A Probabilistic Model for the Nonlinear Electromagnetic Inverse Scattering: TM Case

Lianlin Li, *Senior Member, IEEE*, Long Gang Wang, Jun Ding, P. K. Liu, *Senior Member, IEEE*, M. Y. Xia, *Senior Member*, and Tie Jun Cui, *Fellow, IEEE*

Abstract—Electromagnetic inverse scattering (EMIS) is a non-invasive examination tool, which holds the promising potential in science, engineering, and military applications. In contrast to conventional tomography techniques, the inverse scattering is a quantitative superresolution imaging method since it is capable of accommodating more realistic interactions between the wavefield and the probed scene. In this paper, a full probabilistic formulation of the EMIS is presented for the first time, which is then solved by applying the well-known expectation maximization method. Afterward, the concept of the complex-valued alternating direction method of multipliers has been proposed as an alternative approach to solve the resulting nonlinear optimization problem. Finally, exemplary numerical and experimental results are provided to validate the proposed method.

Index Terms—Complex-valued alternating direction method of multiplier (ADMM), electromagnetic inverse scattering (EMIS), EM, nonlinear optimization problem, probabilistic approach.

I. INTRODUCTION

ELECTROMAGNETIC inverse scattering (EMIS) is a nondestructive investigation tool in civil engineering, military, and medical applications [1]–[3]. As it is capable of accommodating more realistic interactions between the wave field and probed scene, compared to the tomography technique, the EMIS method has been demonstrated to be a quantitative resolution-enhanced imaging technique [3]. In the past decades, many reconstruction algorithms have been developed, which could be roughly categorized into two groups: one group includes deterministic optimization methods, such as the contrast source inversion (CSI) [4], [5], distorted Born iterative method [6], [7], and others [8]–[10], and the other group includes stochastic optimization methods such as the gene algorithm and the particle swarm optimization [11]–[13]. It is well-known that the EMIS problem is ill posedness due to the very limited measurements compared to the unknowns to be recovered [8]. Recently, advanced by the methods in the low-dimensional model-based signal processing, several sparseness-aware algorithms have been developed to deal with

nonlinear [14]–[18] and the linearized [19]–[23] EMIS problems. For instance, Lee *et al.* [14] proposed a deterministic two-step noniterative method, where the contrast currents induced inside the probed object was first reconstructed by solving a joint sparsity penalized least-square problem, and the contrast function was achieved by straightforward manipulations. Interestingly, Poli *et al.* [15] and [16] established a probability formulation of the contrast currents by exploring the idea of multitask Bayesian compressive sensing developed in [24]. It has been demonstrated that these algorithms can produce the reconstruction with enhanced resolution when the probed scene is sparse or has sparse representation in the wavelet domain [21].

In this paper, a full probabilistic model of the nonlinear EMIS problem has been introduced with the aid of Bayesian probabilistic graph, which is different from the probabilistic method in [23] for the treatment of the linearized EMIS problem. Note that the approach developed in [23] is partial probabilistic in the sense that the contrast currents alone are treated in the probabilistic manner. However, both the contrast currents and the contrast are treated in the probabilistic way in this paper. Additionally, our approach is also different from that used in [23], where the Bayesian inference is heuristically exploited for solving the linear inverse problem arising from the deterministic nonlinear optimization EMIS problem at each iteration. Here, we would like to point out the deep neural network model of inverse scattering by Kamilov *et al.* [25], where the technique of artificial neural network in data science was, for the first time, brought to deal with the inverse scattering. In [25], the optical scattering was approximately treated with multiple phase screens, where only the forward propagation was taken into account, and the backward propagation had been explicitly neglected. Our full probabilistic formulation is based on the rigorous interactions between the electromagnetic wavefield and the investigated scenes. Then, the well-known expectation maximization method is used to accomplish our probabilistic inference.

The expectation maximization method has an attractive property that the local minimum of a convex cost function can be found in a reliable way by implementing the alternating iterative strategy [26]–[30]. However, for our problems, the implementation of expectation maximization method suffers from the specification of two artificial regularization factors, which could pose a computational challenge in practice. To address this issue, we introduce the complex-valued alternating direction method of multipliers (ADMMs). The complex-valued ADMM is a natural extension of conventional ADMM for
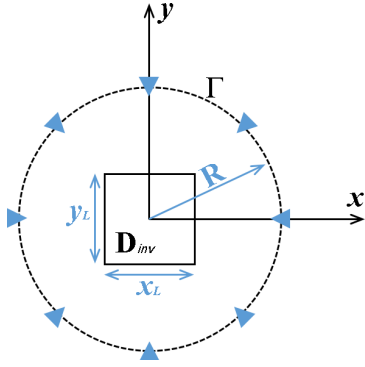
Fig. 1.   Illustrative map of considered EM inverse scattering scenario.

treating a real-valued optimization problem into a complex-valued one. The ADMM is a traditional technique for tackling large-scale optimization problems, and has gained renewed interests at this big-data era due to its unique advantage of very low complexity [31]. Besides, the proposed complex-valued ADMM inherits the advantage of the ADMM that it can be implemented in the parallel manner. In this way, the proposed method could pave an efficient way for solving large-scale EMIS problems.

The rest of this paper is organized as follows. In Section II, the full probabilistic formulation of EMIS is proposed, followed by the probabilistic inference based on the expectation maximization method. Afterward, a complex-valued ADMM algorithm is introduced to solve the resultant optimization problem. In Section III, exemplary numerical and experimental results are provided to show the effectiveness and accuracy of the proposed schemes, where the experimental data are kindly provided by the Institute Fresnel in Marseille (France) [32]. Finally, conclusions are summarized in Section IV.

II. PROBLEM STATEMENT AND METHODOLOGY

A. Problem Statement

For the purpose of illustrating the operational principle of the proposed methodology, we consider a monochromatic multistatic/multiview configuration shown in Fig. 1. For simplicity, we only treat this 2-D scenario; however, the proposed methodology can be readily generalized for much more complicated scenarios. The investigation domain denoted by D_{inv} (inaccessible region) is successively illuminated by TM-polarized incident waves $E_{\text{inc}}^{(n)}$, where $n = 1, 2, \ldots, N$ (with n being the index of the nth illumination). For each illumination, the M receivers distributed uniformly over Γ are used to collect the electrical field scattered from the probed scene. The transmitters and receivers are both located in the observation domain denoted by Γ exterior to D_{inv}. The time dependence factor $\exp(-j\omega t)$ with angular frequency ω is used and suppressed throughout this paper.

For the nth ($n = 1, 2, \ldots, N$) illumination and the mth ($m = 1, 2, \ldots, M$) receiver, the scattered field $E_{\text{sca}}^{(m,n)}$ at the location of r_m is governed by a pair of equations [4], [5]

$$E_{\text{sca}}^{(m,n)}(r_m) = k_0^2 \int_{D_{\text{inv}}} G(r_m, r)\chi(r)E^{(n)}(r)dr \quad (1)$$

and

$$E^{(n)}(r) = E_{\text{inc}}^{(n)}(r) + k_0^2 \int_{D_{\text{inv}}} G(r, r')\chi(r')E^{(n)}(r')dr'$$
$$r, r' \in D_{\text{inv}} \quad (2)$$

where $r = (x, y)$ and $r' = (x', y')$ denote the field and source points, respectively; $E^{(n)}$ represents the total electrical wavefield emerged from the interaction of probed scene with incident wavefield $E_{\text{inc}}^{(n)}$. The contrast function is defined as $\chi = k^2/k_0^2 - 1$, where k and $k_0 = \omega/c$ are the wavenumbers of the object and the background media, respectively. $G(r, r') = (i/4)H_0^{(1)}(k_0|r-r'|)$ is the 2-D Green's function, where $H_0^{(1)}(\cdot)$ is the zeroth-order Hankel function of first kind.

To proceed our methodology, the so-called equivalent current, referred as the contrast current, is introduced, which reads $J^{(n)}(r) = \chi(r)E^{(n)}(r)$ ($n = 1, 2, \ldots, N$) [4], [5], then (1) and (2) become

$$E_{\text{sca}}^{(m,n)}(r_m) = k_0^2 \int_{D_{\text{inv}}} G(r_m, r)J^{(n)}(r)dr \quad (3)$$

and

$$E^{(n)}(r) = E_{\text{inc}}^{(n)}(r) + k_0^2 \int_{D_{\text{inv}}} G(r, r')J^{(n)}(r')dr'$$
$$r, r' \in D_{\text{inv}}. \quad (4)$$

For the purpose of numerical implementations, (3) and (4) are discretized according to the Richmond procedure [4], [5]. To that end, the investigation domain D_{inv} is uniformly divided into K subsquares such that the total electrical field, the contrast currents, and the contrast function are approximately constant within each subsquare. Furthermore, (3) and (4) can be formulated in a compact form, i.e.,

$$E_{\text{sca}}^{(n)} = G_d J^{(n)} + \xi_d \quad (5)$$

and

$$J^{(n)} = \text{diag}(\chi)\left(E_{\text{inc}}^{(n)} + G_s J^{(n)}\right) + \xi_S. \quad (6)$$

The complex-valued ξ_d and ξ_S denote the modeling and/or the measurement errors, respectively, which are assumed to be the Gaussian random numbers in this paper for simplicity.

B. Probabilistic Model of Electromagnetic Inverse Scattering

This section is focused on the probabilistic formulation of nonlinear EMIS. We associate (5) and (6) with a three-layer probabilistic graph $G = (\mathcal{V}, \mathcal{E})$, as shown in Fig. 2, where $\mathcal{V}$ and $\mathcal{E}$ represent the sets of vortices and connection edges, respectively. This probabilistic graph consists of three layers: the first layer corresponds to the observation variables, i.e., the measured wavefield $E_{\text{sca}}^{(n)}$; the second layer, called the hidden layer in the literature of machine learning, is composed of the contrast currents $J^{(n)}$; the third layer is associated with the contrast function χ. Note that the variables at the same layer have no direct connection, implying to be statistically independent.

The expectation maximization algorithm is employed to infer the posterior probability of the contrast function χ. Expectation maximization is an iterative solver, which consists
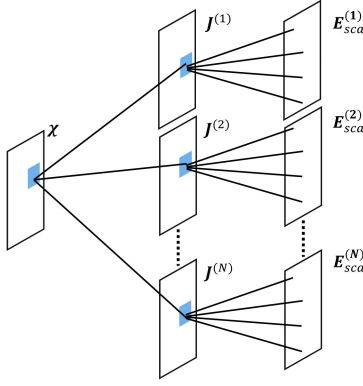
Fig. 2. Three-layer graphical representation of EMIS.

of the alternating between two steps, i.e., E-step and M-step, until convergence [28], [29]. The E-step is intended to find the probability function of latent variables $J^{(n)}$ for estimated χ in last step. In the M-step, the maximum-likelihood solution of χ is achieved based on the probability function of latent variables estimated in E-step. In the following, we will detail the algorithm procedure.

1) E-Step (The Expectation Step): Following the standard expectation maximization strategy [28]–[30], one can arrive at the upper bound of the log-likelihood function $\ln p(E_{\text{sca}}^{(n)}, n = 1, 2, \ldots | \chi)$ of χ as:

$$
\begin{aligned}
&\ln p(E_{\text{sca}}^{(n)}, n = 1, 2, \ldots | \chi) \\
&\leq \mathbb{E}_{J^{(n)} \sim p(J^{(n)}, n=1,2\ldots)} \big[\ln p\big(E_{\text{sca}}^{(n)}, J^{(n)} n = 1, 2, \ldots | \chi\big) \big] \\
&= \mathbb{E}_{J^{(n)} \sim p(J^{(n)}, n=1,2\ldots)} \big[\ln p\big(J^{(n)} n = 1, 2, \ldots | \chi\big) \big] \\
&= \mathbb{E}_{J^{(n)} \sim p(J^{(n)}, n=1,2\ldots)} \\
&\quad \times \left[\eta_s \sum_{n=1}^{N} \big\| J^{(n)} - \text{diag}(\chi)\big(E_{\text{inc}}^{(n)} + G_s J^{(n)}\big) \big\|_2^2 \right]. \quad (7)
\end{aligned}
$$

The holding inequality is used in the first line of (7) [26], [27], and E denotes the operator of mathematical expectation. It is worth mentioning that the prior information on χ can be incorporated by using the following fundamental fact:

$$
\begin{aligned}
&p\big(\chi | E_{\text{sca}}^{(n)}, n = 1, 2, \ldots\big) \\
&\qquad = \ln p\big(E_{\text{sca}}^{(n)}, n = 1, 2, \ldots | \chi\big) + \ln p(\chi)
\end{aligned}
$$

where $p(\chi)$ accounts for the useful prior information on χ, such as the sparse prior. Accordingly, the three-layer graph could be extended for that with more layers. Due to limited paper space, we will leave this part for the future research.

The probability function $p(J^{(n)}, n = 1, 2 \ldots)$ involved in (7) reads

$$
\begin{aligned}
p(J^{(n)}, n = 1, 2 \ldots) &= \prod_{n=1}^{N} p\big(J^{(n)} | \chi, E_{\text{sca}}^{(n)}\big) \\
&= \prod_{n=1}^{N} p\big(E_{\text{sca}}^{(n)}, | J^{(n)}\big) p\big(J^{(n)} | \chi\big). \quad (8)
\end{aligned}
$$

In (8), we have adopted a fundamental fact that the contrast currents $J^{(n)}$ are statistically independent. Two complex-valued Gaussian probability functions involved

in (8) are

$$
p\big(E_{\text{sca}}^{(n)} | J^{(n)}\big) = \left(\frac{\sqrt{\eta_D}}{\pi} \right)^M \exp\left(-\frac{1}{2} \eta_D \big\| E_{\text{sca}}^{(n)} - G_d J^{(n)} \big\|_2^2 \right)
$$
$$(9)$$

and

$$
p(J^{(n)} | \chi) = \left(\frac{\sqrt{\eta_s}}{\pi} \right)^K \exp
$$
$$
\times \left(-\frac{1}{2} \eta_S \big\| J^{(n)} - \text{diag}(\chi)\big(E_{\text{inc}}^{(n)} + G_s J^{(n)}\big) \big\|_2^2 \right)
$$
$$(10)$$

where $\eta_D = 1 / \sum_{n=1}^{N} \| E_{\text{sca}}^{(n)} \|_2^2$ and $\eta_s = 1 / \sum_{n=1}^{N} \| E_{\text{inc}}^{(n)} \|_2^2$ are specified inspired by the CSI method. After substituting (9) and (10) into (8), we can obtain the probability function $p(J^{(n)} | \chi, E_{\text{sca}}^{(n)})$ as

$$
p\big(J^{(n)} | \chi, E_{\text{sca}}^{(n)}\big) = \left(\frac{\sqrt{\eta_D}}{\pi} \right)^M \left(\frac{\sqrt{\eta_s}}{\pi} \right)^K \exp\left(-\frac{1}{2} S_n(J^{(n)}) \right)
$$
$$(11)$$

where $S_n(J^{(n)})$ is defined as

$$
\begin{aligned}
S_n(J^{(n)}) &= \eta_D \big\| E_{\text{sca}}^{(n)} - G_d J^{(n)} \big\|_2^2 \\
&\quad + \eta_S \big\| J^{(n)} - \text{diag}(\chi)\big(E_{\text{inc}}^{(n)} + G_s J^{(n)}\big) \big\|_2^2. \quad (12)
\end{aligned}
$$

Note that for the probability density function of $J^{(n)}$ like (11), it is infeasible to get an analytical expression of the right hand of (7). To tackle this difficulty, we assume $p(J^{(n)} | \chi, E_{\text{sca}}^{(n)})$ to be a delta distribution [23], namely

$$
p\big(J^{(n)} | \chi, E_{\text{sca}}^{(n)},\big) \approx \delta\big(J^{(n)} - J_{\text{MAP}}^{(n)}\big) \quad (13)
$$

where

$$
\begin{aligned}
J_{\text{MAP}}^{(n)} &= \text{argmin}_{J^{(n)}} S_n(J^{(n)}) \\
&= \text{argmin}_{J^{(n)}} \big[\eta_D \big\| E_{\text{sca}}^{(n)} - G_d J^{(n)} \big\|_2^2 + \eta_S \big\| J^{(n)} \\
&\quad - \text{diag}(\chi)\big(E_{\text{inc}}^{(n)} + G_s J^{(n)}\big) \big\|_2^2 \big] = PQ \quad (14)
\end{aligned}
$$

where

$$
P = (\eta_D G_d' G_d + \eta_s (I - \text{diag}(\chi) G_s)' (I - \text{diag}(\chi) G_s))^{-1}
$$
$$
Q = \eta_D G_d' E_{\text{sca}}^{(n)} + \eta_s (I - \text{diag}(\chi) G_s)' \text{diag}(\chi) E_{\text{inc}}^{(n)}.
$$

After substituting (14) into (7), we can estimate the upper bound of the log-likelihood of χ as

$$
\begin{aligned}
&\ln p\big(E_{\text{sca}}^{(n)}, n = 1, 2, \ldots | \chi\big) \\
&\leq \eta_s \sum_{n=1}^{N} \big\| J_{\text{MAP}}^{(n)} - \text{diag}(\chi)\big(E_{\text{inc}}^{(n)} + G_s J_{\text{MAP}}^{(n)}\big) \big\|_2^2. \quad (15)
\end{aligned}
$$

2) M-Step (The Maximum Step): It's straightforward to find the maximum-likelihood solution to χ from (15), which can be derived as

$$
\begin{aligned}
\chi &= \text{argmin}_{\chi} \left[\sum_{n=1}^{N} \big\| J^{(n)} - \text{diag}(\chi)\big(E_{\text{inc}}^{(n)} + G_s J^{(n)}\big) \big\|_2^2 \right] \\
&= \sum_{n=1}^{N} \langle J^{(n)}, J^{(n)} \rangle / \sum_{n=1}^{N} \langle \big(E_{\text{inc}}^{(n)} + G_s J^{(n)}\big) \rangle. \quad (16)
\end{aligned}
$$

TABLE I

ALGORITHM FOR IS-EM

Algorithm 1. IS-EM ($E_{sca}^{(n)}, n = 1, 2, ..., N, \chi$)

Initial Guess: $\chi = 0$
WHILE (Not meet the convergence criterion)
 ■The E-step:
 $P = (\eta_D G_d' G_d + \eta_s (I - diag(\chi)G_s)'(I - diag(\chi)G_s))^{-1}$
  FOR $n = 1: N$
   $Q = \eta_D G_d' E_{sca}^{(n)} + \eta_s (I - diag(\chi)G_s)' diag(\chi)E_{inc}^{(n)}$
   $J^{(n)} = PQ$
  END FOR
 ■The M-step:
 $\chi = \sum_{n=1}^{N} \langle J^{(n)}, J^{(n)} \rangle / \sum_{n=1}^{N} \langle J^{(n)}, (E_{inc}^{(n)} + G_s J^{(n)}) \rangle$
END WHILE

Finally, the whole algorithm procedure (referred to as IS-EM algorithm in this paper) to solve (5) and (6) can be summarized in Table I.

In light of the standard maximum expectation method [23], it can be deduced that Algorithm 1 can be guaranteed to stably converge to a local minimizer of the following cost function:

$$\mathcal{L}(\chi, J^{(n)}, n = 1, 2 \ldots)$$
$$= \eta_D \sum_{n=1}^{N} \left\| E_{sca}^{(n)} - G_d J^{(n)} \right\|_2^2$$
$$+ \eta_s \sum_{n=1}^{N} \left\| J^{(n)} - diag(\chi)(E_{inc}^{(n)} + G_s J^{(n)}) \right\|_2^2. \quad (17)$$

Apparently, the cost function (17) is similar to that used in CSI method. However, the proposed Algorithm 1 differs remarkably from CSI in three important aspects:

First, Algorithm 1 is derived from a probabilistic model of EMIS combined with expectation maximization method, while the CSI method works in a deterministic way. In addition, the convergence property of Algorithm 1 can be naturally guaranteed by the expectation maximization method.

Second, the cost function (17) is a special instance of the proposed probabilistic model in this paper, which specifically corresponds to the use of delta-distribution approximation [23], i.e., (13). In other words, we can obtain more meaningful cost functions other than (17) from the proposed probabilistic model, and we will leave this part for the future work due to the limited paper space.

Third, the implementation procedure of Algorithm 1 is different from the CSI. The contrast currents are updated according to the criterion of maximum posterior probability in Algorithm 1: however, they are updated via the Polak-Ribiere conjugate gradient in CSI.

C. Complex-Valued ADMM Solution to EMIS

Note that Algorithm 1 has two artificial factors, i.e., η_D and η_S, which could be a troublesome issue in some cases. In order to avoid this issue, we explore the spirit of the ADMM technique in minimizing (17). The total electrical field inside the proposed scene is introduced

$$E^{(n)} = E_{inc}^{(n)} + G_s J^{(n)}. \quad (18)$$

TABLE II

ALGORITHM FOR IS-ADMM

Algorithm 2. IS-ADMM ($E_{sca}^{(n)}, n = 1, 2, ..., N, \chi$)

Initial Guess: $\chi = 0, y^{(n)} = 0, E^{(n)} = E_{inc}^{(n)}, n = 1, 2, ..., N$
WHILE (Not meet the convergence criterion)
 $P = ((I - diag(\chi)G_s)'(I - diag(\chi)G_s) + \rho G_d' G_d)^{-1}$
  FOR $n = 1: N$
   ■Update $J^{(n)}$
   $Q = (I - diag(\chi)G_s)' diag(\chi)E_{inc}^{(n)} + \rho G_d' E_{sca}^{(n)} + G_d' y^{(n)}$
   $J^{(n)} = PQ$
   ■Update $y^{(n)}$:
   $y^{(n)} = y^{(n)} + \rho(E_{sca}^{(n)} - G_d J^{(n)})$
  END FOR
 ■ Update χ: $\chi = \sum_{n=1}^{N} \langle J^{(n)}, J^{(n)} \rangle / \sum_{n=1}^{N} \langle J^{(n)}, \Theta^{(n)} \rangle$
END WHILE

Then the problem of minimizing (17) can be casted into a constrained optimization problem as

$$\min_{\chi, J^{(n)}, E^{(n)}} \mathcal{L}_{uc}(\chi, J^{(n)}, E^{(n)}, n = 1, 2 \ldots)$$
$$\text{s.t. } E_{sca}^{(n)} = G_d J^{(n)}, \quad n = 1, 2 \ldots N \quad (19)$$

where the cost function in (19) is

$$\mathcal{L}_{uc}(\chi, J^{(n)}, E^{(n)}, n = 1, 2 \ldots)$$
$$= \frac{1}{2} \sum_{n=1}^{N} \left\| J^{(n)} - diag(\chi)(E_{inc}^{(n)} + G_s J^{(n)}) \right\|_2^2.$$

Usually, the augmented Lagrangian method is a popular technique in treating the real-valued constrained optimization problem. However, all quantities involved in (19) are complex-valued. To overcome this limitation, we extend the conventional augmented Lagrangian method for real value into that for complex value. To this end, the complex-valued augmented Lagrangian optimization problem is casted as

$$\max_{y^{(n)}} \min_{\chi, J^{(n)}, E^{(n)}} \mathcal{L}_{AL}(\chi, J^{(n)}, E^{(n)}, y^{(n)}, n = 1, 2 \ldots) \quad (20)$$

where $\mathcal{L}_{AL}$ is the complex-valued augmented Lagrangian function

$$\mathcal{L}_{AL}(\chi, J^{(n)}, E^{(n)}, y^{(n)}, n = 1, 2 \ldots)$$
$$= \mathcal{L}_{uc}(\chi, J^{(n)}, E^{(n)}, n = 1, 2 \ldots)$$
$$+ \text{Re} \left[\sum_{n=1}^{N} \langle y^{(n)}, E_{sca}^{(n)} - G_d J^{(n)} \rangle \right]$$
$$+ \frac{\rho}{2} \sum_{n=1}^{N} \left\| E_{sca}^{(n)} - G_d J^{(n)} \right\|_2^2. \quad (21)$$

In (21), Re is an operator for calculating the real part of a complex value. Moreover, $y^{(n)}$ represent the Lagrangian complex-valued dual variables, and $\rho > 0$ is the so-called augmented Lagrangian factor. Note that it is not difficult to determine the value of ρ [31]. Then, the same alternating iterative procedure as that used in conventional ADMM is applied to solve (20). The complete algorithm procedure of solving (12) is summarized in Table II. Comparing Algorithm 2 with Algorithm 1, we can see the Algorithm 2 (i.e., IS-ADMM algorithm) is free of artificial factor.
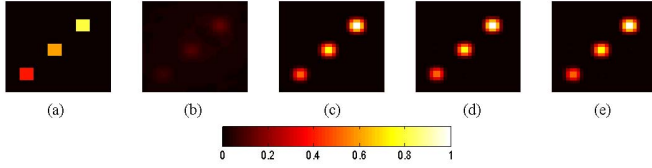
Fig. 3.   Comparisons of the reconstructions by CSI, IS-EM, and IS-ADMM methods. (a) Ground truth. (b) Initial solution obtained by running the BP algorithm. (c) CSI reconstruction. (d) IS-EM result. (e) IS-ADMM result. (The x-axis and y-axis both range from -0.1125 to 0.105 m, and the axis of following figures are the same as that in Fig. 3.)

III. RESULTS AND DISCUSSION

In this section, the performance of the IS-EM and IS-ADMM reconstruction algorithms is investigated for the following three test cases: 1) the reconstruction of three homogeneous lossless dielectric square cylinders against the simulated data; 2) the reconstruction of lossy concentric squares against the simulated data; and 3) the reconstruction of dielectric circular cylinders against the experimental data. For comparison, the corresponding results obtained by the conventional CSI algorithm are provided as well [4]. For the three reconstruction algorithms considered here, their initial solutions are obtained by implementing the conventional back-propagation (BP) algorithm [4], [5]. The simulation parameters of acquisition setup are similar to those used in the database of the Institute Fresnel, Marseille (2005) [32]. The simulation data are generated by running a full-wave solver to Maxwell equations, i.e., CG-FFT [33]. In order to check the tolerance to noise of the proposed methods, the synthetic data computed by the CG-FFT method is corrupted by 30 dB Gaussian random noise in our simulations. Additionally, all simulations are implemented on a personal computer with 3 GHz processor and 1 GB RAM in the MATLAB environment.

A. Case Study 1: Reconstruction of Three Homogeneous Dielectric Square Cylinders Against Simulated Data

We will give the reconstructions of three pure dielectric square cylinders utilizing the IS-EM, IS-ADMM, and CSI methods at the working frequency of 4 GHz. The investigation region under test denoted by Dinv is a square of dimensions 3λ by 3λ (with λ being the working wavelength), which is uniformly discretized into 30 by 30 subsquares. The three pure dielectric square cylinders are 0.4λ by 0.4λ in size, with contrast $\chi = 0.4, 0.6,$ and 0.8, respectively, as shown in Fig. 3(a), where the x-axis is the distance along the x-direction and the y-axis is the distance along the y-direction (the definitions of the x- and y-axes will be used in all of the figures hereafter). In order to avoid the so-called inverse crime, the synthetic data computed by the CG-FFT method are obtained when the investigation domain is discretized into 40×40 cells. Moreover, the Gaussian white noise with 30 dB level has been added to scattered field. Fig. 3(b) shows the result from the BP algorithm, which is used as the initial guess to three iterative algorithms considered here. The reconstructed profiles using the CSI, IS-EM, and IS-ADMM algorithms are plotted in Fig. 3(c)–(e),
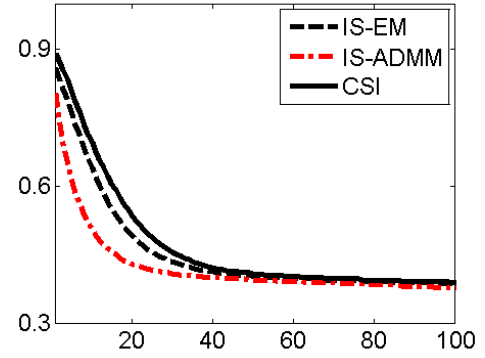

Fig. 4.   Convergence behaviors of three algorithms. The x-axis denotes the iteration index, while the y-axis denotes the MSE error defined by (22).
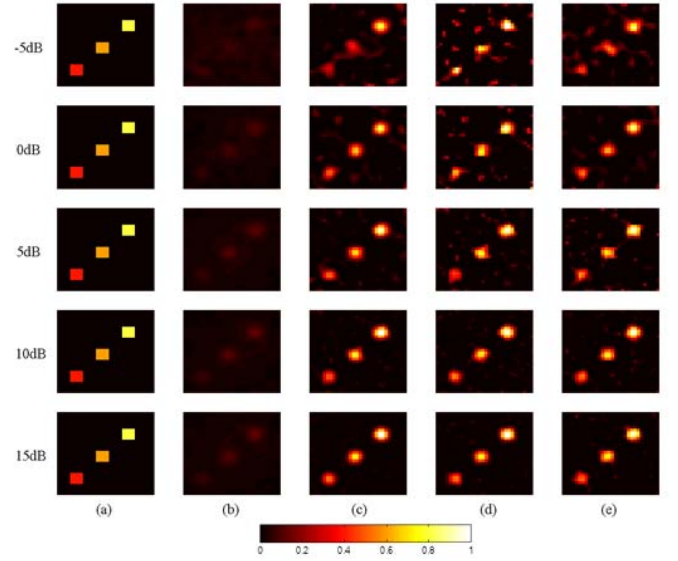

Fig. 5.   Comparisons of the reconstructions by CSI, IS-EM, and IS-ADMM method in different noise, i.e., SNR is $-5, 0, 5, 10,$ and 15 dB, respectively. (a) Ground truth. (b) Initial solution obtained by running the BP algorithm. (c) CSI reconstructions. (d) IS-EM results. (e) IS-ADMM reconstructions.

respectively. Furthermore, the convergence behaviors of the aforementioned three iterative algorithms are compared in Fig. 4. Here, the error is defined as the normalized error, i.e., the mean-square error (MSE), of the reconstructed profile with respect to the exact one is introduced and defined as

$$\text{MSE} = \frac{\|\chi^{\text{rec}} - \chi^{\text{exc}}\|_2}{\|\chi^{\text{exc}}\|_2} \qquad (22)$$

where χ^{rec} is the reconstructed contrast function, and χ^{exc} is the exact one. In the reconstruction process, we choose to stop the iteration when no obvious improvement of the reconstructed image is observed if further iterations are applied. From this set of results, we can see that both IS-ADMM and IS-EM are faster than conventional CSI method, which does make sense since the proposed methods behave as second-order iteration, however, CSI is a first-order iteration method. Overall, the IS-ADMM method performs better than the IS-EM and CSI methods, partly because the IS-ADMM method has no artificial factors involved, and the negative effect on the reconstruction from artificial factors has been omitted.

TABLE III
MSEs of Reconstructions Corresponding to Fig. 5

	BP	CSI	IS-EM	IS-ADMM
-5dB	0.9414	0.5817	0.5677	0.5674
0dB	0.8888	0.4857	0.4750	0.4633
5dB	0.8837	0.4280	0.4180	0.3935
10dB	0.8829	0.3886	0.3746	0.3871
15dB	0.8805	0.3796	0.3459	0.3231

TABLE IV
MSEs of Real and Imagery Reconstructions
Corresponding to Fig. 9

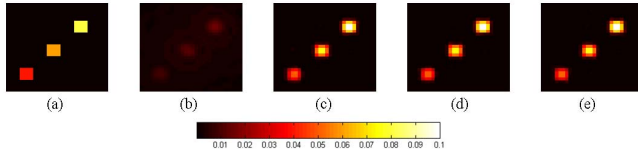|	BP	CSI	IS-EM	IS-ADMM
Real	0.9539	0.1544	0.1434	0.1384
Imag	0.5137	0.1827	0.1567	0.1451

Fig. 6. Comparisons of the reconstructions by CSI, IS-EM, IS-ADMM methods when the contrast function χ of triple cylinders are set as 0.04, 0.06, and 0.08. (a) Ground truth. (b) Initial solution obtained by running the BP algorithm. (c) CSI reconstruction. (d) IS-EM result. (e) IS-ADMM result. Here, the corresponding MSEs of (b)–(e) is 0.8679, 0.2797, 0.2549, and 0.2724, respectively.
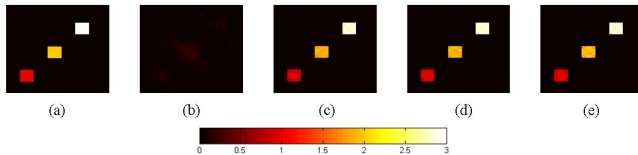

Fig. 7. Comparisons of the reconstructions by CSI, IS-EM, IS-ADMM methods when contrast functions χ of triple cylinders are set as 1, 2, and 3. (a) Ground truth. (b) Initial solution obtained by running the BP algorithm. (c) CSI reconstruction. (d) IS-EM result. (e) IS-ADMM result. Here, the corresponding MSEs of (b)–(e) is 0.9608, 0.0529, 0.0442, and 0.0459, respectively.

To examine the robustness of the proposed methods to the noise, different degrees of noise, i.e., SNR = −5, 0, 5, 10, and 15 dB, have been added to the scattered field. The reconstructions using different methods are shown in Fig. 5. Meanwhile, the corresponding MSE are listed in Table III. From these results, a conclusion can be drawn that the proposed IS-EM and IS-ADMM could outperform than CSI method in different SNRs scenario.

Several experiments, in which the contrast function χ of objects is, respectively, set as (0.04, 0.06, 0.08), (1, 2, 3), and (4, 4.5, 5), are further explored to verify the validity of the proposed algorithm when the objects are weaker or stronger scatterer shown in Figs. 6–8. Other simulation settings are as same as that used in Fig. 3. The reconstructions show that the proposed methods are both appropriate for weak scatter imaging problem and nonweak scatter imaging problem, as illustrated in Figs. 6–8.
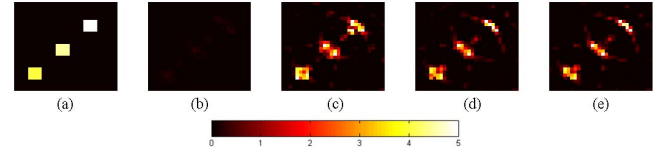

Fig. 8. Comparisons of the reconstructions by CSI, IS-EM, IS-ADMM methods when contrast functions χ of triple cylinders are set as 4, 4.5, and 5. (a) Ground truth. (b) Initial solution obtained by running the BP algorithm. (c) CSI reconstruction. (d) IS-EM result. (e) IS-ADMM result. Here, the corresponding MSEs of (b)–(e) is 0.9971, 0.6873, 0.6973, and 0.7017, respectively.
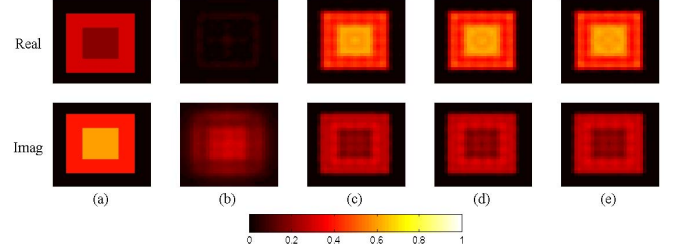

Fig. 9. (a) Ground truth of a lossy concentric square cylinders. Real part (Top row). Imaginary part (Bottom row). (b) Reconstructions using BP algorithm. (c) Reconstructions using CSI. (d) Reconstructions using IS-EM. (e) Reconstructions using IS-ADMM.

B. Case Study 2: Reconstruction of Lossy Concentric Square Cylinders Against the Simulated Data

As a second numerical example, the reconstruction of the lossy concentric square cylinders is investigated under the same computation condition of Example 1. The scattering object to be reconstructed, shown in Fig. 9(a), consists of concentric square cylinders: an inner cylinder of λ by λ, with a complex contrast of $\chi = 0.6 + 0.2i$, surrounded by an outer cylinder, of 2λ by 2λ, with a contrast of $\chi = 0.4 + 0.3i$. Fig. 9(b) shows the real (top) and imaginary (bottom) parts of the reconstructions using the standard BP algorithm, which serves as the initial guess for the CSI, IS-EM, and IS-ADMM algorithms under investigation. Fig. 9(c)–(e) shows the reconstructed profiles using the CSI, IS-EM, and IS-ADMM methods. The MSEs of real and imaginary of reconstructions corresponding to Fig. 9 are reported in Table IV. Similarily, the convergence behaviors of the three reconstruction algorithms are compared in Fig. 10 as well. From this set of results, similar conclusions can be drawn as in case study 1.

C. Case Study 3: Reconstruction With Laboratory-Controlled Data

Finally, the last part of the numerical validation aims to assess the effectiveness of the IS-EM and IS-ADMM algorithms when experimental scattering data are employed. For comparison, the reconstructions by the conventional CSI method are provided as well. The experimental data were kindly provided by the Institute Fresnel, Marseille (2005), and the configuration of the measurement setup has been carefully described in [32]. With reference to Fig. 11(a), the reconstruction employs one set of TM-polarized data, FoamDielExt [32], where white for plastic with relative permittivity $3 \pm 0.3i$, gray for foam with relative permittivity $1.45 \pm 0.15i$, with a working frequency of 4 GHz. All obstacles are assumed to
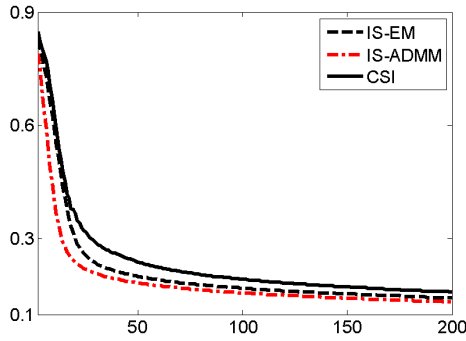
Fig. 10. Convergence behaviors of three reconstruction methods considered in this paper for case study 2. Here, the x-axis denotes the iteration index, and the y-axis is the MSE defined by (22).
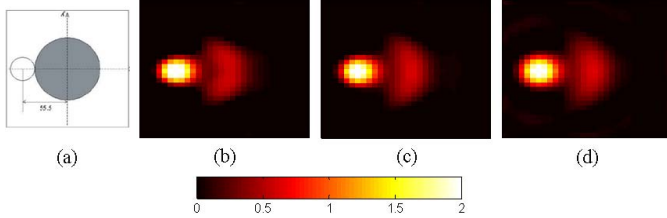

Fig. 11. (a) Target database description of FoamDielExt [28]. The white for plastic with relative permittivity 3 ± 0.3 (the contrast is 2 ± 0.3), and the gray for foam with relative permittivity 1.45 ± 0.15 (the contrast is 0.45 ± 0.15). (b) IS-EM reconstruction. (c) IS-ADMM result. (d) CSI result.

be infinitely long in one direction such that a 2-D assumption can be made and that the line excited source parallel to the z-axis can be modeled [32]. For all computations, the investigation region D_{inv} consists of a square with sides length of $d = 0.17$ m and discretized into a series of subsquares with side length of 0.05λ. It is noted that the scatters are beyond the weak scattering regime due to $k\sigma\chi > 4.0$ for above three databases. Similar to previous cases, the initial guess for the CSI, IS-EM, and IS-ADMM methods is obtained from the BP algorithm. Fig. 11(b)–(d) reports the reconstructions of the real part of contrast function for FoamDielExt by the IS-EM [Fig. 11(b)], IS-ADMM [Fig. 11(c)], and CSI [Fig. 11(d)], respectively. In addition, the MSEs for the IS-EM, IS-ADMM, and CSI methods are 0.20, 0.13, and 0.28, respectively. Once again, similar conclusions can be obtained as previous two cases.

IV. CONCLUSION

In this paper, we present a full probabilistic model of the classical EMIS, from which two novel algorithms have been proposed, i.e., IS-EM and IS-ADMM. As a byproduct, the complex-valued ADMM has been developed to extend the conventional ADMM method solving only for a real-valued optimization problem into that for a complex-valued one. Exemplary numerical and experimental tests have been conducted to validate the proposed procedures, where the experimental data are kindly provided by the Institute Fresnel in Marseille (France). In this paper, two areas of the EMIS and the machine learning are brought together, which opens insightful perspectives on future investigations. First of all, more in-depth efforts should be made when dealing with the specific practical applications, for instance, penetration radar

imaging. Second, the graphic model shown in Fig. 2 can be extended to account for some prior knowledge on both the contrast source and the contrast function, for instance, the sparse prior on the contrast function, and the joint sparse prior on the contrast sources [15]–[24], by which the reconstruction performance could be significantly improved. Finally, the iteration procedure, either in Algorithm 1 or 2, could be further accelerated by leveraging the learning strategy as done in [34] and [35]. As a result, we will report further progress in the near future.

REFERENCES

[1] R. Zoughi, *Microwave Non-Destructive Testing and Evaluation*. Amsterdam, The Netherlands: Kluwer, 2000, pp. 6–7.

[2] D. Lesselier and T. Habashy, "Electromagnetic imaging and inversion on the earth's subsurface," *Inverse Problems*, vol. 16, no. 5, pp. U3–U5, 2000.

[3] P. M. Meaney, M. W. Fanning, D. Li, S. P. Poplack, and K. D. Paulsen, "A clinical prototype for active microwave imaging of the breast," *IEEE Trans. Microw. Theory Techn.*, vol. 48, no. 11, pp. 1841–1853, Nov. 2000.

[4] P. M. van den Berg and A. Abubakar, "Contrast source inversion method: State of art," *Prog. Electromagn. Res.*, vol. 34, no. 11, pp. 189–218, 2001.

[5] L. Li, H. Zheng, and F. Li, "Two-dimensional contrast source inversion method with phaseless data: TM case," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 6, pp. 1719–1735, Jun. 2009.

[6] W. C. Chew and Y. M. Wang, "Reconstruction of two-dimensional permittivity distribution using the distorted Born iterative method," *IEEE Trans. Med. Imag.*, vol. 9, no. 2, pp. 218–225, Jun. 1990.

[7] L. Li, W. Zhang, and F. Li, "Tomographic reconstruction using the distorted Rytov iterative method with phaseless data," *IEEE Geosci. Remote Sens. Lett.*, vol. 5, no. 3, pp. 479–483, Jul. 2008.

[8] M. Pastorino, *Microwave Imaging*. Hoboken, NJ, USA: Wiley, 2010, pp. 18–20.

[9] D. Colton and R. Kress, *Inverse Acoustic and Electromagnetic Scattering Theory*, 2nd ed. Berlin, Germany: Springer, 1998, pp. 7–10.

[10] M. Bertero and P. Boccacci, *Introduction to Inverse Problems in Imaging*. London, U.K.: Institute of Physics Pub., 1998, pp. 1–10.

[11] P. Rocca, M. Benedetti, M. Donelli, D. Franceschini, and A. Massa, "Evolutionary optimization as applied to inverse scattering problems," *Inverse Problems*, vol. 25, no. 12, pp. 123003-1–123003-41, 2009.

[12] P. Rocca, G. Oliveri, and A. Massa, "Differential evolution as applied to electromagnetics," *IEEE Antennas Propag. Mag.*, vol. 53, no. 1, pp. 38–49, Feb. 2011.

[13] M. Pastorino, "Stochastic optimization methods applied to microwave imaging: A review," *IEEE Trans. Antennas Propag.*, vol. 55, no. 3, pp. 538–548, Mar. 2007.

[14] O. Lee, J. M. Kim, Y. Bresler, and J. C. Ye, "Compressive diffuse optical tomography: Noniterative exact reconstruction using joint sparsity," *IEEE Trans. Med. Imag.*, vol. 30, no. 5, pp. 1129–1142, May 2011.

[15] L. Poli, G. Oliveri, P. Rocca, and A. Massa, "Bayesian compressive sensing approaches for the reconstruction of two-dimensional sparse scatterers under TE illuminations," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 5, pp. 2920–2936, May 2013.

[16] L. Poli, G. Oliveri, F. Viani, and A. Massa, "MT—BCS-based microwave imaging approach through minimum-norm current expansion," *IEEE Trans. Antenna Propag.*, vol. 61, no. 9, pp. 4722–4732, Sep. 2013.

[17] R. Scapaticci, P. Kosmas, and L. Crocco, "Wavelet-based regularization for robust microwave imaging in medical applications," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 4, pp. 1195–1202, Apr. 2015.

[18] N. Anselmi, M. Salucci, G. Oliveri, and A. Massa, "Wavelet-based compressive imaging of sparse targets," *IEEE Trans. Antenna Propag.*, vol. 63, no. 11, pp. 4889–4899, Nov. 2015.

[19] V. Pascazio and G. Ferraiuolo, "Statistical regularization in linearized microwave imaging through MRF-based MAP estimation: Hyperparameter estimation and image computation," *IEEE Trans. Image Process.*, vol. 12, no. 5, pp. 572–582, May 2003.

[20] L. Poli, G. Oliveri, and A. Massa, "Microwave imaging within the first-order Born approximation by means of the contrast-field Bayesian compressive sensing," *IEEE Trans. Antennas Propag.*, vol. 60, no. 6, pp. 2865–2879, Jun. 2012.

LI _et al._: PROBABILISTIC MODEL FOR NONLINEAR EMIS 5991

[21] G. Oliveri, N. Anselmi, and A. Massa, "Compressive sensing imaging of non-sparse 2D scatterers by a total-variation approach within the Born approximation," _IEEE Trans. Antennas Propag._, vol. 62, no. 10, pp. 5157–5170, Oct. 2014.

[22] L. Guo and A. M. Abbosh, "Microwave imaging of nonsparse domains using Born iterative method with wavelet transform and block sparse Bayesian learning," _IEEE Trans. Antennas Propag._, vol. 63, no. 11, pp. 4877–4888, Nov. 2015.

[23] G. Oliveri, P. Rocca, and A. Massa, "A Bayesian-compressive-sampling-based inversion for imaging sparse scatterers," _IEEE Trans. Geosci. Remote Sens._, vol. 49, no. 10, pp. 3993–4006, Oct. 2011.

[24] S. Ji, D. Dunson, and L. Carin, "Multitask compressive sensing," _IEEE Trans. Signal Process._, vol. 57, no. 1, pp. 92–106, Jan. 2009.

[25] U. S. Kamilov _et al._, "Learning approach to optical tomography," _Optica_, vol. 2, no. 6, pp. 517–522, 2015.

[26] A. J. Izeman, _Modern Multivariate Statistical Techniques: Regression, Classification, and Manifold Learning_. Berlin, Germany: Springer, 2008.

[27] K. P. Murphy, _Machine Learning: A Probabilistic Perspective_. Cambridge, MA, USA: MIT Press, 2012.

[28] M. A. T. Figueiredo and R. D. Nowak, "An EM algorithm for wavelet-based image restoration," _IEEE Trans. Image Process._, vol. 12, no. 8, pp. 906–916, Aug. 2003.

[29] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," _J. Roy. Stat. Soc. B (Methodol.)_, vol. 39, no. 1, pp. 1–38, 1977.

[30] N. Friedman, "The Bayesian structural EM algorithm," in _Proc. 14th Conf. Uncertainty Aritif. Intell._, San Francisco, CA, USA, 1998.

[31] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," _Found. Trends Mach. Learn._, vol. 3, no. 1, pp. 1–122, Jan. 2011.

[32] J.-M. Geffrin, P. Sabouroux, and C. Eyraoud, "Free space experimental scattering database continuation: Experimental set-up and measurement precision," _Inverse Problems_, vol. 21, no. 6, pp. S117–S130, 2005.

[33] M. F. Catedra, R. P. Torres, J. Basterrechea, and E. Gago, _The CG-FFT Method: Application of Signal Processing Techniques to Electromagnetics_. Boston, MA, USA: Artech House, 1995.

[34] M. Borgerding and P. Schniter. (Jul. 2016). "Onsager-corrected deep learning for sparse linear inverse problems," p. 5. [Online]. Available: https://arxiv.org/abs/1607.05966

[35] K. Gregor and Y. LeCun, "Learning fast approximations of sparse coding," in _Proc. Int. Conf. Int. Conf. Mach. Learn._, 2010, pp. 399–406.

Lianlin Li, photograph and biography not available at the time of publication.

Long Gang Wang, photograph and biography not available at the time of publication.

Jun Ding, photograph and biography not available at the time of publication.

P. K. Liu, photograph and biography not available at the time of publication.

M. Y. Xia, photograph and biography not available at the time of publication.

Tie Jun Cui, photograph and biography not available at the time of publication.