the angle $\theta_m = 0^{\circ}$ due to the specular scattering of the array aperture. The other two maximum scattering grating lobes are lower than that of specular scattering because the amplitudes of the element RCS factor at angles $\theta_m = \pm 56.4^{\circ}$ is lower than that at $\theta_m = 0^{\circ}$, as shown in Fig. 4.

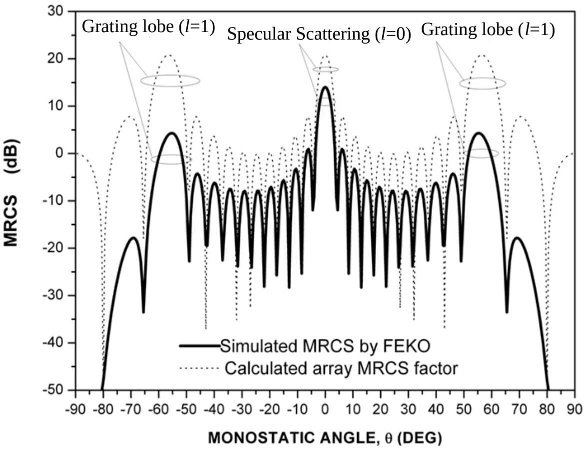


Fig. 3. MRCS of 11-dipole linear array ($d = 0.6\lambda$) simulated by using FEKO and the array MRCS factor calculated by using (17).

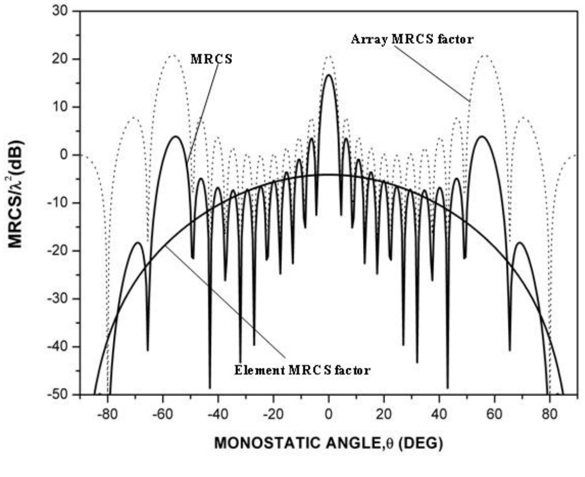


Fig. 4. Calculated array MRCS factor, element MRCS factor and the MRCS of 11-dipole linear array ($d = 0.6\lambda$) by using (15) and (17).

To further investigate the relationship between the scattering grating lobes and the inter-element spacing, the inter-element spacing of the 11-dipole antenna linear array discussed above is extended to $d = 1.4\lambda$. According to (19), the integer *l* should have three values, i.e., 0, 1 and 2, and there are five maximum scattering angles in the scattering pattern. Figure 5 depicts the FEKO full-wave simulated MRCS pattern and calculated MRCS factor pattern of the array antenna, and it can be observed from Fig. 5 that the five maximum scattering grating lobes correspond to different values of the integer *l*.

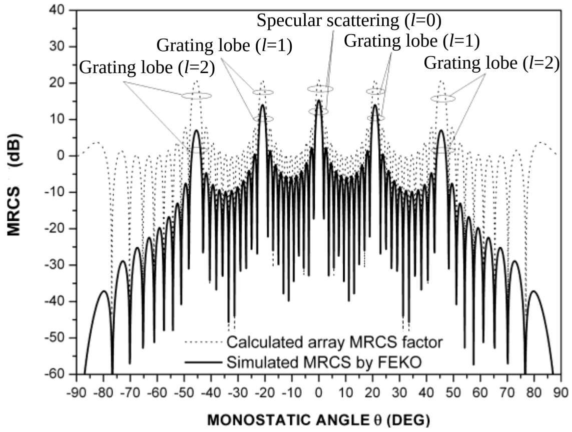


Fig. 5. FEKO full-wave simulated MRCS of 11 half-wave dipoles linear array ($d = 1.4\lambda$) and the calculated array MRCS factor by using (17).

As discussed above, we can conclude that different inter-element spacing makes the grating-lobes in the MRCS pattern of array antenna appear in different incident directions. With the increase of inter-element spacing, more grating lobes will occur in the MRCS pattern and move toward the normal incident direction. This is similar to the radiation of array antenna [19, 21].

IV. CONCLUSION

The scattering grating lobes of an array antenna can be represented by the multiplication of array RCS factor and the element RCS factor. The relationship between scattering grating lobes and inter-element spacing of a typical uniform array antenna were analyzed in this paper. FEKO full-wave simulation results and the calculation results based on the proposed method demonstrated that the new method can fast and accurately determine the number and location of the scattering grating lobes of a practical array antenna. Since the maximum scattering grating lobes of an array antenna depends mainly on the array RCS factor, the conclusions obtained in this paper are also applicable for other types of antennas with the same array distribution (array RCS factor). To eliminate the scattering grating lobes, one can take the formulation here as the base of synthesis methods to optimize the element positions.

ACKNOWLEDGMENT

This work was supported by National Natural Science Foundation of China (No. 61601350, 61501343), China Postdoctoral Science Foundation (No. 2015M58 0813), Natural Science Basic Research Plan in Shaanxi Province of China (No. 2015JQ6217) and Natural Science Foundation of Ningbo of China (No. 2016A610031).

REFERENCES

[1] J. Jian-Ming and J. L. Volakis, "A hybrid finite element method for scattering and radiation by microstrip patch antennas and arrays residing in a cavity," *IEEE Trans. Antennas Propagat*., vol. 39, no. 11, pp. 1598-1604, Nov. 1991.

[2] W. J. Tsay and D. M. Pozar, "Radiation and scattering from infinite periodic printed antennas with inhomogeneous media," *IEEE Trans. Antennas Propagat*., vol. 46, no. 11, pp. 1641-1650, Nov. 1998.

[3] L. Zhang, N. Yuan, M. Zhang, L. W. Li, and Y. B. Gan, "RCS computation for a large array of waveguide slots with finite wall thickness using the MoM accelerated by P-FFT algorithm," *IEEE Trans. Antennas Propagat*., vol. 53, no. 9, pp. 3101-3105, Sep. 2005.

[4] M. Zhang, L. W. Li, and A. Y. Ma, "Analysis of scattering by a large array of waveguide-fed wide-slot millimeter wave antennas using precorrected-FFT algorithm," *IEEE Microw. Wireless Compon. Lett*., vol. 15, no. 11, pp. 772-774, Nov. 2005.

[5] P. J. Tittensor and M. L. Newton, "Prediction of the radar cross section of an array antenna," *Internation Conference on Antennas and Propagation*, vol. 1, pp. 258-262, 1989.

[6] D. C. Jenn and V. Flokas, "In-band scattering from arrays with parallel feed network," *IEEE Trans. Antennas Propagat*., vol. 44, no. 2, pp. 172-178, Feb. 1996.

[7] R. L. Haupt, "Grating lobes in the scattering patterns of edge-loaded strips," *IEEE Trans. Antennas Propagat*., vol. 41, no. 8, pp. 1139-1143, 1993.

[8] W. W. Xu, J. H. Wang, M. I. Chen, et al., "A novel microstrip antenna with composite patch structure for reduction of in-band RCS," *IEEE Antennas and Wireless Propagation*, vol. 14, pp. 139-142, 2015.

[9] C. M. Dikmen and S. Cimen, "Planar octagonal-shaped UWB antenna with reduced radar cross section," *IEEE Trans. Antennas Propagat*., vol. 62, no. 6, pp. 2946-2953, June 2014.

[10] M L. Li, Z. X. Yi, and Y. H. Luo, "A novel integrated switchable absorber and radiator," *IEEE Trans. Antennas Propagat*., vol. 64, no. 3, pp. 944-952, Mar. 2016.

[11] T. Liu, X. Y. Cao, J. Gao, et al., "RCS reduction of waveguide slot antenna with metamaterial absorber," *IEEE Trans. Antennas Propagat*., vol. 61, no. 3, pp. 1479-1484, Mar. 2013.

[12] M. Z. Joozdani, M. K. Amirhosseini, and A. Abdolali, "Wideband radar cross section reduction of patch array antenna with miniaturized hexagonal loop frequency selective surface," *Electronics Letters*, vol. 52, no. 9, pp. 767-768, 2016.

[13] H. Jiang, Z. H. Xue, W. M. Li, et al., "Low-RCS high-gain partially reflecting surface antenna with metamaterial ground plane," *IEEE Trans. Antennas Propagat*., vol. 64, no. 9, pp. 4127-4132, Sep. 2016.

[14] R. L. Haupt and Y. C. Chung, "Optimizing backscattering from arrays of perfectly conducting strips," *IEEE Antennas and Propagation Magazine*, vol. 45, no. 5, pp. 26-33, Oct. 2003.

[15] J. I. Choi, B. H. Lee, and E. J. Park, "Optimum current distribution on resistive strip for arbitrarily prescribed RCS pattern," *International Conference on Microwave and Millimeter Wave Technology*, pp. 363-366, 1998.

[16] R. J. Coe and Akira Ishimaru, "Optimum scattering from an array of half-wave dipoles," *IEEE Trans. Antennas Propagat.*, vol. 18, no. 2, pp. 224-229, Mar. 1970.

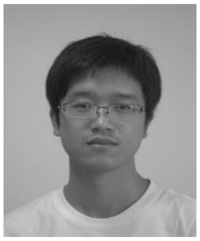
[17] B. Thors and L. Joesfsson, "Radiation and scattering tradeoff design for conformal arrays," *IEEE Trans. Antennas Propagat*., vol. 51, no. 5, pp. 1069-1076, May 2003.

[18] R. C. Hansen, "Relationships between antennas as scatterers and as radiators," *in Proc. IEEE*, vol. 77, no. 5, pp. 659-671, May 1989.

[19] S. M. Rao, D. R. Wilton, and A. W. Glisson, "Electromagnetic scattering by surfaces of arbitrary shape," *IEEE Trans. Antennas Propagat.*, vol. 18, no. 3. pp. 409-418, May 1982.

[20] Y. Liu, D. M. Fu, and S. X. Gong, "A novel model for analyzing the RCS of microstrip antenna," *J. Electromagn. Waves Appl*., vol. 17, no. 9, pp. 1301-1310, 2003.

[21] Y. T. Lo and S. W. Lee, *Antenna Handbook: Theory, Applications, and Design*. Van Nostrand Reinhold Company, New York, 1988.



Shuai Zhang was born in Hubei province, China. He received the B.S. and Doctorate degrees in Electromagnetic Field and Microwave Technique from Xidian University, Xi'an, in 2007 and 2012, respectively. He is currently working as Lecturer at National Key Laboratory of Antennas and Microwave Technology, Xidian University. His current research interests include antenna design, analysis, calculation and control techniques of the radiation and scattering characteristics of array antennas, and optimization methods in electromagnetics.

Compact UWB Microstrip-fed Slot Antenna with Dual-band Rejection by Using EBG and Four-arm Spiral Structures

Ji Ma [1], **Shuai Zhang** [2], **Qiu-ju Li** [3], **and Jun Du** [4]

[1] The 27th Research Institute of China Electronics Technology Group Corporation
Zhengzhou, Henan, 450047, China

[2] National Key Laboratory of Antennas and Microwave Technology
Xidian University, Xi'an, 710071, China
zhangshuai@mail.xidian.edu.cn

[3] Military Representative Office of Army Reside in Zhengzhou, China

[4] Huayin Ordnance Test Center, China

Abstract **─** A microstrip-fed slot antenna with dual rejection bands is designed and prototyped for ultra-wideband (UWB) applications. The rejection band for WiMAX is achieved by symmetrical etching a pair of nested four-arm spiral slots on the ground plane. Self-interdigital electromagnetic band-gap (EBG) structures are embedded to create other rejection bands for WLAN and are placed on both sides of the gradient T-shaped microstrip-fed. The simulated result and measurement data show that the antenna impedance bandwidth (Voltage Standing Wave Ratio is less than 2) covers the entire UWB bandwidth range with two eliminated bands of 3.22-3.81 GHz (0.59 GHz, 16.7%) and 5.11-5.84 GHz (0.73 GHz, 13.4%). In addition, the designed antenna achieves a stable gain and exhibits omni-directional radiation patterns except at rejection frequency bands, which makes it a suitable candidate for UWB applications that will not be interfered by WiMAX/WLAN systems.

Index Terms **─** Dual band rejection, electromagnetic band gap, UWB antenna.

I. INTRODUCTION

Ultra-wideband (UWB) technology has become one of the most promising technologies for its inherent advantages, such as low power consumption, high security, low cost, high speed transmission rate, and so on [1]. Over the past ten years, as a result of accelerating growth of UWB technology, a vast body of literature introduce novel antennas for various UWB system applications [2-7].

UWB communication devices occupy a large frequency spectrum (3.1-10.6 GHz) [8], which covers several other wireless system bands, such as WiMAX (3.3-3.7 GHz) and WLAN (5.15-5.825 GHz). The UWB devices met a hostile electromagnetic environment, which may cause potential interferences to it. Solution to avoid interference is the introduction of rejection bands within the pass band of the UWB antenna.

Various techniques for the design of UWB antennas with notch band characteristics have been reported including etching slot on the radiating patch or on the ground plane [9-11], application of parasitic element near the radiating patch or the ground plane [12-14], and so on. In [15], a compact UWB microstrip-fed slot antenna with the band-stop performance is designed by employing genetic algorithm to optimize the shape of slot. In [16], a compact band-notched UWB circular monopole antenna is introduced with dimensions of 39 mm×35 mm. The notch band for WLAN is realized by means of four mushroom-like electromagnetic band-gap (EBG) structures, which are placed on sides of the feed line. A dual notched band UWB planar antenna with smaller dimensions of 32 mm×26 mm is presented in [17], a T-shaped stub on the radiating element is used to create notch band for WiMAX, and a pair of U-shaped stubs on the side of the feed line are used to generate notch band for WLAN. In [18], the microelectro-mechanical system (MEMS) U-shaped afloat is designed and used instead of a simple slot to activate and deactivate the band notch characteristic of an UWB monopole antenna. However, the utilization of MEMS is difficult for realization and usually results in cost and fabrication error increasing. Generally, the design goal of UWB antenna with band notch characteristic is to achieve small size, low profile, high gain, light weight, low cost and simple structure.

In this paper, a compact design and new structure

Submitted On: October 8, 2016
Accepted On: December 26, 2016

of microstrip-fed slot antenna with much smaller dimensions of 23.8 mm×22 mm is proposed for UWB applications. Two rejection bands of WiMAX and WLAN can be realised in the pass band. A stable gain and omni-directional radiation patterns except at rejection bands are obtained. Detailed controlling mechanisms, simulation results and measurement data are demonstrated.

II. ANTENNA DESIGN

Microstrip-fed wide slot antennas [5, 9-11, 14] have been extensively investigated in the past three decades due to its attractive features, such as low profile, light weight, low cost and wide band. In order to cover the UWB frequency spectrum, the shape of the microstrip-fed, the slot and the ground plane should be properly collocation. Figure 1 shows the configuration of the proposed UWB slot antenna, in which a gradient T-shaped feed line is designed for broadband matching and is printed on the front side of a double-sided FR4 dielectric substrate with a relative permittivity of 4.4, a thickness of 1.6 mm and a loss tangent of 0.02. To achieve 50Ω characteristic impedance, the parameters of the feed line are chosen to be L_1=5 mm, L_2=2.7 mm, L_3=2 mm, W_3=6 mm and W_4=2.2 mm.

Since the EBG structure has a surface wave suppression characteristics [19], various types of these structures have been implemented in different applications such as reduction of mutual coupling between two planar antennas to eliminate spurious responses of a filter [17]. In this paper, we utilize the EBG structure to create a stop band within the pass band of the UWB slot antenna. As shown in Fig. 1, a compact self-interdigital EBG structure resonated at 5.5 GHz is designed and there are two elements on each side of the feed line for filtering the WLAN frequency spectrum. The EBG patch is printed on the front side of FR4 dielectric substrate and its parameters are chosen to be W_5=1.6 mm, W_6=0.2 mm, L_6=4.4 mm and g=0.2 mm.

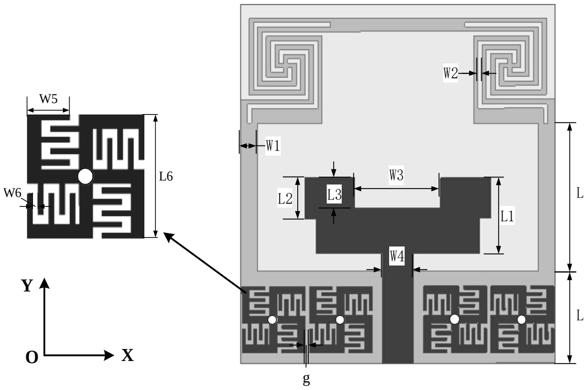
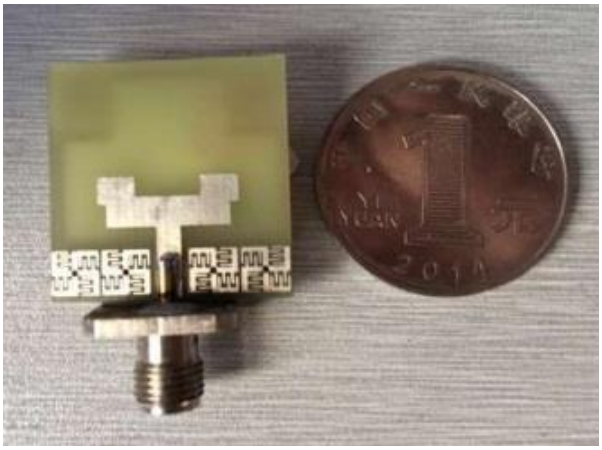


Fig. 1. Configuration of the proposed dual band-notched UWB slot antenna.

The ground plane is printed on the back side of the FR4 dielectric substrate, and we introduce a pair of four-arm spiral slots to generate another rejection band for WiMAX systems. As shown in Fig. 1, the two four-arm spiral slots are symmetrically etched in the front part of the ground plane. The width of the microstrip line and the gap between strips of the four arm spiral slot have the same dimension of 0.3 mm. Figure 2 shows the photograph of the proposed UWB slot antenna, which has a smaller size of 23.8 mm×22 mm×1.6 mm compared to the previous similar antennas [16, 17]. The other antenna parameters are as follows: L_4=9.8 mm, L_5=6.1 mm, W_1=1.5 mm and W_2=0.3 mm.



(a) Front view of the antenna

(b) Back view of the antenna

Fig. 2. Fabricated dual band-notched UWB slot antenna.

III. RESULTS AND DISCUSSION

The simulated results and measured voltage standing wave ratio (VSWR) of the proposed antenna are plotted in Fig. 3. From Fig. 3, two rejection bands of 3.22-3.81 GHz (0.59 GHz, 16.7%) and 5.11-5.84 GHz (0.73 GHz, 13.4%) are observed in the VSWR characteristics, respectively. Taking VSWR less than 2 as a reference, it can be seen that the proposed antenna achieves an ultra-wide operating band ranging from 2.74 to 12.32 GHz with two notched bands. Thus, the proposed UWB slot antenna can availably shield the interference from WiMAX and WLAN systems. It should be mentioned that the discrepancy between the simulation result and measurement data is mainly because of the fabrication error.

The simulated peak gain of the proposed antenna is shown in Fig. 4. Obviously, the gain falls sharply in the two rejection bands since the radiated power is reflected back to the antenna. At the central frequency of the WiMAX band, the simulated peak gain and the measured data are suppressed to -2.91 dB and -3.05 dB, respectively. On the other hand, the simulated gain and the measured data are -2.65 dB and -2.48 dB at the central frequency of the WLAN band, respectively. It clearly indicates the effect of the designed EBG structures and four-arm spiral slots. Other than the rejection bands, the antenna has a stable gain from 1.78 dB to 4.53 dB within the operating band.

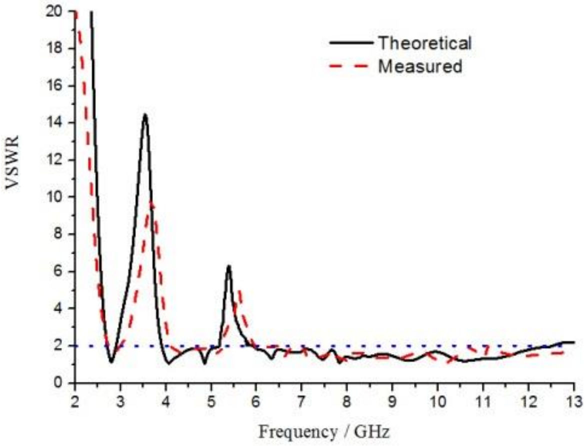


Fig. 3. VSWR variation of the proposed antenna with frequency.

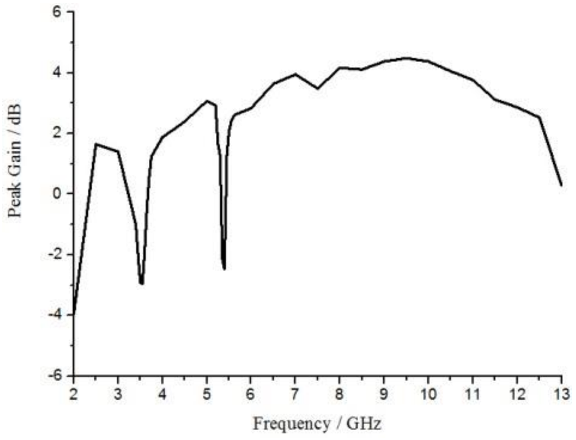
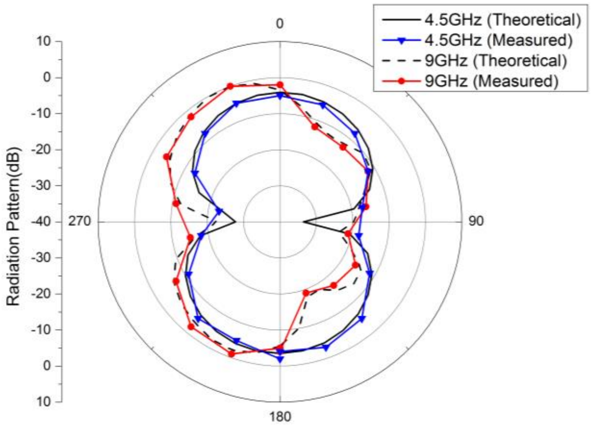
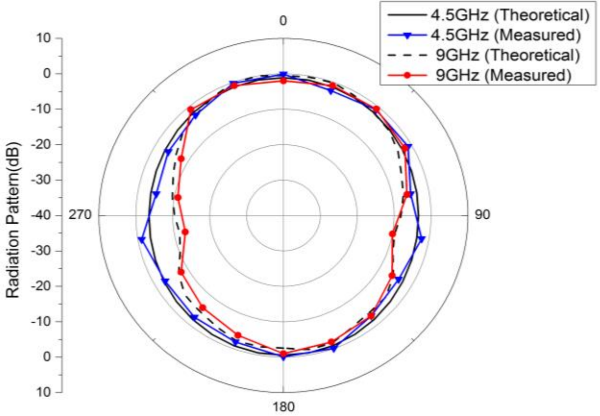


Fig. 4. Peak gain variation of the proposed antenna with frequency.

The simulated results and measured radiation patterns at 4.5 GHz and 9 GHz in the *xoz*- and *yoz*-plane are depicted in Figs. 5 (a) and (b), respectively. There is a slight difference observed between the simulation result and measurement data as a result of the fabrication error. It is found that the radiation patterns are in the omni-directional pattern in the *yoz*-plane (H-plane), whereas in the *xoz*-plane (E-plane), two nulls in the broadside directions are observed that are similar to the typical monopole antennas. Compared with the past dual notched band UWB antennas, the proposed antenna has advantages of small size, low profile, high gain, light weight, low cost and simple structure, which make the proposed antenna a suitable candidate for the UWB applications that is not interfered by WiMAX and WLAN systems.



(a) E-plane pattern in the *xoz* plane



(b) H-plane pattern in the *yoz* plane

Fig. 5. Radiation pattern of the proposed antenna at 4.5 and 9 GHz.

IV. CONCLUSION

A compact UWB slot antenna with dual rejection band characteristics is designed and prototyped. The designed basic UWB antenna consists of a gradient T-shaped feed line and ground plane with a wide rectangle slot. By etching a pair of four-arm spiral slots and embedding self interdigital EBG structures, the proposed slot antenna can achieve two rejection bands of 3.22-3.81 and 5.11-5.84 GHz, which covers the entire WiMAX and WLAN bandwidth ranges, respectively. Meanwhile, the proposed antenna exhibits good performance such as stable gain and omni-directional radiation patterns except at rejection bands are obtained.

ACKNOWLEDGMENT

This work was supported by National Natural Science Foundation of China (No. 61601350, 61501343), China Postdoctoral Science Foundation (No. 2015M580813), Natural Science Basic Research Plan in Shaanxi Province of China (No. 2015JQ6217) and Natural Science Foundation of Ningbo of China (No. 2016A610031).

REFERENCES

[1] I. Oppermann, M. Hamalainen and J. Iinatti, *UWB Theory and Applications*. New York, NY, USA: Wiley, ch. 1, pp. 3-4, 2004.

[2] L. A. Y. Poffelie, P. J. Soh, S. Yan, and G. A. E. Vandenbosch, “A high-fidelity all-texitile UWB antenna with low back radiation for off-body WBAN applications,” *IEEE Transactions on Antennas and Propagation*, vol. 64, no. 2, pp. 757-760, Feb. 2016.

[3] N. Ojaroudi, M. Ojaroudi, N. Ghadimi, and M. Mehranpour, “UWB square monopole antenna with omni-directional radiation patterns for use in circular cylindrical microwave imaging systems,” *Applied Computational Electromagnetics Society (ACES) Journal*, vol. 28, no. 2, pp. 123-129, Feb. 2013.

[4] W. H. Jeong, J. Tak, and J. Choi, “A low-profile IR-UWB antenna with ring patch for WBAN applications,” *IEEE Antennas and Wireless Propagation Letters*, vol. 14, pp. 1447-1450, 2015.

[5] R. V. S. R. Krishna and R. Kumar, “A dual-polarized square-ring slot antenna for UWB, imaging, and radar applications,” *IEEE Antennas and Wireless Propagation Letters*, vol. 15, pp. 195-198, 2016.

[6] S. Koziel and A. Bekasiewicz, “A structure and simulation-driven design of compact CPW-Fed UWB antenna,” *IEEE Antennas and Wireless Propagation Letters*, vol. 15, pp. 750-753, 2016.

[7] A. Amini, H. Oraizi, and M. A. C. Zadeh, “Miniaturized UWB log-periodic square fractal antenna,” *IEEE Antennas and Wireless Propagation Letters*, vol. 14, pp. 1322-1325, 2015.

[8] Federal Communications Commission. First Report and Order. Revision of part 15 of the Commission’s Rules Regarding Ultra-wideband Transmission Systems. USA: FCC, February 2002.

[9] B. Adamchi and J. Nourinia, “Design of compact reconfigurable ultra-wideband slot antenna with switchable single/dual band notch functions,” *IET Microwave, Antennas, & Propagation*, vol. 8, no. 8, pp. 541-548, Aug. 2014.

[10] Y. P. Zhang and C. M. Li, “Design of small dual band-notched UWB slot antenna,” *Electronic Letters*, vo. 51, no. 22, pp. 1727-1728, Oct. 2015.

[11] M. Ojaroudi and N. Ojaroudi, “Ultra-wideband small rectangular slot antenna with variable band-stop function,” *IEEE Transactions on Antennas and Propagation*, vol. 62, no. 1, pp. 490-494, Jan. 2016.

[12] J. Mazloum and N. Ojaroudi, “Utilization of protruded strip resonators to design a compact UWB antenna with WiMAX and WlAN notch bands,” *Applied Computational Electromagnetics Society (ACES) Journal*, vol. 31, no. 2, pp. 204-209, Feb. 2016.

[13] J. Kim, Y. J. Cho, and S. H. Hwang, “Band-notched UWB planar monopole antenna with two parasitic patches,” *Electronic Letters*, vol. 41, no. 14, pp. 783-785, July 2005.

[14] A. Rezaul and T. I. Mohammad, “Design of a dual band-notch UWB slot antenna by means of simple parasitic slits,” *IEEE Antennas and Wireless Propagation Letters*, vol. 12, pp. 1412-1415, 2013.

[15] A. Nemati and B. A. Ganji, “UWB monopole antenna with switchable band-notch characteristic using a novel MEMS afloat,” *Applied Computational Electromagnetics Society (ACES) Journal*, vol. 30, no. 12, pp. 1306-1312, Feb. 2015.

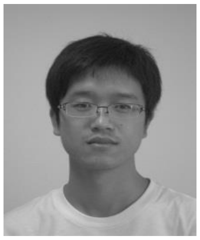
[16] W. Jiang and W. Che, “A novel UWB antenna with dual notched bands for WiMAX and WLAN applications,” *IEEE Antennas and Wireless Propagation Letters*, vol. 11, pp. 293-296, 2012.

[17] M. Yazdi and N. Komjani, “Design of a band-notched UWB monopole antenna by means of an EBG structure,” *IEEE Antennas and Wireless Propagation Letters*, vol. 10, pp. 170-173, 2011.

[18] J. Zolghadr, Y. Cai, and N. Ojaroudi, “UWB slot antenna with band-notched property with time domain modeling based on genetic algorithm optimization,” *Applied Computational Electromagnetics Society (ACES) Journal*, vol. 31, no. 8, pp. 926-932, Feb. 2016.

[19] S. Barth and A. K. Iyer, “A miniaturized uniplanar metamaterial-based EBG for parallel-plate mode suppression,” *IEEE Transactions on Microwave Theory and Techniques*, vol. 64, no. 4, pp. 1176-1185, Feb. 2016.

Ji Ma was born in Henan Province, China. He received the B.S. and Doctorate degrees in Electromagnetic Field and Microwave Technique from Xidian University, Xi'an, in 2008 and 2013, respectively. He is currently working as Engineer at the 27th Research Institute of China Electronics Technology Group Corporation, Zhengzhou, Henan, China. His current research interests include antenna design, analysis and computational electromagnetics.



Shuai Zhang was born in Hubei province, China. He received the B.S. and Doctorate degrees in Electromagnetic Field and Microwave Technique from Xidian University, Xi'an, in 2007 and 2012, respectively. He is currently working as Lecturer at National Key Laboratory of Antennas and Microwave Technology, Xidian University. His current research interests include antenna design, analysis, calculation and control techniques of the radiation and scattering characteristics of array antennas, and optimization methods in electromagnetics.

A Low Complexity High Performance Robust Adaptive Beamforming

Wenxing Li [1], Xiaojun Mao [*1,3], Zhuqun Zhai [2], and Yingsong Li [1]

[1] College of Information and Communications Engineering
Harbin Engineering University, Harbin 86 150001, People's Republic of China
liwenxing@hrbeu.edu.cn, wwwmaoxiaojun@126.com, liyingsong@hrbeu.edu.cn

[2] Naval Academy of Armament, Beijing 86 100161, People's Republic of China
zzz_iii@163.com

[3] Shanghai Radio Equipment Research Institute, Shanghai 200090, People's Republic of China
wwwmaoxiaojun@126.com

Abstract **─** In this paper, we propose a low complexity adaptive beamforming with joint robustness against array steering vector (ASV) mismatch and array manifold errors. The proposed robust beamforming is based on the diagonal loading technique, and the diagonal loading factor can be adjusted adaptively according to the input signal-to-noise ratios (SNRs). The eigenvalue corresponding to the desired signal is identified by projecting the presumed ASV of the desired signal onto the eigenvectors of the sample covariance matrix. We find that the ratio of the eigenvalues corresponding to the desired signal and the noise can be used to accurately reflect the input SNR when the input SNRs is large enough, and the ratio is also the diagonal loading factor. Then, an orthogonal subspace is obtained with the steering vectors associated with the region where the desired signal may locate at. Also, the actual steering vector is estimated as a linear combination of this orthogonal subspace. In order to further reduce the computation complexity, we use the traditional diagonal loading method in low SNRs. Simulations results demonstrate that the proposed beamformer provides strong robustness against a variety of array mismatches, and the complexity is significantly reduced compared with popular existing methods.

Index Terms **─** Adaptive arrays, diagonal loading, low complexity, robust beamforming.

I. INTRODUCTION

Adaptive beamforming has been widely used in radar, sonar, mobile communications, radio astronomy and other fields due to its ability on suppressing the interferences and noise [1-4]. The standard Capon beamformer (SCB), which is one of the representative adaptive beamformers, has excellent resolution and jammer rejection performance in ideal case. Unfortunately, the SCB lacks robustness in the presence of model mismatches, especially, when the desired signal is presented in the training data. Therefore, the behavior of the SCB significantly degraded when the mismatch existed in the array steering vectors (ASV) or the array manifold [5-10].

Many robust adaptive beamforming methods have been developed over the past several decades [11-19]. Among these methods, the diagonal loading method is the most common one due to its lowest complexity, where a fixed value is added to the diagonal of sample covariance matrix [11]. However, it doesn't provide any guidance to select the optimal diagonal loading factor, and thus it cannot provide sufficient robustness. Robust Capon beamformer (RCB) was proposed in [12], which exploited a spherical or ellipsoidal uncertainly set to limit the ASV of the desired signal. This method has been proved belongs to a kind of diagonal loading approach except that the loading factor can now be determined precisely based on the uncertainty set. However, the performance of the RCB is mainly determined by the uncertain parameter set, and uncertain parameter set is difficult to be known accurately in practice. The RCB is equivalent with the worst-case beamformer proposed in [13].

Robust adaptive beamforming based on steering vector estimation has been proposed in [14]. To achieve an accurate steering vector, one needs to maximize the beamformer output power, and the ASV prevented from converging to any interference ASVs or their linear combinations, which is a quadratically constrained quadratic programming (QCQP) problem and can be converted to semi-definite programming (SDP). Certainly, the large amount of calculation is needed and long time to converge. In [15], robust adaptive beamforming based on interference-plus-noise (IPN) covariance matrix reconstruction and ASV estimation has been proposed,

Submitted On: September 29, 2016
Accepted On: March 23, 2017

the IPN covariance matrix was reconstructed by utilizing the Capon spectrum to integrate over a region separated from the SOI direction. This method enjoys good performance in the case of ASV direction errors. Unfortunately, the array manifold information must be known accurately in this method. As a result, this method lacks robustness in the presence of array calibration errors, especially in low signal-to-noise ratios (SNRs) [16-17]. Then, a variable loading method is proposed in [19], which can improve the robustness by deliberately preventing the weight vector converging to the noise components, where the loading factor is set in an ad hoc manner.

In this paper, we propose a novel variable diagonal loading beamforming to achieve high performance and low complexity. The diagonal loading factor is associated with the input SNRs and can be adaptively adjusted. In our proposed beamformer, the eigenvalue corresponding to the desired signal is identified by projecting the presumed ASV of the desired signal onto the eigenvectors of the sample covariance matrix. The input SNR is estimated by using the eigenvalues corresponding to the desired signal and noise, respectively, which is also the diagonal loading factor in high SNR. The traditional diagonal loading method is used in low SNR for further reducing the computation complexity. A correlation matrix of the ASV associated with the region where the desired signal may locate is constructed and an orthogonal subspace is obtained by Eigen-decomposition of this matrix. The actual ASV of the desired signal is subsequently expressed as a linear combination of the orthogonal subspace. The proposed method with low complexity can obtain a closed-form expression of the weight vector. Simulation results show the proposed beamforming can provide strong joint robustness against ASV mismatch and array manifold errors.

II. THE SIGNAL MODEL

We consider a uniform linear array (ULA) with *N* unidirectional antennas with spacing *d*. Assuming that there are $M+1$ signals arriving from the directions θ_p, $p=0,1,\cdots,M$. The received data of the array can be expressed as:

$$\mathbf{X}(k)=\mathbf{A}\mathbf{S}(k)+\mathbf{N}(k), \tag{1}$$

where $\mathbf{X}(k)=[x_1(k)\ x_2(k)\ \cdots\ x_N(k)]^T$ is a $N\times 1$ array observations data vector. $(\bullet)^T$ denotes the transpose. k is the time index. $\mathbf{S}(k)=[s_0(k)\ s_1(k)\ \cdots\ s_M(k)]^T$, $s_p(k)$ denotes the complex waveform of the *p*th signal. Here, $s_0(k)$ is considered as the desired signal, while $s_i(k)$, $i=1,2,\cdots,M$ are the interferences. $\mathbf{N}(k)=[n_1(k)\ n_2(k)\ \cdots\ n_N(k)]^T$ is a vector of the additive white sensor noise, $\mathbf{A}=[\mathbf{a}(\theta_0)\ \mathbf{a}(\theta_1)\ \cdots\ \mathbf{a}(\theta_M)]$, where $\mathbf{a}(\theta_p)=[1\ \mathrm{e}^{j\beta_p}\ \cdots\ \mathrm{e}^{j(N-)1\beta_p}]^T$ represents a steering vector in the θ_p direction, and β_p is the wave number that can be represented as $\beta_p=2\pi d\sin(\theta_p)/\lambda$.

We assume that the signal and noise are statistically independent. The output of the beamformer $y(k)$ is given by:

$$y(k)=\mathbf{w}^{\mathrm{H}}\mathbf{X}(k), \tag{2}$$

where $\mathbf{w}$ is the $N\times 1$ complex weight vector and $(\bullet)^H$ represents the Hermitian transpose.

The minimum variance distortionless response (MVDR) beamformer is formulated as the following linearly constrained quadratic optimization problem:

$$\min_{\mathbf{w}}\ \mathbf{w}^H\mathbf{R}_{i+n}\mathbf{w}\quad \text{subject to}\ \ \mathbf{w}^H\bar{\mathbf{a}}_0=1, \tag{3}$$

where $\bar{\mathbf{a}}_0$ is the presumed ASV of the desired signal, $\mathbf{R}_{i+n}$ is the IPN covariance matrix. In practice, $\mathbf{R}_{i+n}$ is commonly replaced by the sample covariance matrix:

$$\hat{\mathbf{R}}=\frac{1}{K}\sum_{k=1}^{K}\mathbf{X}(k)\mathbf{X}^H(k), \tag{4}$$

where K is the number of snapshots. Thus, the optimal solution to (3) is:

$$\mathbf{w}=\frac{\hat{\mathbf{R}}^{-1}\bar{\mathbf{a}}_0}{\bar{\mathbf{a}}_0^H\hat{\mathbf{R}}^{-1}\bar{\mathbf{a}}_0}. \tag{5}$$

The solution (5) is commonly referred to as the sample matrix inverse (SMI). The standard MVDR beamformer has a good interference rejection performance by producing sharp nulls at the direction of interferences in the ideal case. However, the standard MVDR beamformer lacks robustness against the ASV or manifold errors, which case seriously performance degradation.

III. THE PROPOSED ALGORITHM

A. Input SNR estimation and diagonal loading

In this section, we propose a novel low complexity variable diagonal loading beamformer. In this method, the eigenvalue corresponding to desired signal is identified firstly. The sample covariance matrix $\hat{\mathbf{R}}$ defined by (4) can be decomposed as:

$$\hat{\mathbf{R}}=\sum_{i=1}^{N}\lambda_i\boldsymbol{\upsilon}_i\boldsymbol{\upsilon}_i^H=\mathbf{U}_s\boldsymbol{\Lambda}_s\mathbf{U}_s^H+\mathbf{U}_n\boldsymbol{\Lambda}_n\mathbf{U}_n^H, \tag{6}$$

where λ_i, $i=1,2,\cdots,N$ are the eigenvalues are of $\hat{\mathbf{R}}$, and $\boldsymbol{\upsilon}_i$, $i=1,2,\cdots,N$ are the corresponding eigenvectors. $\mathbf{U}_s=[\boldsymbol{\upsilon}_1\ \boldsymbol{\upsilon}_2\ \ldots\ \boldsymbol{\upsilon}_{M+1}]$ represents the signal-plus-interference (SPI) subspace, which is composed by the *M* interferences and a desired signal. $\mathbf{U}_n=[\boldsymbol{\upsilon}_{M+2}\ \boldsymbol{\upsilon}_{M+3}\ \ldots\ \boldsymbol{\upsilon}_N\ ]$ represents the noise subspace, $\boldsymbol{\Lambda}_s=diag\{\lambda_1,\lambda_2,\cdots,\lambda_{M+1}\}$ are the eigenvalues of the SPI, and $\mathbf{\Lambda}_n = diag\{\lambda_{M+2}, \lambda_{M+3}, \cdots, \lambda_N\}$ are the eigenvalues of noise. As we know, the eigenvectors in $\mathbf{U}_s$ and the ASVs of the SPI lie in the same subspace. What's more, the mismatch between $\mathbf{a}(\theta_0)$ and $\bar{\mathbf{a}}_0$ is not too large in fact. We project the presumed desired signal ASV $\bar{\mathbf{a}}_0$ onto the eigenvectors to get $p(i)$, i=1,2,$\cdots$,N. $p(i)$ can be expressed as:

$$p(i) = \left|\mathbf{\upsilon}_i^H \bar{\mathbf{a}}_0\right|^2, \quad i=1,2,\cdots,N. \tag{7}$$

The projections $p(i)$ can be arranged in descending order, as $p_{[N]} \geq p_{[N-1]} \geq \cdots \geq p_{[1]}$. Meanwhile, the corresponding eigenvectors can be arranged as $\mathbf{\upsilon}_{[N]}$,$\mathbf{\upsilon}_{[N\text{-}1]}$,$\cdots$,$\mathbf{\upsilon}_{[1]}$, and the corresponding eigenvalues can be arranged as $\lambda_{[N]}, \lambda_{[N-1]}, \cdots, \lambda_{[1]}$. Note that the eigenvalues $\lambda_{[N]}, \lambda_{[N-1]}, \cdots, \lambda_{[1]}$ are the same as (6), but they have been reordered according to $p_{[N]} \geq p_{[N-1]} \geq \cdots \geq p_{[1]}$.

As well known, when $\mathbf{\upsilon}_i$ is the eigenvector corresponds to the desired signal, the maximum of the projections $p_{[N]}$ can be obtained [8]. That is to say the eigenvector corresponding to the desired signal can be identified by $p_{[N]}$. Easy to see that $\mathbf{\upsilon}_{[N]}$ and $\lambda_{[N]}$ are the corresponding eigenvector and eigenvalue, respectively. There is no doubt that $\lambda_{[1]}$ is the eigenvalue corresponding to the noise. Subsequently, we can use the parameter γ to reflect the input SNR directly [20], which can be expressed as:

$$\gamma = 10 * \log(\frac{\lambda_{[N]}}{\lambda_{[1]}}). \tag{8}$$

When the input SNR is very small, $\gamma < 0$. On the contrary, the large value of γ can be obtained in high SNR environment. Here, we give an example to discuss the relationship between γ and the input SNR.

We consider a ULA of $N = 10$ antennas spaced at a half wavelength distance. Additive noise is modelled as independent complex Gaussian noise with zero mean and unit variance. Two independent interferences are from the directions of 30° and -50°, respectively. The interference-to-noise ratios (INRs) of the interferences are 30 dB. The desired signal is assumed from the direction of 3°. The random DOA estimation mismatch is distributed in $[-5^\circ, 5^\circ]$. The number of snapshots is K=100. 200 repetitions are executed to obtain each simulated point.

Figure 1 shows the values of γ versus input SNRs. We can observe form Fig. 1 that the values of γ are quite consistent with the optimal SINR when the input SNR> -10 dB. The relationship between γ and SNR is almost linear. We can say the parameter γ can reflect the input SNR exactly as long as SNR is large enough. When SNR$\leq$-10 dB, γ failed to reflect the input SNR due to the overestimated of signal subspace. Luckily, the traditional diagonal loading method can achieve the same performance as other methods in low SNRs. Naturally, we prefer to use the traditional diagonal loading in low SNR due to its low complexity.

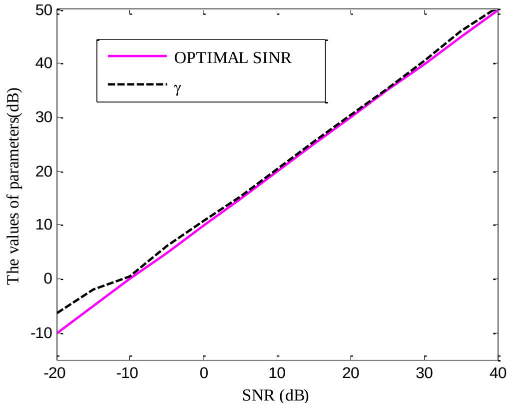


Fig. 1. Values of γ versus the SNRs.

Thus, the proposed diagonal method can be expressed as:

$$\tilde{\mathbf{R}} = \begin{cases} \hat{\mathbf{R}} + \dfrac{\lambda_{[N]}}{\lambda_{[1]}}\mathbf{I} & \gamma > 0 \\ \hat{\mathbf{R}} + \eta\mathbf{I} & \gamma \leq 0 \end{cases}, \tag{9}$$

where $\mathbf{I}$ is an identity matrix, and η is a fixed value. We usually set η as twice as the noise power. We can see that a new covariance matrix can be obtained by using the proposed diagonal method. The diagonal loading factor in (9) is changed according to the SNR. When $\gamma > 0$, the diagonal loading factor is positive to the SNR. The large diagonal loading factor can increase the noise power and the proportion of the desired signal in the covariance matrix is reduced. As a result, the effect of the desired signal is reduced in high SNRs. When $\gamma \leq 0$, the diagonal loading factor is set to be a fixed value.

B. Desired signal steering vector estimation

In practical applications, the presumed ASV of the desired signal may not precisely, and the mismatches cause significant performance degradation. Using the similar idea of [18], we can obtain accurate ASV of the desired signal. We assume that the Θ is the angular sector in which the desired signal is located. Define the correlation matrix of the steering vector:

$$\mathbf{C} = \int_\Theta \mathbf{a}(\theta)\mathbf{a}^H(\theta)d\theta, \tag{10}$$

where $\mathbf{a}(\theta)$ is the ASV associated with a presumed direction θ located in Θ. The matrix can be decomposed as:

$$\mathbf{C}=\sum_{n=1}^{N}\mu_n\bar{\mathbf{u}}_n\bar{\mathbf{u}}_n^H, \tag{11}$$

where μ_n are the eigenvalues of $\mathbf{C}$ in descending, and $\bar{\mathbf{u}}_n$ are the corresponding eigenvectors. The Q largest eigenvalues can be extracted, and the corresponding eigenvectors can be expressed as $\bar{\mathbf{u}}_n$, n=1,2,$\cdots$,Q. Then a column orthogonal matrix $\mathbf{U}$ can be constructed as $\mathbf{U}=[\bar{\mathbf{u}}_1,\bar{\mathbf{u}}_2,\cdots,\bar{\mathbf{u}}_Q]$. As we all know, the Q eigenvalues can contain most of the energy of $\mathbf{C}$ as long as Q is large enough. That is to say any ASV whose direction located in Θ can be expressed as a linear combination of columns of $\mathbf{U}$. So the actual ASV of the desired signal can be estimated as:

$$\tilde{\mathbf{a}}_0=\mathbf{U}\mathbf{r}, \tag{12}$$

where $\mathbf{r}$ is defined as a rotating vector. In order to obtain the vector $\mathbf{r}$, we can maximize the output power of the desired signal. Taking into the norm constraint, the optimization problem can expressed as:

$$\begin{aligned}&\min_{\mathbf{r}} && \mathbf{r}^H\hat{\mathbf{R}}_U\mathbf{r}\\ &\text{subject to} && \mathbf{r}^H\mathbf{r}=N,\end{aligned} \tag{13}$$

where $\hat{\mathbf{R}}_U=\mathbf{U}^H\hat{\mathbf{R}}^{-1}\mathbf{U}$. The constraint $\mathbf{r}^H\mathbf{r}=N$ is aims to avoid scaling ambiguity. The problem (13) can be solved by Lagrange multiplier methodology, and the cost function is given by:

$$L(\mathbf{r},\mathbf{U})=\frac{1}{2}\mathbf{r}^H\hat{\mathbf{R}}_U\mathbf{r}+\hat{\mu}(N-\mathbf{r}^H\mathbf{r}), \tag{14}$$

where $\hat{\mu}$ denotes the Lagrange multiplier. Computing the gradient of $L(\mathbf{r},\mathbf{U})$ with respect to $\mathbf{r}$, and then setting it equal to zero. We can get:

$$\hat{\mathbf{R}}_U\mathbf{r}=\hat{\mu}\mathbf{r}. \tag{15}$$

It is easy to understand that $\mathbf{r}$ can be regarded as the eigenvector of $\hat{\mathbf{R}}_U$ which corresponding to the smallest eigenvalue. $\mathbf{r}$ can be obtained as follows:

$$\mathbf{r}=\bar{\nu}[\hat{\mathbf{R}}_U], \tag{16}$$

where $\bar{\nu}[\bullet]$ is the operator that extracts the eigenvector corresponding to the smallest eigenvalue. Then, the estimated ASV of the desired signal can be obtained by substituting this solution into (12). Taking into consideration of the norm constraint, we obtain:

$$\tilde{\mathbf{a}}_0=\frac{\sqrt{N}}{\|\mathbf{r}\|}\mathbf{U}\mathbf{r}\cdot \tag{17}$$

Once the $\tilde{\mathbf{a}}_0$ is obtained, according to (5), the weighting vector can be achieved. It is worth noting that this method suffers performance degeneration in low SNR due to the Eigen-decompose operator. To address this problem, we can utilize the presumed ASV $\bar{\mathbf{a}}_0$ for $\gamma\le 0$. Then the weighting vector can be expressed as:

$$\tilde{\mathbf{w}}=\begin{cases}\dfrac{(\hat{\mathbf{R}}+\dfrac{\lambda_{[N]}}{\lambda_{[1]}}\mathbf{I})^{-1}\tilde{\mathbf{a}}_0}{\tilde{\mathbf{a}}_0^H(\hat{\mathbf{R}}+\dfrac{\lambda_{[N]}}{\lambda_{[1]}}\mathbf{I})^{-1}\tilde{\mathbf{a}}_0} & \gamma>0\\[2ex] \dfrac{(\hat{\mathbf{R}}+\eta\mathbf{I})^{-1}\bar{\mathbf{a}}_0}{\bar{\mathbf{a}}_0^H(\hat{\mathbf{R}}+\eta\mathbf{I})^{-1}\bar{\mathbf{a}}_0} & \gamma\le 0\end{cases}. \tag{18}$$

It can be seen from (18) that the diagonal loading factor and ASV of the desired signal in the proposed method can be adjusted by introducing the parameter γ. The expression of the weight vector is an attractive closed-form without any iteration process. As a result, the complexity is significantly reduced, especially in low SNR.

C. Complexity analysis

The main computational complexity of the proposed method is the Eigen-decomposition operation and matrix inverse operation. Its overall computational complexity is $O(N^3)$. The RCB algorithm also needs Eigen-decomposition operation, which has a complexity of $O(N^3)$. The worst-case beamforming and the beamformer of [14] (SDP-RAB) have at least the complexity of $O(N^{3.5})$. The beamformer of [15] (IPN-RAB) has a complexity of $O(SN^2)$, where S is the number of sampling points in the area eliminating desired signal. Typically, $S\ge N$. The computational complexity of the beamformer in [19] is $O(N^3)$.

IV. SIMULATION

In this section, the basic simulation conditions are the same as above unless otherwise is specified. The possible angular sector of the SOI is set to $\Theta=[-5^\circ,11^\circ]$, so the complement sector is $\bar{\Theta}=[-90^\circ,-5^\circ)\bigcup(11^\circ,90^\circ]$. We set $Q=2$.The proposed beamformer is investigated and compared with the diagonally loaded SMI (LSMI) [11], RCB [12], SDP-RAB[14], IPN-RAB[15], the beamformer of [18] and the beamformer of [19]. The optimal parameter $\varepsilon=0.3N$ is used for the RCB, while the diagonal loading factor of LSMI is selected as twice as the noise power. CVX software is used to solve these convex optimization problems [21].

A. Simulation Example 1: The ASV was known exactly

In this example, we consider the situation that the actual ASVs are exactly known. That means the presumed ASVs and array manifold knowledge are consistent with the actual.

Figure 2 shows the output SINR of the tested beamformers versus input SNR for K=100. It can be seen from the Fig. 2 that the IPN-RAB enjoys the best performance with a high complexity when the ASV is known exactly. The proposed method outperforms the

RCB, beamformer of [18] and beamformer of [19], which all have the same computational complexities. As it can be observed that the beamformer of [18] suffers significantly performance degradation due to the subspace swap phenomenon in low SNRs. The proposed method overcomes this problem by using the traditional diagonal loading method in low SNRs without eigenvalue decomposition of the sample covariance matrix. The output SINR of the proposed beamformer exceeds beamformer of [19] 3.94 dB when the input SNR is 30 dB. Figure 3 corresponds to the output SINR performance versus the number of the snapshots with input SNR=20 dB. It can be observed that the proposed beamformer offers good performance with low computational complexity. The inaccurate eigenvalue decomposition of the covariance matrix in (6) and (16) is the main reason why the proposed method suffers little performance degradation when the number of snapshots is very small.

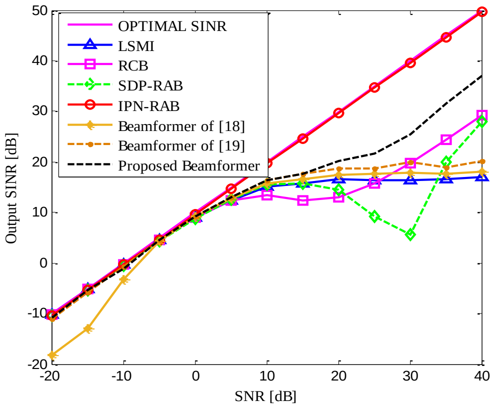


Fig. 2. Output SINR versus the input SNR when the ASVs are exactly known.

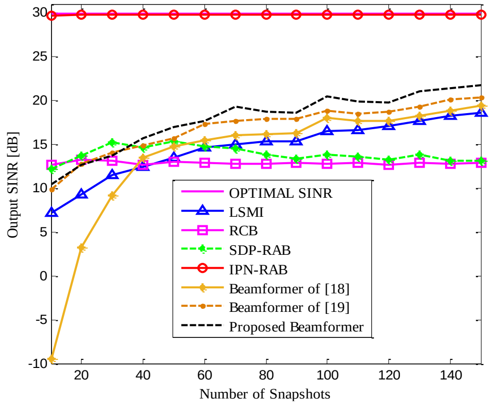


Fig. 3. Output SINR versus the number of snapshots when the ASVs are exactly known.

B. Simulation Example 2: Signal look direction mismatch

In this example, we consider the situation that the random desired signal direction error is occurred. The random DOA estimation mismatch is distributed in $[-5^\circ,5^\circ]$ for each simulation run. So, the actual desired signal DOA is uniformly distributed in $[-2^\circ,8^\circ]$. It is worth noting that the DOA mismatch is changed in each run but remain fixed in each snapshot.

Figure 4 displays the mean output SINR of the tested methods versus the SNR for K=100. As it can be observed, the LSMI suffers significant performance degradation in high SNRs. The performance of the proposed beamformer is only next to the IPN-RAB method. The proposed method provides better performance compared with the LSMI, RCB, beamformer of [18] and beamformer of [19]. In Fig. 5, the output SINR is shown with respect to the number of snapshots for SNR=20 dB. Similar to the previous example, the proposed beamformer has good performance except that the number of the snapshots is very small.

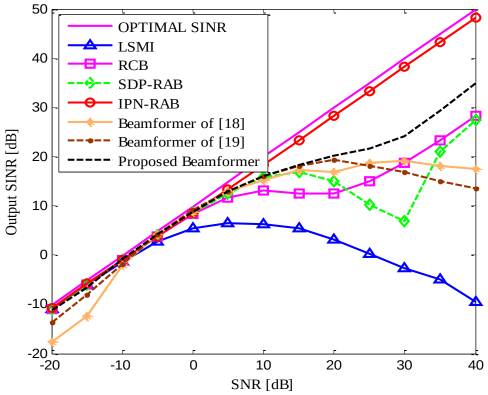


Fig. 4. Output SINR versus the input SNR in the case of look direction mismatch.

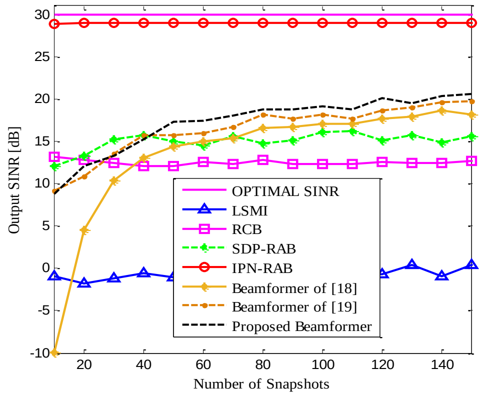


Fig.5. Output SINR versus the number of snapshots in the case of look direction mismatch.

C. Simulation Example 3: Desired ASV mismatch due to wavefront distortion

In this simulation, we consider the ASV of the desired signal is distorted by the effects of wave propagation due to the inhomogeneous medium [14]. In particular, the independent-increment phase distortions are accumulated from the components of the presumed ASV. Assuming that the phase increments are fixed in each simulation runs and are independently chosen from a Gaussian random generator with zero mean and standard deviation 0.04. Figure 6 shows the output SINR of the beamformers versus input SNR for K=100. Figure 7 shows the output SINR of the beamformers versus the number of snapshots for fixed input SNR=20 dB.

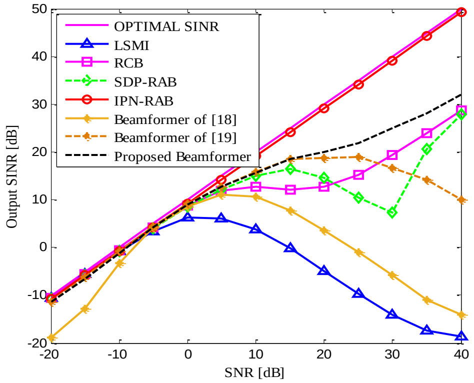


Fig. 6. Output SINRs versus input SNR in the case of wavefront distortion.

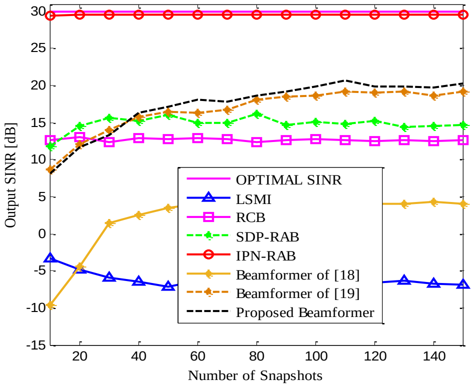


Fig. 7. Output SINRs versus the number of snapshots in the case of wavefront distortion.

It can be observed from the Fig. 6 that the beamformer of [18] lacks robustness against the wavefront distortion, and the SDP-RAB method suffers performance degradation when the $10\text{dB} \le \text{SNR} \le 30\text{dB}$. The IPN-RAB method outperforms the proposed method at high SNRs, but significantly more complicated. The proposed method is superior to the beamformer of [19] without increasing the calculation load. In particular, we can find from Fig. 7 that the output SINR of the proposed method exceeds the beamformer of [19] 1.15 dB when K=150. From the discussions mentioned above, the performance of our proposed method has little deterioration in small number of snapshots due to the inaccurate estimation of the SOI component.

D. Simulation Example 4: Effect of the error in the knowledge of the array geometry

In this simulation, the effect of the element position errors on the performance of the tested beamformers is investigated. We assume the difference between the presumed and actual positions of each element is modelled as a uniform random variable distributed in the interval $[-0.075\lambda, 0.075\lambda]$, where λ represents the wavelength. The actual DOA of SOI is 5°, and hence, the DOA mismatch is 2°.

Figure 8 shows the output SINR of the beamformers versus input SNR for K=100. It can be seen from the Fig. 8 that the IPN-RAB suffers serious performance degradation in low SNRs due to the inaccurate array manifold information. The proposed beamformer provides strong robustness in the presence of the element positions errors both at low and high SNRs. Figure 9 displays the output SINR performance of all the tested beamformer versus the number of training snapshots for SNR=20 dB. As is shown in the picture, we can see clearly that the proposed method enjoys the best performance when $K \geq 60$.

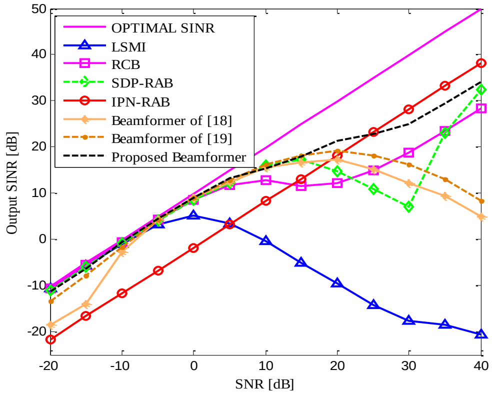


Fig. 8. Output SINR of beamformers versus input SNR for the case of element position errors.

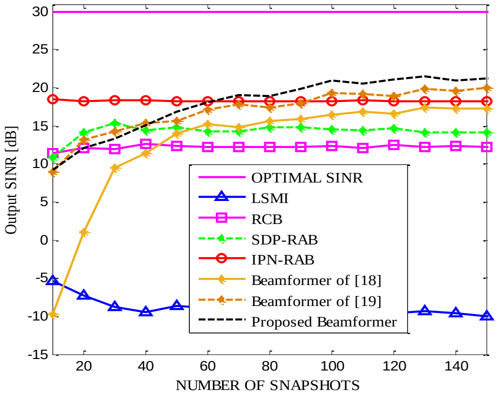


Fig. 9. Output SINR versus the number of snapshots for the case of element position errors.

E. Simulation Example 5: Mismatch due to arbitrary ASV errors

In this simulation, we study the performance of the proposed beamformer when arbitrary ASV errors are considered. Here, the ASV mismatch is comprehensive and arbitrary-type, which may be caused by direction errors, calibration errors, gain and phase perturbations, and so on. The actual ASVs of can be modelled as [17]:

$$\bar{\mathbf{a}}(\theta)=\mathbf{a}(\theta)+\hat{\mathbf{e}}, \tag{19}$$

where $\bar{\mathbf{a}}(\theta)$ denoted the presumed ASVs, $\hat{\mathbf{e}}$ is a zero-mean complex random vector with the variance σ_e^2. In this example, all the array imperfections are generated as Gaussian variables with the given variance, $\sigma_e^2 = 2$.

The output SINR of the beamformers versus input SNR for K=100 is displayed in Fig. 10. We can notice that the proposed method can improve the output SINR of the beamformer efficiently, and fit to be used in complex environment. This means the proposed beamformer is effective in the presence of the arbitrary ASV errors. Figure 11 displays the output SINR performance of all the tested beamformer versus the number of training snapshots for SNR=20 dB. Obviously, the proposed beamformer has the best output SINR among all the tested beamformers when $K \geq 60$.

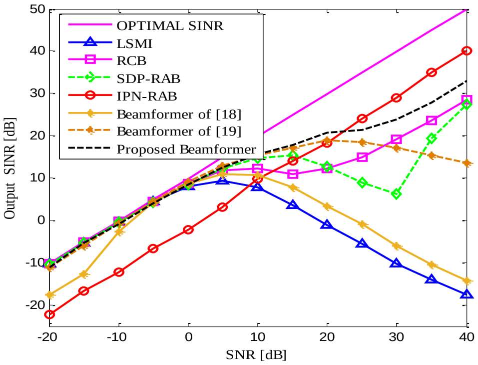


Fig. 10. Output SINR of beamformers versus input SNR with arbitrary ASV errors.

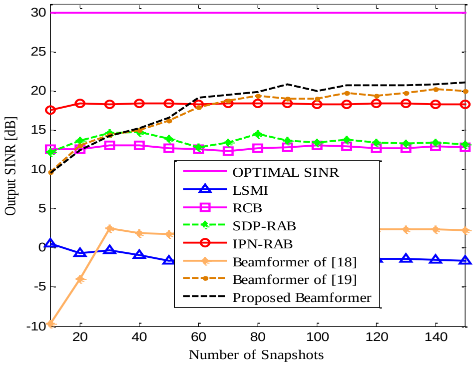


Fig. 11. Output SINR of beamformers versus snapshots with arbitrary ASV errors.

V. CONCLUSION

In this paper, a robust adaptive beamforming method with low complexity has been proposed and its performance has been investigated. The proposed beamformer is realized based on variable diagonal loading method, and the diagonal loading factor is selected according to the input SNR. The expression of the weighting vector in a closed-form has been provided. The proposed method has a low complexity and possesses good operability and excellent performance. The simulation results demonstrated that the proposed beamformer can provide superior performance against unknown arbitrary-type mismatches compared to the existing popular methods.

REFERENCES

[1] H. L. Van Trees, *Optimum Array Processing*. Hoboken, NJ, USA: Wiley, 2002.

[2] R. L. Haupt, "Adaptive Arrays," *Applied Computational Electromagnetics Society Journal*, vol. 24, no. 6, pp. 541-549, Dec. 2009.

[3] W. Li, X. Mao, W. Yu, and C. Yue, "Robust adaptive array beamforming based on modified norm constraint algorithm," *Applied Computational Electromagnetics Society Journal*, vol. 29, no. 12, pp. 1060-1066, Dec. 2014.

[4] W. Li, Y. Li, and W. Yu, "On adaptive beamforming for coherent interference suppression via virtual antenna array," *Progress In Electromagnetics Research*, vol. 125, pp. 165-184, 2012.

[5] S. A. Vorobyov, "Principles of minimum variance robust adaptive beamforming design," *Signal Process*, vol. 93, no. 12, pp. 3264-3277, 2013.

[6] X. Mao, W. Li, Y. Li, and Z. Zhai, "Robust adaptive beamforming against signal steering vector mismatching and jammer motion," *International Journal of Antennas and Propagation*, Article ID 780296, 12 pages, 2015.

[7] Z. L. Yu, M. H. Er, and W. Ser, "A novel adaptive beamformer based on semidefinite programming (SDP) with magnitude response constraints," *IEEE Trans. Antennas Propagat.*, vol. 56, no. 7, pp. 1297-1307, May 2008.

[8] W. Jia, W. Jin, S. Zhou, and M. Yao, "Robust adaptive beamforming based on a new steering vector estimation algorithm," *Signal Process*, vol. 93, no. 9, pp. 2539-2542, Sep. 2013.

[9] B. Liao, K. M. Tsui, and S. C. Chan, "Robust beamforming with magnitude response constraints using iterative second-order cone programming," *IEEE Trans. Antennas Propag.*, vol. 59, no. 9, pp. 3477-3482, Sep. 2011.

[10] D. Xu, R. He, and F. Shen, "Robust beamforming with magnitude response constraints and conjugate symmetric constraint," *IEEE Commun. Let.*, vol. 17, no. 3, pp. 561-564, May 2013.

[11] B. D. Carlson, “Covariance matrix estimation errors and diagonal loading in adaptive arrays,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. 24, no. 4, pp. 397-401, July 1988.

[12] J. Li, P. Stoica, and Z. Wang, “On robust Capon beamforming and diagonal loading,” *IEEE Trans. Signal Process*, vol. 51, no. 7, pp. 1702-1715, July 2003.

[13] S. A. Vorobyov, A. B. Gershman, and Z.-Q. Luo, “Robust adaptive beamforming using worst-case performance optimization: A solution to the signal mismatch problem,” *IEEE Trans. Signal Process*, vol. 51, no. 2, pp. 313-324, Feb. 2003.

[14] A. Khabbazibasmenj, S. A. Vorobyov, and A. Hassanien, “Robust adaptive beamforming based on steering vector estimation with as little as possible prior information,” *IEEE Trans. Signal Process*, vol. 60, no. 6, pp. 2974-2987, June 2012.

[15] Y. Gu and A. Leshem, “Robust adaptive beamforming based on interference covariance matrix reconstruction and steering vector estimation,” *IEEE Trans. Signal Process*, vol. 60, no. 7, pp. 3881-3885, July 2012.

[16] Y. Kai, Z. Zhao, and Q.H. Liu, “Robust adaptive beamforming against array calibration errors,” *Progress In Electromagnetics Research*, vol. 140, pp. 341-351, 2013.

[17] L. Huang, J. Zhang, X. Xu, and Z. Ye, “Robust adaptive beamforming with a novel interference-plus-noise covariance matrix reconstruction method,” *IEEE Trans. Signal Process*, vol. 63, no. 7, pp. 1643-1650, Apr. 2015.

[18] Z. Wei, J. Wang, and S. Wu, “Robust Capon beamforming against large DOA mismatch,” *Signal Process*, vol. 93, no. 4, pp. 804-810, 2013.

[19] J. Zhuang, Q. Ye, Q. Tan, and A. H. Ali, “Low-complexity variable loading for robust adaptive beamforming,” *Electronics Letters*, vol. 52, no. 5, pp. 338-340, Mar. 2016.

[20] W. Li, X. Mao, Z. Z, and L. Li, “High performance robust adaptive beamforming in the presence of array imperfections,” *International Journal of Antennas and Propagation*, Article ID 3743509, 12 pages, 2016.

[21] M. Grant, S. Boyd, and Y. Y. Ye, CVX: Matlab Software for Disciplined Convex Programming. June 2014. [Online]. Available: http://cvxr.com/cvx/

Optical Cloak Design Exploiting Efficient Anisotropic Adjoint Sensitivity Analysis

L. S. Kalantari and M. H. Bakr

Department of Electrical and Computer Engineering
McMaster University, Hamilton, Ontario, L8S 4K1, Canada
kalantl@mcmaster.ca, mbakr@mail.ece.mcmaster.ca

Abstract ─ We propose in this work a novel optimization-based wideband invisibility cloaking approach at optical frequencies. We exploit the memory efficient anisotropic adjoint variable method (AVM) to significantly accelerate the sensitivity analysis with respect to the large number of parameters. Minimizing the cloaking objective function involves the gradients estimation with respect to a massive number of parameters at every iteration for time-intensive electromagnetic simulations. The AVM evaluates the required gradients over one thousand time faster than conventional methods. The significant reduction in the computational cost enables wideband optimization-based cloak design at optical region.

Index Terms ─ Adjoint variable method, computationally intensive structures, gradient-based optimization, invisibility cloaks, massive parameters, sensitivity analysis acceleration.

I. INTRODUCTION

Recently, hardware acceleration techniques and development of novel computational methods were utilized to efficiently model electrically large structures [1] – [4]. These methods enable the optimization-based design of time intensive high frequency electromagnetics structures. Several optimization techniques require estimating the gradient of an objective function with respect to all the design parameters [5]. These methods iteratively solve the optimization problem using the estimated gradients.

Some electromagnetics problems such as invisibility cloaking may have a massive number of optimizable parameters. Calculating the gradient of the objective function with respect to all those parameters may be prohibitive using conventional sensitivity analysis approaches. Gradient estimation is one of the challenges of big data analysis [6]. In conventional sensitivity analysis methods such as central finite differences (CFD), the number of required simulations to estimate the gradients scales linearly with the number of parameters. For instance, for a problem with N number of optimizable parameters, $2N$ simulations are required to find the gradient for only one iteration. This is cumbersome for huge N, particularly when the considered electromagnetic simulation is time intensive. Adjoint sensitivity analysis approaches [7] find the sensitivities with respect to all optimizable parameters regardless of their number using at most one extra simulation. Using the AVM approach, optimization-based design of electrically large structures with massive parameters, such as invisibility cloak design at optical frequencies, becomes feasible.

Invisibility cloaking have been studied through both analytical [8] – [11], scattering cancellation (SC) [8], transformation optics (TO) [9] – [11], and optimization-based [12] – [16] design approaches. Optimization methods are utilized to design multi-layered covers for structures with geometric symmetry at microwave region [12] – [15]. The resultant cloaks are usually narrowband. For wider bandwidth, the optimization process is repeated for every frequency in the band of interest [14], [15]. Optimization-based cloaking utilizes either global optimization approaches or gradient-based approaches. Global methods require too many simulations. Gradient-based methods require a possibly large number of sensitivities per iteration. This is cumbersome at optical frequencies where smaller mesh sizes make the simulations even more time intensive. These constraints limit optimization-based cloaking to narrowband structures [12], [13], [16] at microwave region [12] – [15].

Instead of following the traditional layer-based cloaks, we recently showed the strength of voxel-by-voxel cloak design in the microwave frequency regime [17], [18]. However, to solve similar problems at optical frequencies, the number of required gradients per iteration is larger than [17], [18]. Also, each simulation takes much longer time and the efficient memory usage become more critical. Thus, at optical frequencies, wideband optimization-based cloaking using classical approaches are computationally formidable.

In this work, exploiting an AVM method, we significantly accelerate the design of optical cloaks. The gradient of the cloaking objective function with respect to the massive number of parameters is estimated using

Submitted On: December 28, 2016
Accepted On: January 28, 2017

only one extra simulation. Each voxel of the cloak shell is allowed to have its independent constitutive parameters. We optimally design the generally anisotropic permittivity tensor elements of each voxel to provide sufficient invisibility for the whole structure at the two observation ports.

The manuscript is structured as follows: the formulation is presented in Section II. A wideband version of the structure studied in [11] is cloaked in Section III. We suggest a non-magnetic cloak for both transverse electric (TE) and transverse magnetic (TM) illuminations. The numerical results illustrate the efficiency of the method in wideband all-dielectrics cloak design within the optical frequency regime.

II. FORMULATION

A. Cloaking problem formulation

We aim to match the electric field of an empty domain $\boldsymbol{E}^{inc}(\boldsymbol{r},t)$ (Fig. 1 (a)), to the electric field resulting from the optimally cloaked object, $\boldsymbol{E}(\boldsymbol{r},t,\boldsymbol{p}^*)$, (Fig. 1 (b)), over the input and output spatial observation ports (Ω_{in} and Ω_{out}) and over a time domain window (T_m). Here, $\boldsymbol{r}$ is the position vector and t is time. The vector $\boldsymbol{p}$ includes all the constitutive parameters of the utilized cloak voxels. $\boldsymbol{p}^*$ is the optimal cloak constitutive parameters vector that makes $\boldsymbol{E}(\boldsymbol{r},t,\boldsymbol{p}^*)$ match $\boldsymbol{E}^{inc}(\boldsymbol{r},t)$ for all time. The following objective function measures the object invisibility:

$$\begin{aligned} F &= \int_0^{T_m}\int_{\Omega_{in}} \left\| \boldsymbol{E}(\boldsymbol{r},t,\boldsymbol{p}) - \boldsymbol{E}^{inc}(\boldsymbol{r},t)\right\|_2^2 d\Omega_{in}dt \\ &+ \int_0^{T_m}\int_{\Omega_{out}} \left\| \boldsymbol{E}(\boldsymbol{r},t,\boldsymbol{p}) - \boldsymbol{E}^{inc}(\boldsymbol{r},t)\right\|_2^2 d\Omega_{out}dt = \int_0^{T_m} G dt. \end{aligned} \quad (1)$$

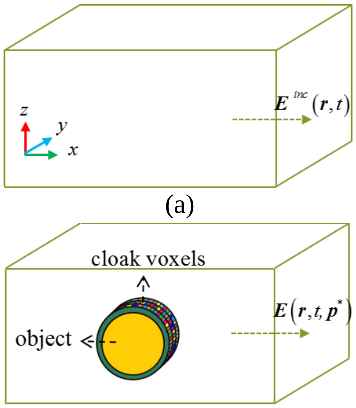
Here, G is the corresponding kernel of F. Ideally, if $\boldsymbol{E}(\boldsymbol{r},t,\boldsymbol{p})$ is matched to $\boldsymbol{E}^{inc}(\boldsymbol{r},t)$ over all the space and time samples, (1) would have the ideal value of zero. However, (1) can be minimized by adjusting the vector $\boldsymbol{p}$ to reach a minimum that is close to the ideal value. We utilize an all-dielectric cloak, as it provides low loss and wider band response. Such a cloak can be fabricated at optical frequencies [10], [11]. The cloak is divided into a number of voxels. Each voxel has an independent permittivity tensor, $\boldsymbol{\varepsilon}_{rj}=\mathrm{diag}(\varepsilon_{rxxj}\ \varepsilon_{ryyj}\ \varepsilon_{rzzj})$, $j=1,2,\ldots, N_v$, where N_v is the number of cloak voxels. All tensors are non-dispersive with diagonal elements greater than one. The vector of optimizable parameters is given by:

$$\boldsymbol{p} = \left[\boldsymbol{p}_1^T\ \boldsymbol{p}_2^T\ \ldots\ \boldsymbol{p}_{N_v}^T\right]^T \in R^{N\times 1}. \quad (2)$$

Here, $\boldsymbol{p}_j=[\varepsilon_{rxxj}\ \varepsilon_{ryyj}\ \varepsilon_{rzzj}]^T$, $j=1,2,\ldots, N_v$ and N is the size of $\boldsymbol{p}$. To minimize (1), we utilize the steepest descent method with line search [5]. It requires estimating the gradient of (1) with respect to the N elements of $\boldsymbol{p}$ at each iteration:

$$\frac{\partial F}{\partial \boldsymbol{p}} = \left[\left(\frac{\partial F}{\partial \boldsymbol{p}_1}\right)^T \left(\frac{\partial F}{\partial \boldsymbol{p}_2}\right)^T \cdots \left(\frac{\partial F}{\partial \boldsymbol{p}_{Nv}}\right)^T\right]^T \in R^{N\times 1}. \quad (3)$$

As the number of cloak voxels may be large, estimating (3) using conventional CFD approaches is prohibitive.



(a)

(b)

Fig. 1. The computational domain where, (a) the object is absent, and (b) a cloak shell composed of a number of optimally designed voxels covers the object.

B. AVM gradient estimation

To overcome the massive computational cost of estimating gradient vector (3), we developed a fast memory efficient anisotropic AVM algorithm. In optical frequencies, due to smaller mesh size, the number of optimizable parameters are inherently massive. Utilizing memory efficient AVM is thus mandatory as compared to [17], [18]. It estimates all gradient elements of (3) for an arbitrary wideband F, using only one extra simulation [7]. The cost of the AVM approach does not scale with the number of parameters as in classical techniques. This makes AVM well suited for applications with a relatively large number of parameters that appear in big data problems [4]. We utilize a transmission line modeling (TLM)-based AVM approach. TLM replaces the computational domain by a network of transmission lines. The electric and magnetic field components are mapped to circuit voltages and currents, v_u and i_u, $u=x,y,z$. Also, for each node of the computational domain, reflected voltages and currents, $\boldsymbol{v}^r$ and $-\boldsymbol{i}^r$, that correspond to the reflected electric and magnetic fields are defined. The total voltages and current and extra storages vectors in the original system are updated at each iteration using:

$$\begin{bmatrix}\boldsymbol{v}\\ \boldsymbol{i}\\ \boldsymbol{Q}_e\\ \boldsymbol{Q}_m\end{bmatrix} = \begin{bmatrix} 2\boldsymbol{T}_e & \boldsymbol{0} & \boldsymbol{T}_e & \boldsymbol{0}\\ \boldsymbol{0} & 0.5\boldsymbol{I} & \boldsymbol{0} & 0.25\boldsymbol{I}\\ 2\boldsymbol{I}+2\boldsymbol{\kappa}_e\boldsymbol{T}_e & \boldsymbol{0} & \boldsymbol{\kappa}_e\boldsymbol{T}_e & \boldsymbol{0}\\ \boldsymbol{0} & 4\boldsymbol{I} & \boldsymbol{0} & \boldsymbol{I}\end{bmatrix} \begin{bmatrix}\boldsymbol{v}^r\\ -\boldsymbol{i}^r\\ z^{-1}\boldsymbol{Q}_e\\ z^{-1}\boldsymbol{Q}_m\end{bmatrix}. \quad (4)$$

Here, z is the time shift operator, $\boldsymbol{\kappa}_e=\boldsymbol{\sigma}_e+8\boldsymbol{I}-4\boldsymbol{\varepsilon}_r \in R^{3\times 1}$, $\boldsymbol{T}_e=(\boldsymbol{\sigma}_e+4\boldsymbol{\varepsilon}_r)^{-1}$, and $\boldsymbol{\sigma}_e$ is the electric conductivity tensor. $\boldsymbol{Q}_e$ and $\boldsymbol{Q}_m$ are the electric and magnetic extra storages vector, respectively. AVM finds the sensitivities of (1) with respect to the elements of (2), using the formula:

$$\frac{\partial F}{\partial p_i} = -\sum_j \sum_k \boldsymbol{\Lambda}_{i,j,k}^T \zeta_{i,j,k}, \quad i=1,2,\ldots,N. \tag{5}$$

The vectors $\zeta_{i,j,k,}=\begin{bmatrix} \boldsymbol{\zeta}^{v^T} & \boldsymbol{\zeta}^{i^T} & \boldsymbol{\zeta}^{Q_e{}^T} & \boldsymbol{\zeta}^{Q_m{}^T} \end{bmatrix}_{i,j,k}^T \in R^{12\times 1}$ represent the AVM original responses. They are evaluated during original simulation for the ith parameter, the jth cloak voxel, and the kth time-step. For all dielectric cloaks, the responses $\boldsymbol{\zeta}^i$ and $\boldsymbol{\zeta}^{Q_m}$ are zero. The responses $\boldsymbol{\zeta}^v$ and $\boldsymbol{\zeta}^{Q_e}$ are given by:

$$\boldsymbol{\zeta}^v = 2\frac{\partial \boldsymbol{T}_e}{\partial p_i}\boldsymbol{v}^r + \frac{\partial \boldsymbol{T}_e}{\partial p_i} z^{-1}\boldsymbol{Q}_e, \tag{6}$$

$$\boldsymbol{\zeta}^{Q_e} = \frac{\partial \boldsymbol{\kappa}_e}{\partial p_i}\boldsymbol{v} + \boldsymbol{\kappa}_e \boldsymbol{\zeta}^v, \tag{7}$$

the indices i, j, k are dropped for simplicity. The vectors $\boldsymbol{\Lambda}_{i,j,k} = \begin{bmatrix} \boldsymbol{v}^{r,\Lambda^T} & -\boldsymbol{i}^{r,\Lambda^T} & \boldsymbol{B}_e^{\Lambda^T} & \boldsymbol{B}_m^{\Lambda^T} \end{bmatrix}_{i,j,k}^T \in R^{12\times 1}$ are the adjoint responses. They are evaluated during adjoint simulation. $\boldsymbol{v}^{r,\Lambda}$ and $-\boldsymbol{i}^{r,\Lambda}$ are the adjoint vector of reflected field voltages and currents. $\boldsymbol{B}_e^{\Lambda}$ and $\boldsymbol{B}_m^{\Lambda}$ are the adjoint electric and magnetic extra storages. They are updated backward in time using the adjoint simulation:

$$\begin{bmatrix} \boldsymbol{v}^{\Lambda} \\ \boldsymbol{i}^{\Lambda} \\ \boldsymbol{B}_e^{\Lambda} \\ \boldsymbol{B}_m^{\Lambda} \end{bmatrix} = \begin{bmatrix} 2\boldsymbol{T}_e^T & \boldsymbol{0} & (2\boldsymbol{I}+2\boldsymbol{\kappa}_e\boldsymbol{T}_e)^T & \boldsymbol{0} \\ \boldsymbol{0} & 0.5\boldsymbol{I} & \boldsymbol{0} & 4\boldsymbol{I} \\ \boldsymbol{T}_e^T & \boldsymbol{0} & (\boldsymbol{\kappa}_e\boldsymbol{T}_e)^T & \boldsymbol{0} \\ \boldsymbol{0} & 0.25\boldsymbol{I} & \boldsymbol{0} & \boldsymbol{I} \end{bmatrix} \begin{bmatrix} \boldsymbol{v}^{r,\Lambda} \\ -\boldsymbol{i}^{r,\Lambda} \\ z\boldsymbol{B}_e^{\Lambda} \\ z\boldsymbol{B}_m^{\Lambda} \end{bmatrix}. \tag{8}$$

In the adjoint simulation the vector of nodal excitation depend on the objective function [7]. In our cloaking problem G in (1), is composed of 2 surface integrals over Ω_{in} and Ω_{out}. Thus, in the adjoint simulation, 2 excitation ports, located at Ω_{in} and Ω_{out}, are applied. This is in contrast to the case of [17], [18] where the objective function matches the electric fields at one observation plane only.

C. Possible fabrication

The effective medium theory (EMT) and alternating layered structures [10], [11] have been applied to fabricate all dielectric cloaks at optical frequencies. Mixing different dielectrics with proper filling factors produce anisotropic materials with desired anisotropy tensor. For instance, consider a composite that is made of dielectrics ε_1 and ε_2 with filling factors f_1 and f_2. The EMT formula is [11]:

$$f_1\frac{\varepsilon_1-\varepsilon_{eff}}{\varepsilon_1+\kappa\varepsilon_{eff}} + f_2\frac{\varepsilon_2-\varepsilon_{eff}}{\varepsilon_2+\kappa\varepsilon_{eff}} = 0, \tag{9}$$

where ε_{eff} is the isotropic effective permittivity and κ is the screening factor [11]. The desired anisotropic permittivity tensor can be written as [10]:

$$\begin{bmatrix} \varepsilon_{\parallel} & 0 \\ 0 & \varepsilon_{\perp} \end{bmatrix}, \ \varepsilon_{\parallel} = f_1\varepsilon_1 + f_2\varepsilon_2, \ \varepsilon_{\perp} = \frac{\varepsilon_1\varepsilon_2}{f_2\varepsilon_1 + f_1\varepsilon_2}. \tag{10}$$

Here, index || and ⊥ show the cases with the electric field polarized parallel and perpendicular to the interfaces of the ε_1 and ε_2 layers. Thus, by properly selecting the dielectrics components, changing the filling factors and orientation of the composite components, the desired cloak parameters are generated using (9) and (10).

III. NUMERICAL EXAMPLE

We design a wideband dielectric cloak for the example presented in [11] for both the TE and TM illuminations. The cloaked object is a copper cylinder with a radius R_0=21.5Δl and electric conductivity $\sigma_e=5.6\times 10^7$ Sm^{-1}. The object is placed at the center of a waveguide with dimension $d_1\times d_2$, where d_1=758Δl and d_2=380Δl (see Fig. 2). Δl=21.09 nm is the TLM space step [7]. The object invisibility is measured by (1), where Ω_{in} and Ω_{out}, are located at 5Δl and 750Δl from the starting edge of the waveguide in x direction. For x and y boundaries, 70 layers of a perfectly matched layer are utilized. The excitation is a wideband Gaussian modulated sinusoidal (wavelength 632.8 nm and 127 nm bandwidth). The simulations are conducted on an Intel®Xeon® CPU×5670 @ 2.93 GHz (48.0 GB of RAM).

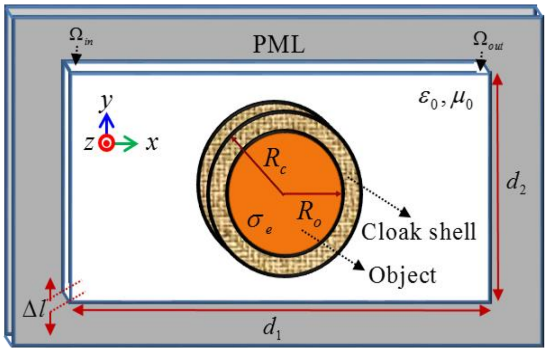


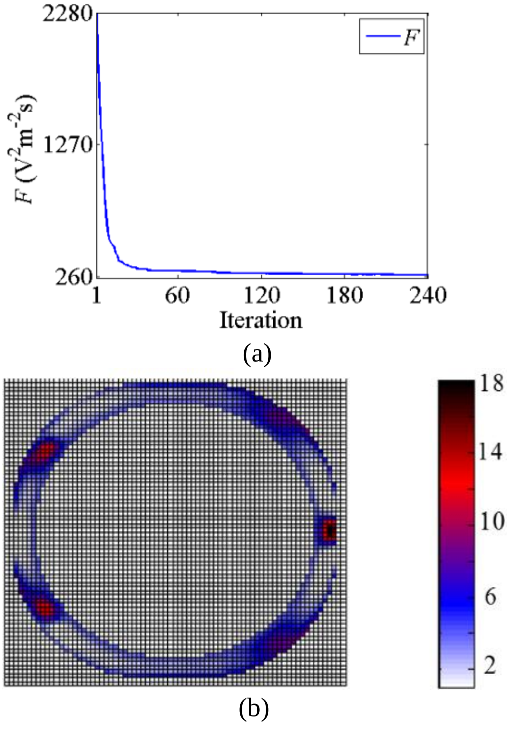
Fig. 2. The copper cylinder is covered by a cloak shell composed of a number of voxels.

A. TE illumination

We cover the object with a shell of thickness 5Δl (R_c=26.5Δl in Fig. 2). The shell is composed of 1320 voxels. To force TE illumination, E_z (electric field polarized along z-axis) incident field is applied combined with electric wall in z boundaries. Here, E_z is the only electric field in the domain and (1) is only sensitive to ε_{rzzj}, j=1,2,…,1320. The vector $\boldsymbol{p}$ has thus 1320 elements, ε_{rzzj}. Setting $\varepsilon_{rxxj}=\varepsilon_{ryyj}=\varepsilon_{rzzj}$, the cloak uses dielectric materials with isotropic permittivities greater than 1.

To minimize (1), the gradient vector (3), which is composed of 1320 components is estimated per iteration. The AVM calculates (3) using only 2 simulations in 118 minutes. The CFD requires 2640 simulations where each simulation takes 50.5 minutes. Therefore, in total it takes 92.5 days to estimate (3) for only one iteration using CFD! For this example at each iteration, AVM is 1129 times faster than CFD. Thus, minimizing (1) without AVM is simply prohibitive because of the large number of parameters and the intensive simulation time.

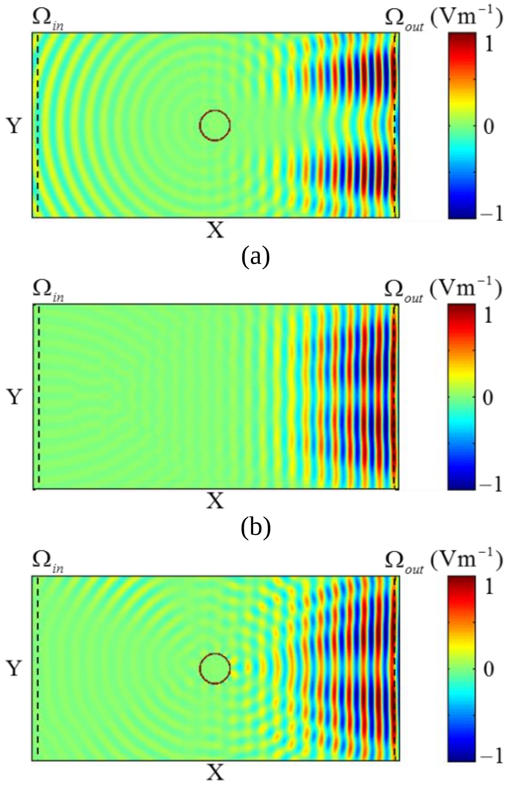
For the bare and cloaked object, the *F* has the values 1843 and 273 [$V^2m^{-2}s$], respectively. Figure 3 (a) illustrates the reduction of (1) per iteration. The optimal cloak parameters are within the range 1< ε_{rzzj} <18, $\forall j$. Their distributions are presented in Fig. 3 (b). The snapshots of time-dependent E_z distributions throughout the waveguide, at the same time sample, for the bare object, empty space, and clocked object are shown in Figs. 4 (a), (b), and (c), respectively. Applying the optimally designed cloak matches the E_z distribution of the cloaked object (Fig. 4 (c)) to that of the empty space (Fig. 4 (b)) at Ω_{in} and Ω_{out}. Also, for the bare object (Fig. 4 (a)), field distortions are observed at both Ω_{in} and Ω_{out}. However, those distortions have been significantly reduced for the cloaked cylinder (Fig. 4 (c)).



(a)

(b)

Fig. 3. TE illumination: (a) objective function per iteration, and (b) the distributions of ε_{rzzj} of the cloak voxels.



(a)

(b)

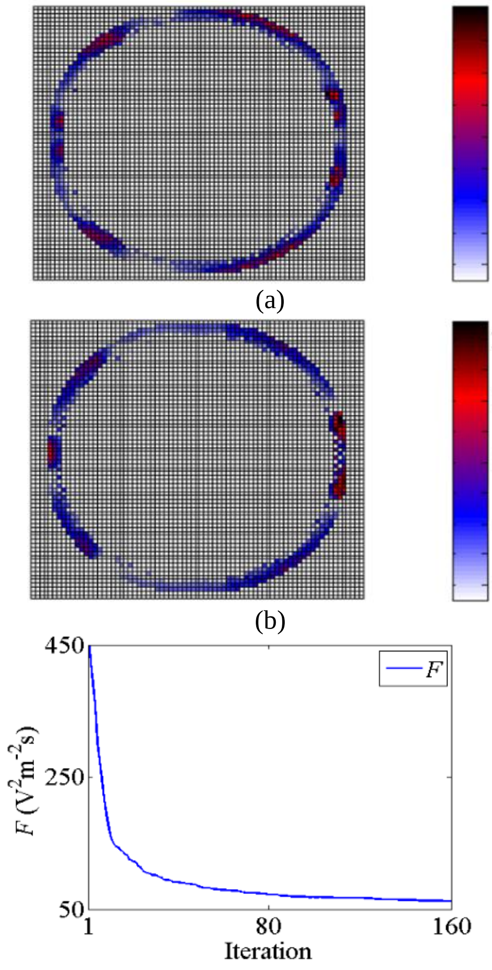
(c)

Fig. 4. TE illumination; the snapshots of E_z distributions throughout the waveguide for the: (a) bare object, (b) empty waveguide, and (c) cloaked object. Dashed lines locate Ω_{in} and Ω_{out}.

B. TM illumination

Here, the object is covered by a shell with thickness 3Δ*l* (R_c=24.5Δ*l*) which is composed of 768 voxels. To provide TM illumination, an incident magnetic field H_z (polarized along *z*-axis) is applied combined with magnetic wall boundary condition in *z*-direction. The electric fields in the domain are E_x and E_y. Applying the optimally designed cloak matches the E_x and E_y distributions of the cloaked object to that of the empty space. Therefore, (1) is only sensitive to ε_{rxxj} and ε_{ryyj} for *j*=1, 2,…,768. For simplicity, ε_{rzzj} are set equal to any of ε_{rxxj} or ε_{ryyj}. The vector ***p*** has 1536 parameters comprised of ε_{rxxj} and ε_{ryyj}, $\forall j$. Finally, the anisotropic cloak material is simplified to uniaxial crystals with diagonal tensor elements greater than one. In optimization process, the gradient vector (3), with 1536 components, is calculated per iteration. AVM approximate all those gradients by running only 2 simulations in 107 minutes. CFD would require 3072 simulations where each simulation takes 50.4 minutes. In total, using CFD, estimation of (3) takes 107.5 days per iteration! For this example AVM is 1449 times faster than CFD. Thus, minimizing (1) without AVM is simply not possible, due to the massive computation cost.

For the bare and cloaked object, *F* is 403.7 and 51.7 [$V^2m^{-2}s$], respectively. The optimal cloak parameters are within the range $1< \varepsilon_{rxxj} <18.2$ and $1< \varepsilon_{ryyj} <19.2$, $\forall j$. Figures 5 (a) and (b) show the optimal ε_{rxxj} and ε_{ryyj} distributions of the cloak voxels. Figure 5 (c) illustrates the reduction of (1) per iteration. Snapshots of the time-dependent H_z distributions throughout the waveguide for the bare object, empty space, and cloaked object, at a same time sample, are shown in Figs. 6 (a), (b), and (c), respectively. It illustrates that the cloaked object H_z distribution (Fig. 6 (c)) is well matched to that of the empty space (Fig. 6 (b)) at Ω_{in} and Ω_{out}. Also, the field distortions at Ω_{in} and Ω_{out} are reduced for the cloaked object (Fig. 6 (c)) as compared to the bare object (Fig. 6 (a)).

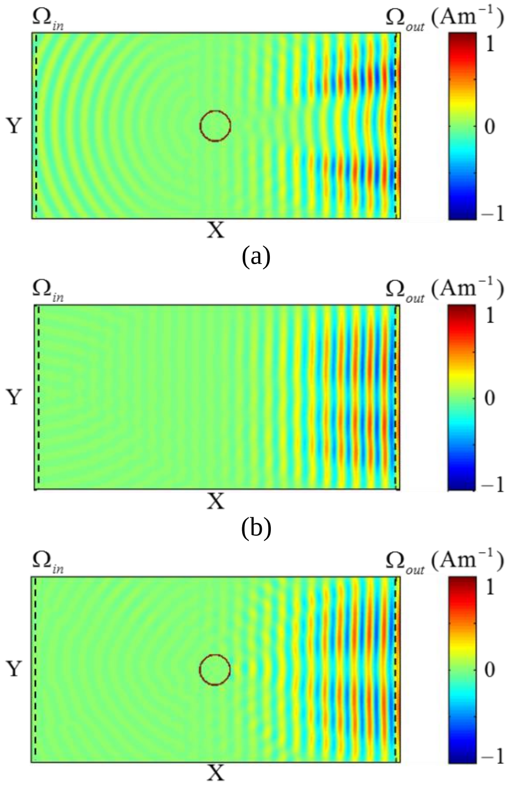


(a)

(b)

(c)

Fig. 5. TM illumination; the distributions of the optimal: (a) ε_{rxxj} and (b) ε_{ryyj} of the cloak voxels. (c) Objective function per iteration.



(a)

(b)

(c)

Fig. 6. TM illumination; the snapshots of the H_z distributions throughout the waveguide for the: (a) bare object, (b) empty space, and (c) cloaked object.

IV. CONCLUSION

Exploiting a fast anisotropic memory efficient AVM simulator, an all-dielectric wideband cloaks are designed at optical frequency region for both TE and TM illuminations. Designing such cloaks using classical approaches is prohibitive at optical frequencies due to the formidable computation cost. The sensitivities of a cost function with respect to a huge number of parameters are required per iteration. Our anisotropic AVM simulator significantly accelerates the sensitivities computation cost enables wideband invisibility cloaking at optical frequencies.

REFERENCES

[1] W. Yu, X. Yang, Y. Liu, R. Mittra, D. Chang, C. Liao, M. Akira, W. Li, and L. Zhao, "New development of parallel conformal FDTD method in computational electromagnetics engineering," *IEEE Antennas and Propag. Magzine*, vol. 53, pp. 15-41, 2011.

[2] W. Yu, X. Yang, Y. Liu, R. Mittra, Q. Rao, and A. Muto, "High performance conformal FDTD techniques," *IEEE Microwave Magazine*, vol. 11, pp. 42-55, 2010.

[3] W. Yu, W. Li, A. Elsherbeni, and Y. R. Samii, *Advanced Computational Electromagnetic Methods and Applications*. Norwood, MA: Artech House, 2015.

[4] E. Haber, *Computational Methods in Geophysical Electromagnetics*. Philadelphia, PA: Society for Inducterial and Applied Mathematics, 2015.

[5] M. H. Bakr, *Nonlinear Optimization in Electrical Engineering with Applications in Matlab©*. London, UK: IET, 2013.

[6] J. Fan, F. Han, and H. Liu, “Challenges of big data analysis,” *Natl. Sci. Rev.*, vol. 1, pp. 293-314, 2014.

[7] L. S. Kalantari, O. S. Ahmed, M. H. Bakr, and N. Nikolova, “Adjoint sensitivity analysis of 3D problems with anisotropic materials,” *IEEE Int. Microw. Symp. Digest*., Tampa Bay, FL, pp. 1-3, 2014.

[8] A. Alù and N. Engheta, “Achieving transparency with plasmonic and metamaterial coatings,” *Phys. Rev. E*, vol. 72, pp. 016623, 2005.

[9] J. B. Pendry, D. Schurig, and D. R. Smith, “Controlling electromagnetic fields,” *Science*, vol. 312, no. 5781, pp. 1780-1782, 2006.

[10] J. Zhang, L. Liu, Y. Luo, S. Zhang, and N. A. Mortensen, “Homogeneous optical cloak constructed with uniform layered structures,” *Opt. Express*, vol. 19, no. 9, pp. 8625-8631, 2011.

[11] W. Cai, U. K. Chettiar, A. V. Kildishev, and V. M. Shalaev, “Optical cloaking with metamaterials,” *Nat. Photon*., vol. 1, pp. 224-227, 2007.

[12] B. I. Popa and S. A. Cummer, “Cloaking with optimized homogeneous anisotropic layers,” *Phys. Rev. A*., vol. 79, pp. 023806, 2009.

[13] J. Andkjaer and O. Sigmund, “Topology optimized low-contrast all-dielecteric optical cloak,” *Appl. Phys. Lett*., vol. 98, pp. 021112, 2011.

[14] X. Wang and E. Semouchkina, “A route for efficient non-resonance cloaking by using multilayer dielectric coating,” *Appl. Phys. Lett*., vol. 102, pp. 113506, 2013.

[15] Y. Urzhumov, N. Landy, T. Driscoll, D. Basov, and D. R. Smith, “Thin low-loss dielectric coatings for free-space cloaking,” *Opt. Lett*., vol. 38, pp. 1606-1608, 2013.

[16] A. Håkansson, “Cloaking of objects from electromagnetic fields by inverse design of scattering optical elements,” *Opt. Express*, vol. 15, no. 7, pp. 4328-4334, 2007.

[17] L. S. Kalantari and M. H. Bakr, “Cloaking exploiting anisotropic adjoint sensitivity analysis,” *IEEE Antennas and Propag. Soc. Int. Symp. Digest.*, Vancouver, BC, pp. 61-62, 2015.

[18] L. S. Kalantari and M. H. Bakr, “Wideband cloaking of objects with arbitrary shape exploiting adjoint sensitivities,” *IEEE Trans. Antennas. Propag.*, vol. 64, no. 5, pp. 1963-1968, 2016.

Laleh Seyyed-Kalantari received the B.Sc. and M.Sc. degree with honors in Electrical Engineering (Telecommunications) from Sadjad Institute of Higher Education, Mashhad, Iran, in 2006 and University of Sistan and Baluchestan, Zahedan, Iran in 2009, respectively. She is currently pursuing the Ph.D. degree in Electrical Engineering at McMaster University, Hamilton, ON, Canada. In 2011, she joined the Computational Electromagnetic Laboratory (CEML), McMaster University. Her research interests include numerical method in electromagnetics, optimization of microwave and photonic devices, genetic algorithm and soft computing, metamaterials and cloaking.

Seyyed-Kalantari is a recipient of the Ontario Graduate Scholarship (OGS) Award in 2013, OGS and Queen Elizabeth II Graduate Scholarship in Science and Technology Award in 2014, and Research in Motion Ontario Graduate Scholarship Award in 2015.

Mohamed H. Bakr received a B.Sc. degree in Electronics and Communications Engineering from Cairo University, Egypt in 1992 with distinction (honors). In June 1996, he received a Master's degree in Engineering Mathematics from Cairo University. He earned the Ph.D. degree in September 2000 from the Department of Electrical and Computer Engineering, McMaster University. In November 2000, he joined the Computational Electromagnetics Research Laboratory (CERL), University of Victoria, Victoria, Canada as an NSERC Post Doctoral Fellow. He is currently with the Department of Electrical and Computer Engineering, McMaster University.

Bakr received a Premier’s Research Excellence Award (PREA) from the province of Ontario, Canada, in 2003. He was a recipient of an NSERC DAS Award in 2011 and a co-recipient of the 2014 Chrysler Innovation Award. His research areas of interest include optimization methods, computer-aided design and modeling of microwave circuits and photonic devices, neural network applications, smart analysis of high frequency structures, and design of electric motors and power circuits.

Analysis and Design of Class E Power Amplifier with Finite DC-Feed Inductance and Series Inductance Network

Chuicai Rong [1,2], Xiansuo Liu [1], Yuehang Xu [1]*, Ruimin Xu [1], and Mingyao Xia [1]

[1] School of Electronic Engineering
University of Electronic Science and Technology of China, Chengdu, 611731, China
*yuehangxu@uestc.edu.cn

[2] School of Physics and Electronic Information
Gannan Normal University, Ganzhou, Jiangxi, 341000, China
chuicair@126.com

Abstract **─** With the increasing operation frequency, it is essential to take into account the parasitic parameters of transistor for high efficiency microwave power amplifier design. In this paper, a class E power amplifier with finite dc-feed inductance and series inductance network is analyzed including the parasitic inductance of transistor. The analytical design expressions are derived. And the effects of series inductance on the load network parameter are obtained. The results suggest that this new topology can be used in broadband power amplifiers design by making full use of transistor's output parasitic inductance. A GaN HEMT power amplifier is designed with the proposed topology for validation purpose. Experimental results show that the amplifier can realize from 2.5 GHz to 3.5 GHz (33.3%) with measured drain efficiency larger than 60% and output power larger than 34 dBm. The measured performance shows good agreement with the theoretical performance predicted by the equations.

Index Terms **─** Broadband, class E power amplifier, finite dc-feed inductance, parasitic inductance.

I. INTRODUCTION

One of the most important features of RF power amplifier (PA) is power efficiency. By increasing the efficiency, PA will consume less supply power and requires less heat sinking. This allows a reduction of battery size and an increase in battery life. The switch mode class E PA [1] is a good candidate for high efficiency PA due to its design simplicity.

The class E PA with finite dc-feed inductance [2, 3] is one important topology of the class E PA. It has smaller inductance than the RF-choke and thus has lower loss [4] due to a smaller electrical series resistance (ESR). It can obtain greater power capability than other class E topology. And the larger load resistance makes the design of the matching network easier. These advantages make this topology widely attracted. In [5], the effects of dc-feed inductance, the quality factor (Q_L) of the series-tuned circuit, and the switching-device on resistance have been analyzed. In [6], the maximum frequency of the class E PA with finite dc-feed inductance is discussed. In [7], an arbitrary duty-cycle and finite dc-feed inductance is discussed. In [8], the power dissipation in each component is calculated. In [9], load transformation networks for wideband operation is investigated. In [10], the analytical expression of the switch peak voltage is presented. With the increasing operation frequency, it is essential to take into account all the device parasitic parameters [11, 12]. In [13, 14], the normalized optimum load network parameters versus normalized bond-wire inductance for parallel-circuit class E PA are presented. But the parallel-circuit class E PA is only one kind of the class E power amplifier with finite dc-feed inductance. To get the general results, it is necessary to further study the effect of the device output series inductance on the load network parameters of the class E power amplifier with finite dc-feed inductance.

In this paper, a theoretical description of the class E PA with finite dc-feed inductance and series inductance network is presented. The analysis takes into account the transistor's output parasitic inductance on the load network parameters of the class E PA with finite dc-feed inductance. Thus, the analysis can provide useful and accurate design to the class E PA in higher operation frequency. Finally, a design case is constructed in the laboratory in order to verify the theoretical predictions for demonstration purpose.

II. CIRCUIT DESCRIPTION

The class E power amplifier with finite dc-feed inductance and series inductance is shown in the Fig. 1. The load network consists of the shunt capacitance C_0, a series inductance L_{series}, a parallel inductance L_0, a

Submitted On: March 3, 2017
Accepted On: March 18, 2017

series reactive element jX, and a load R. The shunt capacitance C_0 represents the intrinsic device output capacitance. The series inductance L_{series} can be considered as an adjustment parameter which include the bond-wire inductance and lead inductance. A parallel inductance L_0 represents the finite DC-feed inductance and the series reactive element jX can be positive (inductance) or negative (capacitance) or zero. The active device is considered to be an ideal switch.

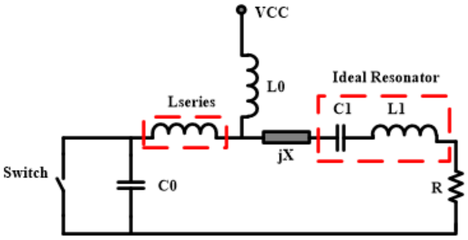


Fig. 1. Circuit of the class E power amplifier with finite dc-feed inductance and series inductance network.

To simplify analysis of the class E power amplifier with finite dc-feed inductance and series inductance, several assumptions are introduced in [13, 14]. For an idealized theoretical analysis, the moments of the switch-on is $\omega t = 0$ and switch-off is $\omega t = \pi$ with period of repeatability of the input driving signal $T = 2\pi$. Nominal conditions for voltage across the switch prior to the start of switch–on at the moment $\omega t = 2\pi$ are:

$$v(\omega t)\big|_{\omega t=2\pi} = 0, \tag{1}$$

$$\left.\frac{dv(\omega t)}{dt}\right|_{\omega t=2\pi} = 0. \tag{2}$$

The output current flowing through the load R is written as sinusoidal by:

$$i_R(\omega t) = I_R \sin(\omega t + \phi), \tag{3}$$

where I_R is the load current amplitude and ϕ is the initial phase shift.

When the switch is turned on for $0 \le \omega t \le \pi$, the voltage on the switch is zero. The current flowing through the switch can be written as:

$$i(\omega t) = \frac{V_{cc}}{\omega L_0(1+\alpha)}\omega t + \frac{\omega L_0 I_R}{\omega L_0(1+\alpha)}\left[\sin(\omega t+\phi) - \sin\phi\right], \tag{4}$$

where $\alpha = L_{series}/L_0$.

When switch is off for $\pi \le \omega t \le 2\pi$, the current $i(\omega t) = 0$ and the current $i_{c_0}(\omega t) = i_{L_0}(\omega t) + i_R(\omega t)$ flowing the capacitance C_0 can be rewritten as:

$$\omega C_0 \frac{dv(\omega t)}{d(\omega t)} = \frac{1}{\omega L_0}\int_{\pi}^{\omega t}\left[V_{cc} - v(\omega t) - v_{L_{series}}(\omega t)\right]d(\omega t) + i_{L_0}(\pi) + I_R \sin(\omega t + \phi). \tag{5}$$

Differentiating both sides of (5), the second-order differential equation becomes:

$$\omega^2 C_0 L_0 (1+\alpha)\frac{d^2 v(\omega t)}{d(\omega t)^2} + v(\omega t) - V_{cc} - \omega L_0 I_R \cos(\omega t + \phi) = 0. \tag{6}$$

Under the initial off-state conditions,

$$v(\pi) = 0, \tag{7}$$

The current $i_{L_0}(\pi)$ flowing through the finite inductance L_0 is:

$$i_{L_0}(\pi) = \frac{V_{cc}\pi}{\omega L_0(1+\alpha)} - \frac{1-\alpha}{1+\alpha} I_R \sin\phi. \tag{8}$$

The current flowing through the capacitance C_0 is:

$$i_{C_0}(\pi) = i_{L_0}(\pi) + i_R(\pi) = \frac{V_{cc}\pi - 2\omega L_0 I_R \sin\phi}{\omega L_0(1+\alpha)}, \tag{9}$$

$$i_{C_0}(\omega t) = \omega C_0 \frac{dv(\omega t)}{d(\omega t)}, \tag{10}$$

$$\left.\frac{dv(\omega t)}{d(\omega t)}\right|_{\omega t=\pi} = \frac{V_{cc}\pi - 2\omega L_0 I_R \sin\phi}{\omega^2 C_0 L_0 (1+\alpha)} = \frac{V_{cc}(\pi - 2p\sin\phi)}{Q^2}, \tag{11}$$

where

$$Q^2 = \omega^2 C_0 L_0 (1+\alpha) = \frac{\omega^2}{\omega_0^2}(1+\alpha) = \chi^2(1+\alpha), \tag{12}$$

$$\omega_0 = \frac{1}{\sqrt{C_0 L_0}},\ \chi = \frac{\omega}{\omega_0},\ p = \frac{\omega L_0 I_R}{V_{cc}}, \tag{13}$$

where χ is the normalized frequency. With the initial off-state conditions (7) and (11), the general solution of (6) can be obtained in the normalized forms:

$$\frac{v(\omega t)}{V_{cc}} = C_1 \cos\left(\frac{\omega t}{Q}\right) + C_2 \sin\left(\frac{\omega t}{Q}\right) + 1 - \frac{p}{Q^2-1}\cos(\omega t + \phi), \tag{14}$$

$$C_1 = -\left[\cos\left(\frac{\pi}{Q}\right) + \frac{\pi}{Q}\sin\left(\frac{\pi}{Q}\right)\right] - \frac{Qp}{Q^2-1}\left[\frac{\cos\phi}{Q}\cos\left(\frac{\pi}{Q}\right) - \frac{Q^2-2}{Q^2}\sin\phi\sin\left(\frac{\pi}{Q}\right)\right], \tag{15}$$

$$C_2 = \left[\frac{\pi}{Q}\cos\left(\frac{\pi}{Q}\right) - \sin\left(\frac{\pi}{Q}\right)\right] - \frac{Qp}{Q^2-1}\left[\frac{\cos\phi}{Q}\sin\left(\frac{\pi}{Q}\right) - \frac{Q^2-2}{Q^2}\sin\phi\cos\left(\frac{\pi}{Q}\right)\right]. \tag{16}$$

Applying nominal conditions of (1) and (2), the optimum parameters ϕ and p as functions of Q are:

$$\tan\phi = -\frac{\pi + \pi\cos\left(\frac{\pi}{\phi}\right) + 2Q\sin\left(\frac{\pi}{Q}\right)}{2(Q^2-1)\left[1-\cos\left(\frac{\pi}{Q}\right)\right] + Q\pi\sin\left(\frac{\pi}{Q}\right)}, \tag{17}$$

$$p=\frac{(1-Q^2)\left[Q\sin(\frac{\pi}{Q})+\pi\cos(\frac{\pi}{Q})\right]}{Q\cos\phi\sin(\frac{\pi}{Q})+\left[Q^2+(2-Q^2)\cos(\frac{\pi}{Q})\right]\sin\phi}, \quad (18)$$

where Q is a function of α and χ. Figure 2 shows the initial phase shift ϕ versus α and χ. With the increasing of χ, the initial phase shift ϕ decreases. With the increasing of α, the initial phase shift ϕ decreases and the gradient of the initial phase shift ϕ is slow in broadband. Thus, it is easy to match for load network.

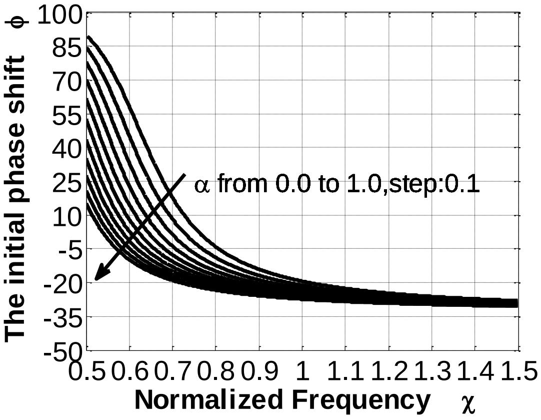


Fig. 2. Initial phase shift ϕ versus α and χ.

The normalized load-network parameters inductance L_0, capacitance C_0, and resistance R are presented as functions of parameters p,ϕ,α,Q, as below:

$$\frac{\omega L_0}{R}=\frac{p(1+\alpha)}{\frac{\pi}{2p}+\frac{2\cos\phi}{\pi}-\sin\phi}, \quad (19)$$

$$\omega C_0 R=\frac{Q^2}{p(1+\alpha)^2}\left[\frac{\pi}{2p}+\frac{2\cos\phi}{\pi}-\sin\phi\right], \quad (20)$$

$$\frac{RP_{out}}{V_{cc}^2}=\frac{\left[\frac{\pi^2}{2p}+2\cos\phi-\pi\sin\phi\right]^2}{2\pi^2(1+\alpha)^2}. \quad (21)$$

From the viewpoint of mathematics, they are only functions of parameters α and χ. Figure 3 shows the parameter $\omega L_0/R$ versus α and χ. With the increasing of χ, the parameter $\omega L_0/R$ increases. With the increasing of α, the parameter $\omega L_0/R$ increases.

Figure 4 shows the parameter $\omega C_0 R$ versus α and χ. The parameter $\omega C_0 R$ has maximum of 0.7021 when $\chi=0.681$ and $\alpha=0$. Then the maximum frequency of the class E power amplifier with finite dc-feed inductance and series inductance is $f_{\max}=0.7021/(2\pi C_0 R)$. With the increasing of α, the parameter $\omega C_0 R$ decreases.

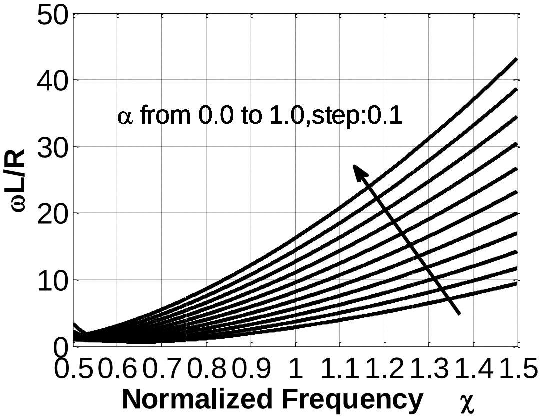


Fig. 3. $\omega L_0/R$ versus α and χ.

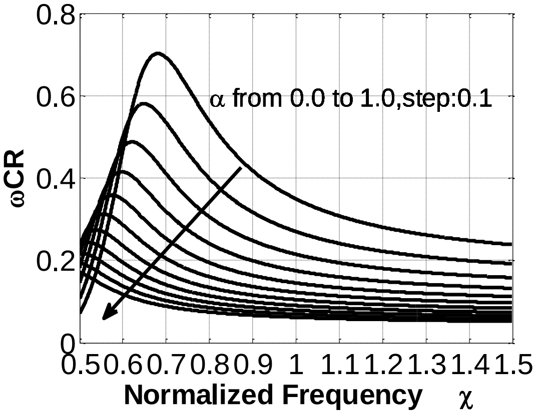


Fig. 4. $\omega C_0 R$ versus α and χ.

Figure 5 shows the parameter RP_{out}/V_{cc}^2 versus α and χ. The parameter RP_{out}/V_{cc}^2 has maximum of 1.3633 when $\chi=0.709$ and $\alpha=0$. Then the maximum load resistance of the class E power amplifier with finite dc-feed inductance and series inductance is $R_{\max}=1.3633V_{cc}^2/P_{out}$. With the increasing of α, the parameter RP_{out}/V_{cc}^2 decreases.

The Equations (22) through (25) below present the analytical expressions of the voltage across the reactance X. The Equations (26) through (29) below present the analytical expressions of the voltage across the resistance *R*.

$$V_X=\frac{1}{\pi}\left(V_{X1}+V_{X2}+V_{X3}\right), \quad (22)$$

$$V_{X1}=\frac{\alpha(\frac{\pi p}{2}-2\sin\phi)}{1+\alpha}+\frac{\pi p(\alpha\chi^2-1)}{2(Q^2-1)}+2\sin\phi, \quad (23)$$

$$V_{X2} = \frac{C_1}{2}\left(1-\frac{\alpha\chi^2}{Q^2}\right)\left\{\begin{array}{l}\frac{Q}{1+Q}\left[\sin\left(\frac{2\pi}{Q}+\phi\right)+\sin\left(\frac{\pi}{Q}+\phi\right)\right]\\ +\frac{Q}{1-Q}\left[\sin\left(\frac{2\pi}{Q}-\phi\right)+\sin\left(\frac{\pi}{Q}-\phi\right)\right]\end{array}\right\}, \tag{24}$$

$$V_{X3} = -\frac{C_2}{2}\left(1-\frac{\alpha\chi^2}{Q^2}\right)\left\{\begin{array}{l}\frac{Q}{1+Q}\left[\cos\left(\frac{2\pi}{Q}+\phi\right)+\cos\left(\frac{\pi}{Q}+\phi\right)\right]\\ +\frac{Q}{1-Q}\left[\cos\left(\frac{2\pi}{Q}-\phi\right)+\cos\left(\frac{\pi}{Q}-\phi\right)\right]\end{array}\right\}, \tag{25}$$

$$V_R = \frac{1}{\pi}\left(V_{R1}+V_{R2}+V_{R3}\right), \tag{26}$$

$$V_{R1} = -\frac{2\cos\phi}{1+\alpha}, \tag{27}$$

$$V_{R2} = \frac{C_1}{2}\left(1-\frac{\alpha\chi^2}{Q^2}\right)\left\{\begin{array}{l}\frac{Q}{1-Q}\left[\cos\left(\frac{2\pi}{Q}-\phi\right)+\cos\left(\frac{\pi}{Q}-\phi\right)\right]\\ -\frac{Q}{1+Q}\left[\cos\left(\frac{2\pi}{Q}+\phi\right)+\cos\left(\frac{\pi}{Q}+\phi\right)\right]\end{array}\right\}, \tag{28}$$

$$V_{R3} = -\frac{C_2}{2}\left(1-\frac{\alpha\chi^2}{Q^2}\right)\left\{\begin{array}{l}\frac{Q}{1+Q}\left[\sin\left(\frac{2\pi}{Q}+\phi\right)+\sin\left(\frac{\pi}{Q}+\phi\right)\right]\\ -\frac{Q}{1-Q}\left[\sin\left(\frac{2\pi}{Q}-\phi\right)+\sin\left(\frac{\pi}{Q}-\phi\right)\right]\end{array}\right\}, \tag{29}$$

$$\frac{X}{R} = \frac{V_X}{V_R}. \tag{30}$$

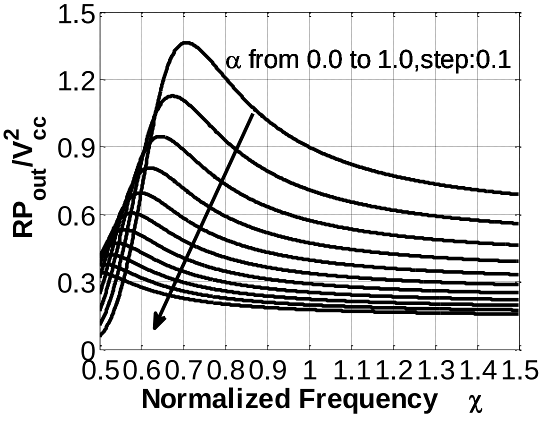


Fig. 5. RP_{out}/V_{cc}^2 versus α and χ.

Figure 6 shows the parameter X/R versus α and χ. The X/R is equal to zero when $\chi < 0.708$ and $\alpha = 0$, which is the parallel-circuit class E PA. When $\chi < 0.708$ and $\alpha = 0$, the reactance X is positive (inductive reactance). When $\chi < 0.708$ and $\alpha = 0$, the reactance X is negative (capacitive reactance). When $\chi < 0.708$, with the increasing of α, the parameter X/R decreases, even below zero. When $\chi < 0.708$, with the increasing of α, the parameter X/R increases.

In Table 1, the optimized load-network parameters of the different class E modes include class E with shunt filter and the class E with finite dc-feed inductance and series inductance are shown in a normalized form. As can be seen, the class E with finite dc-feed inductance and series inductance offers the larger value of the power output capability c_P and the load R, which is 3.18 times higher than that for class E with shunt capacitance.

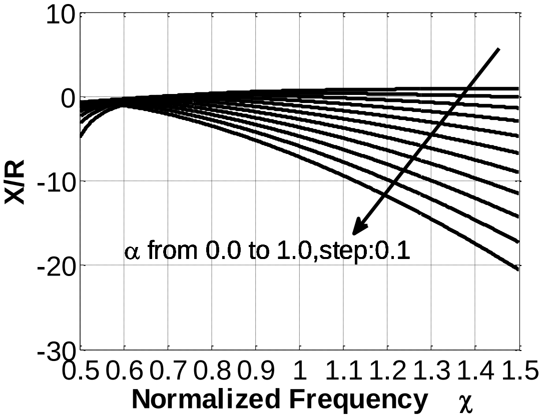


Fig. 6. X/R versus α and χ.

Table 1: Load network parameters for different class E modes

Normalized Load-Network Parameter	Class E with Shunt Capacitance and Shunt Filter[15]	Class E with Finite DC-Feed Inductance and Series Inductance
$\frac{X}{R}$	1.4836	0 $(\alpha = 0,\ \chi = 0.708)$
ωCR	0.261	0.7021 $(\alpha = 0,\ \chi = 0.681)$
$\frac{P_{out}R}{V_{CC}^2}$	0.4281	1.3633 $(\alpha = 0,\ \chi = 0.681)$
$\frac{f_{\max}C_{out}V_{CC}^2}{P_{out}}$	0.097	0.1505 $(\alpha = 0,\ \chi = 0.693)$
c_P	0.09825	0.1049 $(\alpha = 0,\ \chi = 0.6689)$

III. DESIGN CONSIDERATION

For broadband PA design [15], the susceptance of the network is an important parameter. Figure 7 shows the susceptance $Imag[V_{cc}^2 Y_{net}(\chi)/P_{out}]$ of the class E PA with finite dc-feed inductance and series inductance.

The χ when α increase from 0.0 to 1.0 by step 0.1. When $\alpha = 0$, the difference of susceptance is 0.2389 in the frequency range $(0.5 \le \chi \le 1.5)$. When $\alpha = 1$, the difference of susceptance is 0.0381 in the frequency range $(0.5 \le \chi \le 1.5)$. The parameter α can be used to control the difference of susceptance over a wide frequency range.

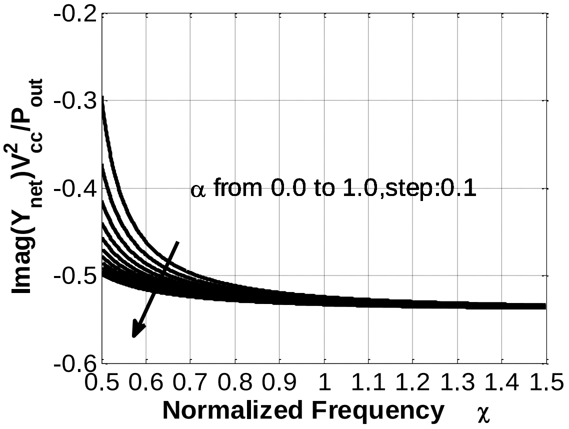


Fig. 7. Susceptance $Imag[V_{cc}^2 Y_{net}(\chi)/P_{out}]$ versus α and χ.

Figure 8 shows the conductance $\mathrm{Re}al[V_{cc}^2 Y_{net}(\chi)/P_{out}]$ of class E PA with finite dc-feed inductance and series inductance versus χ when α increase from 0.0 to 1.0 by step 0.1. When $\alpha = 0$, the difference of conductance is 0.0239 in the frequency range $(0.5 \le \chi \le 1.5)$. When $\alpha = 1$, the difference of conductance is 0.01 in the frequency range $(0.5 \le \chi \le 1.5)$. The parameter α can be used to control the difference of conductance over a wide frequency range.

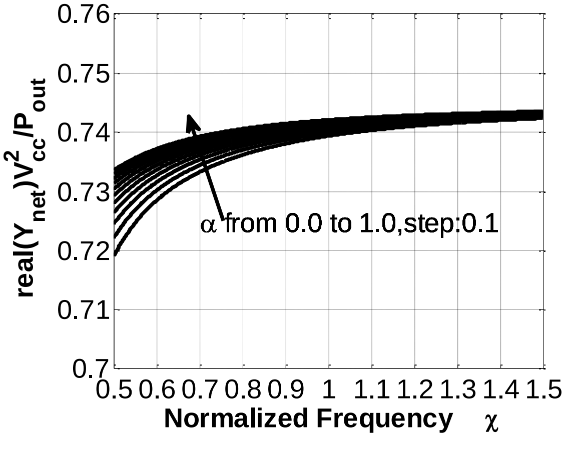


Fig. 8. Conductance $\mathrm{Re}al[V_{cc}^2 Y_{net}(\chi)/P_{out}]$ versus α and χ.

Figure 9 shows load phase angle of class E PA with finite dc-feed inductance and series inductance versus χ when α increase from 0.0 to 1.0 by step 0.1. When $\alpha = 0$, the difference of load phase angle is 13.42° in the frequency range $(0.5 \le \chi \le 1.5)$. When $\alpha = 1$, the difference of load phase angle is 1.65° in the frequency range $(0.5 \le \chi \le 1.5)$. The parameter α can be used to control the difference of load phase angle over a wide frequency range.

In a word, by proper choice of the series inductance αL_0, which produces a zero total variation of the susceptance, the conductance and the load phase angle are controllable over a wide frequency range.

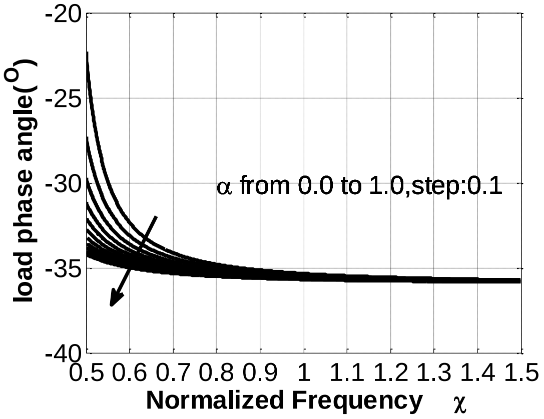


Fig. 9. Load phase angle versus α and χ.

IV. SIMULATION AND IMPLEMENTATION

A complete circuit schematic of class E PA with finite dc-feed inductance and series inductance is shown in Fig. 10. A 0.25μm gate length GaN HEMT with 1.25 mm total gate-width $(C_{ds} = 0.254pF)$ is used to design a Class E PA with finite dc-feed inductance and series inductance. The simulation of amplifier is realized by combing Ansys HFSS and Keysight ADS. The HFSS is used to simulate passive part of matching network. A large signal model is established to simulate the large signal performance of amplifier with HB simulation tool [16]. The total inductance $L_{series} = L_{para} + L_{wire} + L_1$, where L_{para} is the output parasitic inductance of transistor, L_{wire} is the inductance induced by bonding wire for hybrid amplifier, and L_1 is the adjustive inductance. The parasitic output capacitance C_{out} of the transistor, L_{series}, L_2, and the reactance C_3 constitute the double L-type network. The inductance L_1 and L_2 is realized by the high impedance transmission line.

Typically, class E PA achieve high efficiency when the output power gain at 3 dB or 4 dB compression point

[17, 18]. So it is necessary to suppress the second and third harmonic to improve efficiency. Low pass match was used in operation frequency and suppress the harmonics both in the input and output network [19, 20]. Shunt resistance R_1 and capacitance C_2 in the input network was used to improve the low frequency stability. The photo of the class E PA with finite dc-feed inductance and series inductance is shown in Fig. 11.

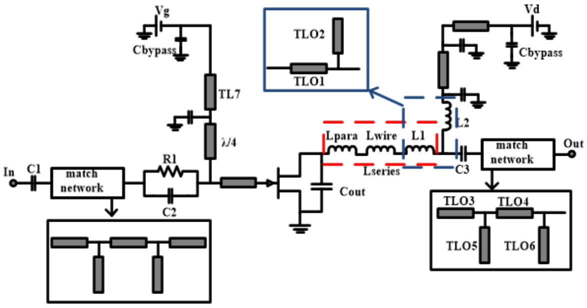


Fig. 10. The circuit schematic of the class E PA with finite dc-feed inductance and series inductance.

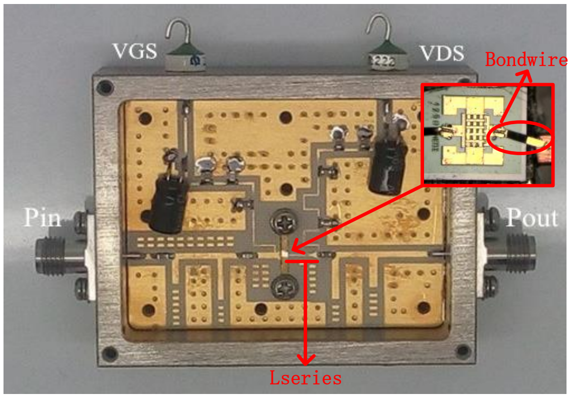


Fig. 11. Photo of fabricated class E PA with finite dc-feed inductance and series inductance.

Figure 12 shows the measured drain efficiency (DE), power added efficiency (PAE), output power, and gain at input power (CW) at 3.1 GHz. The maximum PAE is 63.4% when the input power is 28 dBm.

Figure 13 shows the measured behavior of DE, PAE [25], output power, and output power gain at the input power of 27 dBm. It can be seen that, the output power gain is large than 8.2 dB, while the output power is more than 35.2 dBm between 2.5 GHz and 3.5 GHz (33.3% fractional band width (FBW)).

Figure 14 shows the simulated and the measured power second harmonic over the bandwidth. The maximum power of second harmonic in this frequency band is -22 dBc at 2.5 GHz and the minimum is -55 dBc at 3.0 GHz. Most of the second harmonics power is below -30 dBc.

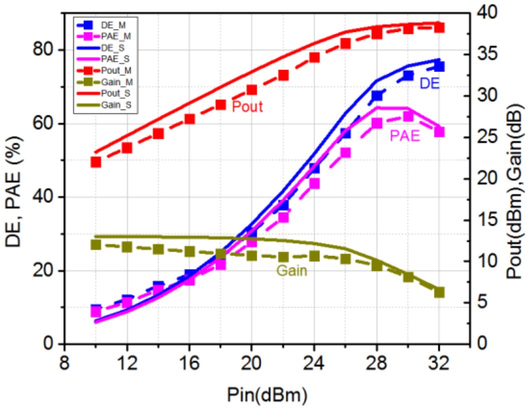


Fig. 12. Simulated and measured DE, PAE, output power and gain versus input power at 3.1 GHz continuous input signal.

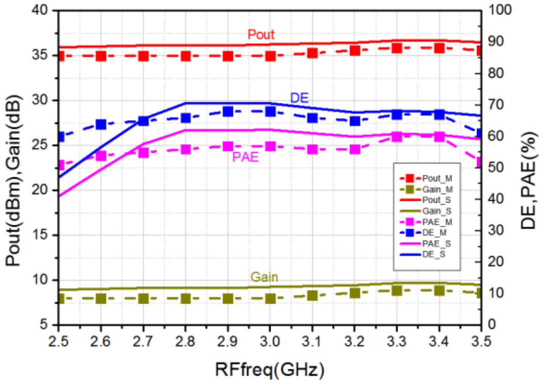


Fig. 13. Simulated and measured frequency dependence of DE, PAE, and Gain characteristics performance.

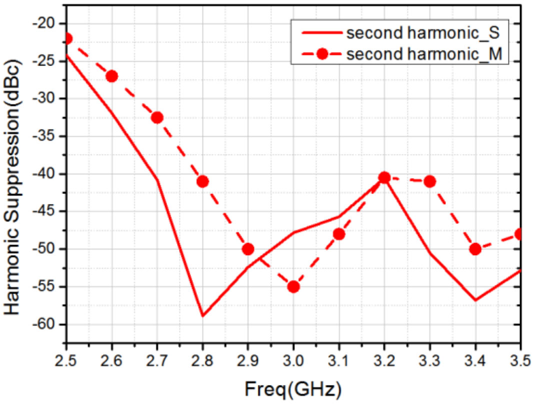


Fig. 14. Simulated and measured second harmonic power.

The measured performance of the proposed PA is compared with other state-of-the-art class E Pas. The results show that the proposed PA can achieve more than

60% DE in such high frequency as summarized in Table 2. The amplifier shows competitive drain efficiency and bandwidth in higher operation frequency. Because of the small gate-width device applied in the present PA, the output power is not very large. However, this report has the highest operation frequency. We can acquire higher output power through increasing the gate-width.

Table 2: Comparison of state of the art GaN PAs

	(GHz)/FBW (%)	DE (%)	Pout
2010 [21]	1.9-2.9 (42%)	>63	45.8 dBm
2011 [22]	2.15-2.5 (15%)	>60	>23 dBm
2011 [23]	0.9-2.2 (84%)	>63	>10 W
2014 [24]	2.52-2.64 (4.6%)	>60	>39 dBm
2015 [25]	1.7-2.8 (48.8%)	>60.3	>19.5 W
2016 [15]	1.4-2.7 (63.4%)	>63	>39.7 dBm
This work	2.5-3.5 (33.3%)	>60	>35.2 dBm

V. CONCLUSION

The class E power amplifier with finite dc-feed inductance and series inductance is analyzed in time domain. Analytical expressions of optimum parameters of the load network are derived. It suggests that the topology can be used in higher operation frequency and broadband PA design with competitive efficiency. A GaN HEMT class E PA with finite dc-feed inductance and series inductance is fabricated and measured. The experimental data and theoretical predictions are found in good agreements. The proposed structure may be useful in the coming 5G communication systems.

ACKNOWLEDGMENT

This work was supported by the National Natural Science Foundation of China (Grant No. 61474020), China Postdoctoral Science Foundation (Grant No. 2015M570775, 2015T80969) and the National Key Project of Science and Technology.

REFERENCES

[1] A. Grebennikov, N. O. Sokal, and M. J. Franco, *Switchmode RF and Microwave Power Amplifiers*. New York, NY, USA: Academic, 2012.

[2] R. Zulinski and J. Steadman, “Class E power amplifiers and frequency multipliers with finite DC-feed inductance,” *IEEE Transactions on Circuits and Systems*, vol. 34, no. 9, pp. 1074-1087, Sep. 1987.

[3] M. Acar, A. J. Annema, and B. Nauta, “Generalized design equations for class-E power amplifiers with finite DC feed inductance,” *2006 European Microwave Conference, Manchester*, pp. 1308-1311, 2006.

[4] D. Milosevic, J. van der Tang, and A. Van Roermund, “Explicit design equations for class-E power amplifiers with small DC-feed inductance,” *Circuit Theory and Design, 2005. Proceedings of the 2005 European Conference on IEEE*, 2005: III/101-III/104, vol. 3, 2005.

[5] C. P. Avratoglou, N. C. Voulgaris, and F. I. Ioannidou, “Analysis and design of a generalized class E tuned power amplifier,” *IEEE Transactions on Circuits and Systems*, vol. 36, no. 8, pp. 1068-1079, Aug. 1989.

[6] C. H. Li and Y. O. Yam, “Maximum frequency and optimum performance of class E power amplifiers,” IEEE *Proceedings - Circuits, Devices and Systems*, vol. 141, no. 3, pp. 174-184, June 1994.

[7] M. Acar, A. J. Annema, and B. Nauta, “Analytical design equations for class-E power amplifiers,” *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 54, no. 12, pp. 2706-2717, Dec. 2007.

[8] J. Y. Hasani and M. Kamarei, “Analysis and optimum design of a class E RF power amplifier,” *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 55, no. 6, pp. 1759-1768, July 2008.

[9] M. D. Wei, D. Kalim, D. Erguvan, S. F. Chang, and R. Negra, “Investigation of wideband load transformation networks for class-E switching-mode power amplifiers,” *IEEE Transactions on Microwave Theory and Techniques*, vol. 60, no. 6, pp. 1916-1927, June 2012.

[10] A. F. Jaimes and F. R. de Sousa, “Simple expression for estimating the switch peak voltage on the class-E amplifier with finite DC-feed inductance,” *2016 IEEE 7th Latin American Symposium on Circuits & Systems (LASCAS)*, Florianopolis, pp. 183-186, 2016.

[11] C. Wang, et al., “An electrothermal model for empirical large-signal modeling of AlGaN/GaN HEMTs including self-heating and ambient temperature effects,” *IEEE Transactions on Microwave Theory and Techniques*, vol. 62, no. 12, pp. 2878-2887, Dec. 2014.

[12] Z. Wen, Y. Xu, C. Wang, et al., “An efficient parameter extraction method for GaN HEMT small-signal equivalent circuit model,” *International Journal of Numerical Modelling Electronic Networks Devices & Fields*, 2015.

[13] A. Grebennikov, “Load network design techniques for class E RF and microwave amplifiers,” *High Frequency Electronics*, vol. 3, no. 7, pp. 18-32, 2004.

[14] A. Grebennikov, “Load network design technique for switched-mode tuned class E power amplifiers,” *High Frequency Electron*, vol. 3, no. 7, pp. 1-23, 2004.

[15] A. Grebennikov, “High-efficiency class-E power amplifier with shunt capacitance and shunt filter,” *IEEE Transactions on Circuits and Systems I:*

Regular Papers, vol. 63, no. 1, pp. 12-22, Jan. 2016.

[16] Y. Xu, C. Wang, H. Sun, et al., "A scalable large-signal multiharmonic model of AlGaN/GaN HEMTs and its application in C-band high power amplifier MMIC," *IEEE Transactions on Microwave Theory and Techniques*, vol. PP, no. 99, pp. 1-11, 2017.

[17] V. Zomorrodian, U. K. Mishra, and R. A. York, "Modeling of CPW based passive networks using Sonnet simulations for high efficiency power amplifier MMIC design," *Applied Computational Electromagnetics Society Journal*, vol. 26, no. 2, pp. 131-140, 2011.

[18] O. Kizilbey, O. Palamutçuogullari, S. B. Yarman, "3.5-3.8 GHz class-E balanced GaN HEMT power amplifier with 20W Pout and 80% PAE," *IEICE Electronics Express*, vol. 10, no. 5, pp. 20130104-20130104, 2013.

[19] S. Hietakangas and T. Rahkonen, "Input impedance of class E switching amplifiers," *2011 Workshop on Integrated Nonlinear Microwave and Millimetre-Wave Circuits*, Vienna, pp. 1-4, 2011. doi: 10.1109/INMMIC.2011.5773341

[20] V. P. Mcginn and V. A. Demir, "Dynamic measurement method for determining the output impedance of an RF power amplifier," *Applied Computational Electromagnetics Society Journal*, vol. 27, no. 4, pp. 302-310, 2012.

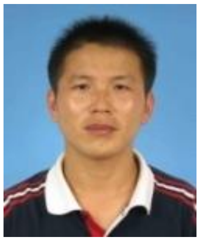
[21] D. Y. T. Wu, F. Mkadem, and S. Boumaiza, "Design of a broadband and highly efficient 45W GaN power amplifier via simplified real frequency technique," *Microwave Symposium Digest (MTT), 2010 IEEE MTT-S International*, Anaheim, CA, pp. 1090-1093, 2010.

[22] M. Thian, V. Fusco, and P. Gardner, "Power-combining class-E amplifier with finite choke," *Circuits & Systems I Regular Papers IEEE Transactions on*, vol. 58, no. 3, pp. 451-457, 2011.

[23] K. Chen and D. Peroulis, "Design of highly efficient broadband class-E power amplifier using synthesized low-pass matching networks," *IEEE Transactions on Microwave Theory and Techniques*, vol. 59, no. 12, pp. 3162-3173, Dec. 2011.

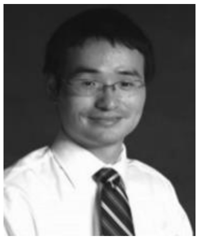
[24] X. Du, J. Nan, W. Chen, and Z. Shao, "'New' solutions of Class-E power amplifier with finite dc feed inductor at any duty ratio," *IET Circuits, Devices & Systems*, vol. 8, no. 4, pp. 311-321, July 2014.

[25] Y. Sun and X. Zhu, "Broadband continuous class-F^{-1} amplifier with modified harmonic-controlled network for advanced long term evolution application," *IEEE Microwave and Wireless Components Letters*, vol. 25, no. 4, pp. 250-252, Apr. 2015.



Chuicai Rong received the Bachelor of Science in Physics from Liaocheng University, Liaocheng, China, in 2000. Master of Science in Theoretical Physics from Qufu Normal University, Qufu, China, in 2006, and is currently working toward the Ph.D. degree at University of Electronic Science and Technology of China (UESTC). His research interests include microwave PA design and RF integrated circuit (RFIC) design.

Xiansuo Liu was born in Anhui Province, China, in 1992. He received the B.S. degree in University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2015, and is currently working toward the M.S. degree in Electromagnetic Field and Microwave Techniques from the UESTC. His current research interests include wideband high efficiency power amplifier and MMIC.



Yuehang Xu (M'11) was born in Zhejiang Province, China, in 1981. He received the B.S. and M.S. degree in Electromagnetic Field and Microwave Techniques from University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2004 and 2007, respectively, and the Ph.D. degree from the UESTC joined with Columbia University New York, NY, USA, in 2010.

He joined the Department of Electronic Engineering, UESTC, in December 2010 and was promoted to Associate Professor in August 2012. He is an author or a co-author of more than 30 scientific papers in international journals and conference proceedings. His current research interests are on modeling and characterization of micro/nano-scale electronic devices for radio frequency applications, especially on GaN HEMTs and graphene electronics.