Sparse Nonlinear Electromagnetic Imaging Accelerated With Projected Steepest Descent Algorithm

Abdulla Desmal, *Student Member, IEEE*, and Hakan Bağcı, Senior *Member, IEEE*

Abstract—An efficient electromagnetic inversion scheme for imaging sparse 3-D domains is proposed. The scheme achieves its efficiency and accuracy by integrating two concepts. First, the nonlinear optimization problem is constrained using L_0 or L_1-norm of the solution as the penalty term to alleviate the ill-posedness of the inverse problem. The resulting Tikhonov minimization problem is solved using nonlinear Landweber iterations (NLW). Second, the efficiency of the NLW is significantly increased using a steepest descent algorithm. The algorithm uses a projection operator to enforce the sparsity constraint by thresholding the solution at every iteration. Thresholding level and iteration step are selected carefully to increase the efficiency without sacrificing the convergence of the algorithm. Numerical results demonstrate the efficiency and accuracy of the proposed imaging scheme in reconstructing sparse 3-D dielectric profiles.

Index Terms—Accelerated steepest descent, electromagnetic imaging, electromagnetic inverse scattering, Landweber iterations, nonlinear ill-posed problem, numerical methods, sparsity.

I. INTRODUCTION

NUMERICAL schemes capable of efficiently and accurately solving the electromagnetic inverse scattering problem [1]–[4] are indispensable in various fields of engineering such as tomography [4], nondestructive testing [5], through-wall imaging [6], [7], crack/mine detection [8], [9], and hydrocarbon reservoir exploration and monitoring [10], [11]. In this context, the electromagnetic inverse scattering problem is defined as finding material properties (such as permittivity and conductivity) in an unknown investigation domain from scattered fields measured away from the domain itself [1]–[4]. Developing efficient and accurate numerical tools for solving this problem is a challenging task. First, scattered fields are a nonlinear function of material properties [1]–[4]. Second, inverse problem is ill-posed because scattered field measurements are noisy and taken at a finite set of points and the scattered field integral operator has a smoothening effect [1]–[4]. It should also be noted here that the difficulty of solving the electromagnetic inverse scattering problem increases significantly as the dimension

of investigation domain increases, i.e., many of the methods developed for reconstruction of material properties in 2-D domains do not work as efficiently and/or accurately for reconstruction of material properties in 3-D domains [1], [2].

Electromagnetic inversion algorithms can be grouped into two main categories: stochastic and deterministic schemes [1]. Stochastic schemes rely on (nonlinear) global optimization techniques to solve the electromagnetic inverse scattering problem and are known to be stable and convergent even when they are applied to highly nonlinear problems [12]–[18]. Despite their advantages, stochastic schemes have become popular only recently because of their high computational requirements that can be met by modern parallel computers. The inversion algorithm developed in this paper belongs to the group of deterministic schemes; therefore, the reader is referred to [1] and [12]–[18] and the references therein for a more detailed discussion on accuracy, efficiency, and applicability of stochastic schemes.

Deterministic electromagnetic inversion algorithms can be classified based on the level of nonlinearity that they can accurately account for [1]. The scattering operator's nonlinearity increases as the contrast of the domain increases. Here, contrast is defined as the difference between the relative permittivities of the investigation domain and the background medium. For scatterers with low contrast, replacing the nonlinear scattering operator with its first-order (linear) approximations (e.g., first-order Born and Rytov approximations [19]) yields results accurate enough for many practical examples. As the level of nonlinearity increases, to maintain the stability and the accuracy of the inversion, more accurate approximations to the nonlinear scattering operator have to be used. These approximations can be higher order, like in the extended-Born or the second-order Born schemes [20], or can be constructed iteratively like in the Born iterative method [21]. For scatterers with high contrast, "fully" nonlinear schemes such as the distorted-Born method [22] and several other methods, which make use of Newton iterations [23], nonlinear conjugate gradient scheme [24], [25], and nonlinear steepest descent algorithm [26], are more accurate and efficient.

The electromagnetic inverse scattering problem is constructed in such way that the data misfit between the (noisy) measured fields and the scattered fields generated by the mathematical model is minimized [1]–[4]. Typically, the solution of this optimization problem (reconstructed image) is designed to be the investigation domain's contrast profile.

The ill-posedness of the inverse problem is alleviated by using a regularization scheme that minimizes the weighted sum of the data misfit and a penalty term. If the penalty term is selected as the L_2-norm of the solution, the (regularized) optimization problem converges to a smooth solution [3]. Consequently, such regularization schemes may not produce accurate results if the investigation domain includes sharp variations/discontinuities and/or sparse content [27], [28]. Regularization schemes using L_0/L_1-norm of the solution as the penalty term have also been developed [29]–[34]. These schemes "promote" sharpness and sparseness in the solution and produce more accurate reconstructions for domains with sharp variations/discontinuities and/or sparse content. It should be noted here that the electromagnetic inversion algorithms developed in [29]–[34] assume sparseness in the actual investigation domain (i.e., spatial sparseness meaning that contrast profile is nonzero only in a small portion of the investigation domain). One also can carry out the reconstruction in a different domain. For example, in [35], an electromagnetic inversion algorithm, which makes use of a sparse optimization scheme operating on the wavelets coefficients of the contrast profile, has been proposed. Another example is the algorithm proposed in [36], which reconstructs the derivative of the contrast profile. Other "sparsity-enhanced" inversion algorithms include compressive sensing techniques that have been used to exploit the fact that measurements can be sampled below the Nyquist rate [37]. This reduces the computation cost without sacrificing from the accuracy of the reconstruction [37].

In [33], a (spatial) sparseness-promoting regularization scheme is used together with an inexact Newton method for 2-D nonlinear electromagnetic imaging. Similarly, in [34], a scheme, which makes use of thresholded nonlinear Landweber iterations (NLW) to solve the sparse Tikhonov minimization problem, has been developed for the same purpose. The numerical results obtained using either one of these schemes have demonstrated that many object features, which cannot be detected by smoothness-promoting regularization schemes, can be reconstructed under sparseness-promoting regularization. The results have also shown that these inversion algorithms are more robust to noise and maintain their stability for investigation domains with higher contrast values (i.e., stronger nonlinearity). Last but not least, their efficiency increases as the investigation domain becomes sparser suggesting that their benefits will be more pronounced when they are used on 3-D domains that are oftentimes significantly sparser than their 2-D counterparts. Despite the advantages listed above, schemes proposed in [33] and [34] have not been extended to 3-D electromagnetic imaging. This is simply due to fact that convergence rate of these algorithms is not high enough to image 3-D domains within a reasonable computation time.

To overcome this bottleneck, in this paper, a new scheme for solving the electromagnetic inverse scattering problem on spatially sparse 3-D investigation domains is developed. This scheme integrates two concepts to achieve accuracy and efficiency. First, the nonlinear optimization problem is constrained using the L_0/L_1-norm of the solution as the penalty term. The resulting Tikhonov minimization problem
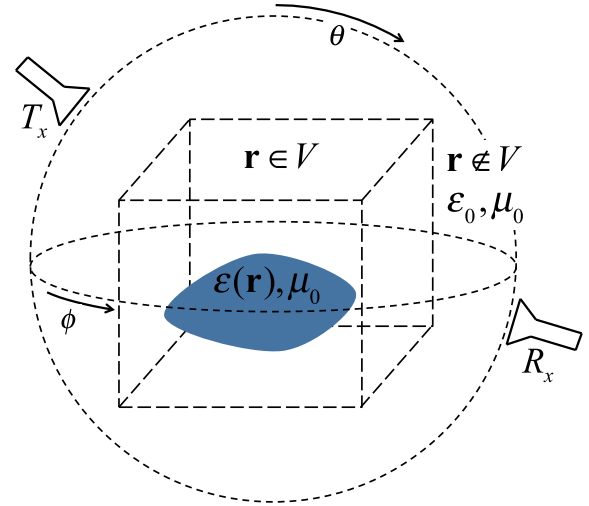

Fig. 1. Description of the 3-D electromagnetic inverse scattering problem.

is solved using the NLW. Second, the NLW is accelerated using a steepest descent algorithm [38], [39]. The sparsity constraint is enforced by thresholding the solution with a projection operator that is applied at every iteration. The resulting projected steepest descent algorithm ensures efficient and convergent solution using carefully selected thresholding level and iteration step. It should also be noted here that this 3-D electromagnetic inversion scheme shares the same benefits as its 2-D counterparts (in comparison with schemes making use of L_2-norm regularization): It is more robust to noise and can handle higher contrast levels without sacrificing from stability, and the reconstructed images are sharper and oftentimes more accurate. Numerical experiments, which demonstrate the accuracy, efficiency, and applicability of the proposed scheme in reconstructing 3-D sparse profiles, are presented.

II. FORMULATION

A. Discretization of the Forward Model

Let V denote the volumetric support of a 3-D investigation domain residing in an unbounded homogenous background medium (Fig. 1). Permittivity and permeability of the investigation domain and the background medium are represented with $\{\varepsilon_r(\mathbf{r})\varepsilon_0 - j\sigma(\mathbf{r})/\omega, \mu_0\}$ and $\{\varepsilon_0, \mu_0\}$, respectively. Here, $\sigma(\mathbf{r})$ and $\varepsilon_r(\mathbf{r})$ are the conductivity and relative permittivity in V. It is assumed that the investigation domain is surrounded by N^R number of receivers, which are located at $\mathbf{r}_m^R$, $m = 1, \ldots, N^R$, and a single transmitter. The transmitter generates the incident field $\mathbf{E}^{\text{inc}}(\mathbf{r})$. Upon excitation, equivalent electric current density $\mathbf{J}(\mathbf{r})$ is induced on V; $\mathbf{J}(\mathbf{r}) = j\omega\varepsilon_0\tau(\mathbf{r})\mathbf{E}(\mathbf{r})$, where $\tau(\mathbf{r}) = \chi(\mathbf{r}) - j\sigma(\mathbf{r})/(\omega\varepsilon_0)$ and $\chi(\mathbf{r}) = \varepsilon_r(\mathbf{r}) - 1$ are the contrast and susceptibility in V, and $\mathbf{E}(\mathbf{r})$ is the total field. In return, $\mathbf{J}(\mathbf{r})$ generates the scattered field $\mathbf{E}^{\text{sca}}(\mathbf{r})$ [40]

$$\mathbf{E}^{\text{sca}}(\mathbf{r}) = -j\omega\mu_0 \int_V \bar{\mathbf{G}}(\mathbf{r}, \mathbf{r}') \cdot \mathbf{J}(\mathbf{r}')dv'$$

$$= k_0^2 \int_V \bar{\mathbf{G}}(\mathbf{r}, \mathbf{r}') \cdot \mathbf{E}(\mathbf{r}')\tau(\mathbf{r}')dv'. \quad (1)$$

Here, $\bar{\mathbf{G}}(\mathbf{r}, \mathbf{r}')$ is the dyadic Green function of the background medium and expressed as

$$\bar{\mathbf{G}}(\mathbf{r}, \mathbf{r}') = \left[\bar{\mathbf{I}} + \frac{\nabla \nabla}{k_0^2}\right] g(\mathbf{r}, \mathbf{r}')$$

where $\bar{\mathbf{I}}$ is the identity dyad, $g(\mathbf{r}, \mathbf{r}') = e^{-jk_0|\mathbf{r}-\mathbf{r}'|}/(4\pi|\mathbf{r}-\mathbf{r}'|)$ is the scalar Green function, and $k_0 = \omega\ (\varepsilon_0\mu_0)^{1/2}$ is the wavenumber. Total, scattered, and incident fields satisfy $\mathbf{E}(\mathbf{r}) = \mathbf{E}^{\text{inc}}(\mathbf{r}) + \mathbf{E}^{\text{sca}}(\mathbf{r})$, which leads to [40]

$$\mathbf{E}^{\text{inc}}(\mathbf{r}) = \mathbf{E}(\mathbf{r}) - k_0^2 \int_V \bar{\mathbf{G}}(\mathbf{r}, \mathbf{r}') \cdot \mathbf{E}(\mathbf{r}')\tau(\mathbf{r}')dv'. \quad (2)$$

Enforcing (2) on V, i.e., for $\mathbf{r} \in V$, and evaluating (1) at the receiver locations $\mathbf{r}_m^R$, $m = 1, \ldots, N^R$, provide a relation between τ $(\mathbf{r})$, $\mathbf{r} \in V$, and $\mathbf{E}(\mathbf{r}_m^R)$, $m = 1, \ldots, N^R$, which is termed as the forward model. Forward model is discretized as described next.

First, V is divided into N number of equal cubic cells with dimension Δd. Total field components, $E_u(\mathbf{r})$, $u \in \{x, y, z\}$, and the contrast τ $(\mathbf{r})$ are approximated using

$$E_u(\mathbf{r}) = \sum_{n=1}^{N} p_n(\mathbf{r})\{\bar{E}_u\}_n$$

$$\tau(\mathbf{r}) = \sum_{n=1}^{N} p_n(\mathbf{r})\{\bar{t}\}_n. \quad (3)$$

Here, $\bar{E}_u$ and $\bar{t}$ are $N \times 1$ vectors with entries $\{\bar{E}_u\}_n = E_u(\mathbf{r}_n)$ and $\{\bar{t}\}_n = \tau(\mathbf{r}_n)$, where $\mathbf{r}_n$, $n = 1, \ldots, N$, are the centers of the cells. In (3), $p_n(\mathbf{r})$ is the pulse basis function defined as

$$p_n(\mathbf{r}) = \begin{cases} 1 & \mathbf{r} \in V_n \\ 0 & \text{otherwise} \end{cases}$$

where V_n is the volumetric support of the nth cell. Inserting (3) into (2) and evaluating the resulting equation at $\mathbf{r}_m$, $m = 1, \ldots, N$ yield

$$\bar{E}^{\text{inc}} = (\bar{I}_{3N} - \bar{F}D\{\bar{P}\bar{t}\})\bar{E}. \quad (4)$$

Here, $\bar{E} = [\bar{E}_x, \bar{E}_y, \bar{E}_z]^t$ and $\bar{E}^{\text{inc}} = [\bar{E}_x^{\text{inc}}, \bar{E}_y^{\text{inc}}, \bar{E}_z^{\text{inc}}]^t$ are $3N \times 1$ vectors, where $\bar{E}_u$ and $\bar{E}_u^{\text{inc}}$ are $N \times 1$ vectors with entries $\{\bar{E}_u\}_n = E_u(\mathbf{r}_n)$ and $\{\bar{E}_u^{\text{inc}}\}_n = E_u^{\text{inc}}(\mathbf{r}_n)$, $n = 1, \ldots, N$, respectively. $\bar{P} = [\bar{I}_N, \bar{I}_N, \bar{I}_N]^t$ is a $3N \times N$ matrix, $\bar{I}_{3N}$ and $\bar{I}_N$ are $3N \times 3N$ and $N \times N$ identity matrices, $D\{.\}$ is an operator that converts the input vector into a diagonal matrix with entries equal to those of the input vector, and finally $\bar{F}$ is a $3N \times 3N$ matrix with $N \times N$ blocks $\bar{G}_{uv}$, $\{u, v\} \in \{x, y, z\}$

$$\bar{F} = \begin{bmatrix} \bar{G}_{xx} & \bar{G}_{xy} & \bar{G}_{xz} \\ \bar{G}_{yx} & \bar{G}_{yy} & \bar{G}_{yz} \\ \bar{G}_{zx} & \bar{G}_{zy} & \bar{G}_{zz} \end{bmatrix}. \quad (5)$$

In (5), the entries of $\bar{G}_{uv}$ are given by

$$\{\bar{G}_{uv}\}_{mn} = \begin{cases} \int_{V_n} k_0^2 g(\mathbf{r}_m, \mathbf{r}') + \partial_u \partial_v g(\mathbf{r}, \mathbf{r}')\big|_{\mathbf{r}=\mathbf{r}_m} dv', & u = v \\ \int_{V_n} \partial_u \partial_v g(\mathbf{r}, \mathbf{r}')\big|_{\mathbf{r}=\mathbf{r}_m} dv', & u \neq v. \end{cases}$$

Inserting (3) into (1) and evaluating the resulting equation at the receiver locations $\mathbf{r}_m^R$, $m = 1, \ldots, N^R$, yield

$$\bar{E}^{\text{sca}} = \bar{H}D\{\bar{E}\}\bar{P}\bar{t} \quad (6)$$

where $\bar{E}^{\text{sca}} = [\bar{E}_x^{\text{sca}}, \bar{E}_y^{\text{sca}}, \bar{E}_z^{\text{sca}}]^t$ is a $3N^R \times 1$ vector, where $\bar{E}_u^{\text{sca}}$ is an $N \times 1$ vector with entries $\{\bar{E}_u^{\text{sca}}\}_m = E_u^{\text{sca}}(\mathbf{r}_m^R)$, $n = 1, \ldots, N^R$, and $\bar{H}$ is a $3N^R \times 3N$ matrix with $N^R \times N$ blocks $\bar{G}_{uv}^R$, $\{u, v\} \in \{x, y, z\}$

$$\bar{H} = \begin{bmatrix} \bar{G}_{xx}^R & \bar{G}_{xy}^R & \bar{G}_{xz}^R \\ \bar{G}_{yx}^R & \bar{G}_{yy}^R & \bar{G}_{yz}^R \\ \bar{G}_{zx}^R & \bar{G}_{zy}^R & \bar{G}_{zz}^R \end{bmatrix}. \quad (7)$$

In (7), the entries of $\bar{G}_{uv}^R$ are given by

$$\{\bar{G}_{uv}^R\}_{mn} = \begin{cases} \int_{V_n} k_0^2 g(\mathbf{r}_m^R, \mathbf{r}') + \partial_u \partial_v g(\mathbf{r}, \mathbf{r}')\big|_{\mathbf{r}=\mathbf{r}_m^R} dv', & u = v \\ \int_{V_n} \partial_u \partial_v g(\mathbf{r}_m^R, \mathbf{r}')\big|_{\mathbf{r}=\mathbf{r}_m} dv', & u \neq v. \end{cases}$$

Inverting (4) for $\bar{E}$ and inserting the resulting expression in (6) yield the discretized complete nonlinear electromagnetic forward operator

$$f(\bar{t}) = \bar{E}^{\text{sca}} = \bar{H}D\{(\bar{I}_{3N} - \bar{F}D\{\bar{P}\bar{t}\})^{-1}\bar{E}^{\text{inc}}\}\bar{P}\bar{t}. \quad (8)$$

Several observations about the discretized forward operator $f(\cdot)$ are in order. First, it is clearly seen from (8) that $f(\bar{t})$ is a nonlinear function of the contrast $\bar{t}$. Second, forward operator $f(\bar{t})$ is constructed under the assumption of single transmitter operated with a single polarization. However, usually measurements are carried out with multiple transmitters (located at different positions) operated with different polarizations. Let vector $\bar{E}_{t,s}^{\text{inc}}$ store the samples of the incident field generated by the transmitter t, $t = 1, 2, \ldots, N^T$, where N^T is the number of transmitters, and with polarization s, $s = 1, 2, \ldots, N^P$, where N^P is the number of polarizations of a given transmitter. For each incident field $\bar{E}_{t,s}^{\text{inc}}$, scattered field $\bar{E}_{t,s}^{\text{sca}}$ is generated and can be computed using (8). Consequently, for multiple transmitters with multiple polarization configurations, the operator $f(\bar{t})$ should be constructed by cascading (8) for every possible pair of $\bar{E}_{t,s}^{\text{inc}}$ and $\bar{E}_{t,s}^{\text{sca}}$. Third, the inversion operation in (8) is carried out using a biconjugate gradient stabilized iterative solver [41]. The memory requirement and computational cost of executing the matrix–vector product required by this iterative solver both scale with $O(N^2)$. This cost is significantly reduced using the fact that blocks of the matrix $\bar{F}$, $\bar{G}_{uv}$, $\{u, v\} \in \{x, y, z\}$ are Toeplitz matrices. Consequently, the matrix $\bar{F}$ can be multiplied by a given vector in $O(N \log N)$ operations using fast Fourier transform as described in [42] and [43]. The memory requirement of this operation scales with $O(N)$. Finally, due to the smoothing effect of the integral operator in (1), finding the contrast $\bar{t}$ from using a finite set of noisy measurements at the receiver locations is an ill-posed problem [1]–[4]. Effects of ill-posedness is alleviated by casting the problem of finding $\bar{t}$ from measured scattered fields in the form of a sparse nonlinear optimization problem as described in the next section.

B. Sparse Nonlinear Optimization Problem

Solving the sparse nonlinear optimization problem

$$\bar{t} = \min_{\bar{t}} \frac{1}{2} \| f(\bar{t}) - \bar{E}^{\text{meas}} \|_2^2, \quad \|\bar{t}\|_0 \leq l_0 \qquad (9)$$

yields the samples of the contrast $\tau(\mathbf{r})$, which are stored in the vector $\bar{t}$, from noisy scattered fields measured at the receiver locations and stored in the vector $\bar{E}^{\text{meas}}$. In (9), the nonlinear least-square minimization the data misfit $\| f(\bar{t}) - \bar{E}^{\text{meas}} \|_2^2$ is constrained by the condition $\|\bar{t}\|_0 \leq l_0$, where l_0 is a positive integer. Since $\|\bar{t}\|_0 = \sum_{n=1}^{N} |\{\bar{t}\}_n|^0$, $\|\bar{t}\|_0$ counts the number of the nonzero entries in $\bar{t}$. This means that enforcing the constraint $\|\bar{t}\|_0 \leq l_0$ promotes the sparseness in the solution of (9) [27]. Consequently, l_0 should be estimated using the *a priori* knowledge of the sparseness level in the investigation domain, i.e., the number of zero-valued contrast samples to N. However, L_0-norm function is nonconvex. Therefore, solving the optimization problem (9) comes with a very high computational cost [44]. Alternatively, one can use the L_1-norm instead of the L_0-norm to constrain the minimization of the data misfit leading to the optimization problem

$$\bar{t} = \min_{\bar{t}} \frac{1}{2} \| f(\bar{t}) - \bar{E}^{\text{meas}} \|_2^2, \quad \|\bar{t}\|_1 \leq l_1 \qquad (10)$$

where l_1 is a positive real number. Since $\|\bar{t}\|_1 = \sum_{n=1}^{N} |\{\bar{t}\}_n|$, l_1 should be estimated based on *a priori* knowledge of the values of the contrast samples as well as the sparseness level in the investigation domain. Clearly, unlike the L_0-norm constraint, L_1-norm constraint also depends on the values of the contrast samples. Optimization problem (10) is the best convex norm approximation to the optimization problem (9) and replacing (9) with (10) is termed as convex relaxation [28]. The effectiveness of the L_1-norm constrained optimization in recovering sparse solutions (with smaller computational cost) has been well proven [27], [28].

Using the method of Lagrange multipliers [45], (9) and (10) are converted into

$$\bar{t} = \min_{\bar{t}} \frac{1}{2} \| f(\bar{t}) - \bar{E}^{\text{meas}} \|_2^2 + \lambda \|\bar{t}\|_q \qquad (11)$$

where $q \in \{0, 1\}$, and λ is a regularization parameter that specifies the tradeoff between minimizing the data misfit and the penalty term. It should be noted here that (11) is also known as the sparse nonlinear Tikhonov minimization problem [45]. The next section describes a projected accelerated steepest descent (PASD) algorithm to solve (11) [or equivalently (9) or (10)].

C. Projected Accelerated Steepest Descent Algorithm

Optimization problem (11) is iteratively solved using a PASD algorithm [38]. PASD enforces the sparsity constraint using a projection operator on the NLW iteration. The efficiency is achieved without sacrificing the convergence by using carefully selected thresholding level and step for the projection operator and iterations, respectively. A detailed mathematical study of the selection criteria and the convergence of PASD is carried out in [38]. In this paper, this algorithm is adopted to solve the sparse electromagnetic inverse scattering problem. The resulting PASD reads

Step 1: initialize $l_q, \alpha, \gamma, \bar{t}_{(0)}, \upsilon < 1$

Step 2: $r = \max\{2\alpha, 2\gamma \sqrt{\Gamma\{\bar{t}_{(0)}\}}\}$

Step 3.0: for $p = 0, 1, 2, \ldots$.

Step 3.1: $\beta_{(p)} = \max \left\{ \dfrac{\alpha}{\gamma \sqrt{\Gamma\{\bar{t}_{(p)}\}}}, \sqrt{\dfrac{\Gamma\{\bar{t}_{(0)}\}}{\Gamma\{\bar{t}_{(p)}\}}} \right\}$

Step 3.2: $\bar{t}_{(p+1)} = P_q \left(\bar{t}_{(p)} + \dfrac{\beta_{(p)}}{r} \partial_{\bar{t}_{(p)}} f^*(\bar{E}^{\text{meas}} - f(\bar{t}_{(p)})) \right)$

Step 3.3: does $\beta_{(p)}$ satisfy condition (13b) ?

　　　　yes, continue

　　　　no, $\beta_{(p)} = \upsilon \beta_{(p)}$ and back to Step 3.2

Step 3.4: end.

At Step 1, several parameters of the algorithm are initialized. Here, $l_q, q \in \{0, 1\}$ represents the estimated sparsity threshold [see (9), (10), or (11)], $\alpha \geq \sup_{\bar{t} \in B_q} \|\partial_{\bar{t}} f(\bar{t})\|_2$, $\gamma \geq 2 \sup_{\{\bar{t}, \bar{h}\} \in B_q} \|\partial_{\bar{t}}^2 f(\bar{t}, \bar{h})\|_2 / \|\bar{h}\|_2^2$, $B_q = \{\|\bar{t}\|_q \leq l_q\}$ is a ball of radius l_q, and $\partial_{\bar{t}} f(\bar{t})$ and $\partial_{\bar{t}}^2 f(\bar{t}, \bar{h})$ are the first- and second-order Frechet derivatives that are computed using

$$\partial_{\bar{t}} f(\bar{t}) = \bar{H} \bar{K} D\{\bar{T} \bar{E}^{\text{inc}}\} \qquad (12)$$
$$\partial_{\bar{t}}^2 f(\bar{t}, \bar{h}) = -\bar{H} \bar{K} (D\{\bar{h}\} \bar{F} \bar{K} D\{\bar{T} \bar{E}^{\text{inc}}\} + D\{\bar{T} \bar{F} D\{\bar{h}\} \bar{T} \bar{E}^{\text{inc}}\})$$

where $\bar{K} = (\bar{I}_{3N} - D\{\bar{P}\bar{t}\}\bar{F})^{-1}$ and $\bar{T} = (\bar{I}_{3N} - \bar{F}D\{\bar{P}\bar{t}\})^{-1}$. In addition, at Step 1, $\bar{t}_{(0)}$ is an initial guess of the solution and $\upsilon < 1$ is a shrinking factor.

At Step 2, r is specified using the estimated parameters α and γ and the data misfit of the initial guess given by $\Gamma(\bar{t}_0)$. Here, data misfit function is defined as $\Gamma(\bar{t}) = 0.5\| f(\bar{t}) - \bar{E}^{\text{meas}} \|_2^2$.

Step 3 and its substeps describe the iterations. At Step 3.1, $\beta_{(p)}$ is selected. Step 3.2 is the projected steepest descent iteration (also known as the projected NLW iteration). Here, $\beta_{(p)}/r$ is the iteration step, $P_q(\cdot)$ is the projection operator, and $\partial_{\bar{t}_{(p)}} f^*(\cdot)$ is the complex adjoint of $\partial_{\bar{t}_{(p)}} f(\cdot)$. Parameters $\beta_{(p)}$ and r should satisfy the conditions

$$\sup\{\beta_{(p)}, \forall p\} < \infty \text{ and } \inf\{\beta_{(p)}, \forall p\} \geq 1 \qquad (13a)$$

$$\beta_{(p)} \| f(\bar{t}_{(p+1)}) - f(\bar{t}_{(p)}) \|_2^2 \leq \frac{r}{2} \|\bar{t}_{(p+1)} - \bar{t}_{(p)}\|_2^2 \qquad (13b)$$

$$\beta_{(p)} \gamma \sqrt{\Gamma(\bar{t}_{(p)})} \leq \frac{r}{2} \qquad (13c)$$

to ensure the convergence of the algorithm [38]. Specific selections of r and $\beta_{(p)}$ at Steps 2 and 3.1 ensure that the conditions in (13a) and (13c) are satisfied at any given iteration [38]. However, r and $\beta_{(p)}$ should be explicitly checked against condition (13b). This is done at Step 3.3 of the algorithm. If the condition (13b) is satisfied, PASD proceeds to the next iteration. If not, $\beta_{(p)}$ (and consequently the iteration step) is reduced by multiplying it with υ, $\upsilon < 1$, and $\bar{t}_{(p+1)}$ is recomputed by repeating Step 3.2. This operation is iterated until $\beta_{(p)}$ satisfies condition (13b).

At Step 3.2, the operator $P_q(\cdot)$ projects the result of each steepest descent iteration into the ball B_q to enforce the sparsity constraint. Different projection operators are used for

$q = 0$ and $q = 1$ as described next. For $q = 0$, $P_0(\cdot)$ is computed using [38], [39]

Step 1) $\bar{x} = \text{sort}(|\bar{z}|)$
Step 2) $\mu = \{\bar{x}\}_{(l_0+1)}$
Step 3) $P_0(\bar{z}) = \text{Thr}_0^\mu(\bar{z})$.

At Step 1, vector $\bar{x}$ stores the absolute values of entries of the input vector $\bar{z}$, which are sorted from the largest to the smallest. At Step 2, threshold level μ is determined. At Step 3, the complex hard-thresholding function with threshold level μ [32]–[34]

$$\left\{\text{Thr}_0^\mu(\bar{z})\right\}_n = \begin{cases} 0 & |\{\bar{z}\}_n| < \mu \\ \{\bar{z}\}_n & \text{else} \end{cases}$$

is used as $P_0(\bar{z})$. This projection operation sets the entries of $\bar{z}$, which satisfy $|\{\bar{z}\}_k| < |\{\bar{z}\}|_{(l_0+1)}$, to zero. If it is known *a priori* that l_0 number of entries in $\bar{t}$ (among a total of N number of entries) are zero, then there are $N!/(l_0!(N - l_0)!)$ possible sparse solutions [27]. The projection operator $P_0(\cdot)$ described above selects l_0 entries of the vector $\bar{z}$, starting with the one with largest absolute value. The chances of these entries to be in the actual solution are higher [27]. However, to achieve the best sparse solution, one still needs to compute the data misfit [first term in (11)] for all possible sparse solutions. This makes the L_0-norm sparse optimization problem computationally expensive and inefficient [27]. It should also be noted here, even though $P_0(\cdot)$ selects a set of entries, which is more likely to be in the actual solution, it fails to provide efficient sparse reconstruction for investigation domains with high levels of permittivity and/or involving geometrically complicated scatterers [32]–[34]. For such investigation domains, some entries of the solution cannot "grow" enough to overcome the threshold level of $P_0(\cdot)$, especially during the first few iterations. Consequently, these entries are eventually eliminated from the solution. One way of tackling this problem is to use a large value for l_0, say $l_0^{\max}$, during the first few iterations of PASD. As the iterations evolve, the value for l_0 is reduced until it reaches a desired level of sparseness, say $l_0^{\min}$. This approach ensures that some entries of the solution are not prematurely eliminated at the earlier iterations since the threshold level is gradually increased. Indeed, as shown by the numerical results in Section III, it leads to a more efficient sparse recovery by allowing the entries of the solution to grow enough to surpass the projection operation at the later iterations.

For $q = 1$, $P_1(\cdot)$ is computed using [38], [39]

Step 1) $\bar{x} = \text{sort}(|\bar{z}|)$
Step 2) Find k such that
$$\left\|\text{Thr}_1^{\{\bar{x}\}(k)}(\bar{x})\right\|_1 \leq l_1 \leq \left\|\text{Thr}_1^{\{\bar{x}\}(k+1)}(\bar{x})\right\|_1$$
Step 3) $\mu = \{\bar{x}\}_{(k)} - \left(l_1 - \left\|\text{Thr}_1^{\{\bar{x}\}(k)}(\bar{x})\right\|_1\right)/k$
Step 4) $P_1(\bar{z}) = \text{Thr}_1^\mu(\bar{z})$.

Step 1 in the computation of $P_1(\cdot)$ is the same as that in the computation of $P_0(\cdot)$. At Step 2, a search is preformed to find index k such that $\|\text{Thr}_1^{\{\bar{x}\}(k)}(\bar{x})\|_1 \leq l_1 \leq \|\text{Thr}_1^{\{\bar{x}\}(k+1)}(\bar{x})\|_1$ is

satisfied. Step 3 computes the threshold μ for the projection operation. Finally, at Step 4, $P_1(\bar{z})$ is computed using $\text{Thr}_1^\mu(\bar{z})$. Here, $\text{Thr}_1^\mu(\bar{z})$ is the complex soft-thresholding function given by [32]–[34]

$$\left\{\text{Thr}_1^\mu(\bar{z})\right\}_n = \{\bar{z}\}_n \frac{\max\{|\{\bar{z}\}_n - \mu|, 0\}}{\max\{|\{\bar{z}\}_n - \mu|, 0\} + \mu}.$$

In summary, the projection operation and the conditions in (13) ensure that PASD confines the iterative search of the solution inside the ball B_q while using the largest possible iteration step that does not sacrifice the convergence of the algorithm. One can think of PASD as a different version of the nonlinear iterative shrinkage thresholding (NIST) algorithm (i.e., thresholded NLW), where each iteration is in the form

$$\bar{t}_{(p+1)} = \text{Thr}_q^\lambda\left(\bar{t}_{(p)} + \left(1/\sigma_{\max(p)}^2\right)\partial_{\bar{t}_{(p)}} f^*(\bar{E}^{\text{meas}} - f(\bar{t}_{(p)}))\right). \tag{14}$$

Here, $\sigma_{\max(p)}$ is the maximum singular value of the operator $\partial_{\bar{t}_{(p)}} f(\cdot)$. It should be noted here that $\sigma_{\max(p)}^2$ normalizes the adjoint operator $\partial_{\bar{t}_{(p)}} f^*(\cdot)$ at each iteration to ensure convergence. However, unlike PASD, this normalization operation may not produce the largest step that guarantees the convergence of NIST or NLW [38]. Consequently, PASD is more efficient than NIST because: 1) the normalization process of NIST requires the estimation/computation of $\sigma_{\max(p)}$ and 2) NIST converges slowly due to the suboptimal selection of the iteration step.

It should also be noted here that one can use $q = 2$ in (11). The resulting optimization problem can be solved using PASD after the projection operator is dropped, i.e., the projection operator is replaced by an identity operator. This turns PASD into truncated NLW. It is well known that NLW starts by recovering the slowly varying components of the solution (i.e., low spatial frequency components) and as the iterations continue, components with faster variations (i.e., high spatial frequency components) are recovered [45]. But these higher frequency components are more susceptible to noise in the measured data. Therefore, the iterations are truncated before the noise starts corrupting the solution. In other words, truncation acts like a regularization scheme [1]–[4]. Consequently, higher frequency components of the solution are not recovered. This means that solving the optimization problem (11) with $q = 2$ using truncated NLW promotes the smoothness in the solution.

III. NUMERICAL RESULTS

In this section, several numerical experiments, which demonstrate the efficiency and applicability of PASD, are described. Its performance and accuracy are compared with those of NIST and truncated NLW. In all examples considered here, it is assumed that the investigation domain resides in free space. The incident field generated by each transmitter t, $t = 1, \ldots, N^T$, is approximated by a plane wave polarized in $\hat{\mathbf{a}}_s$ direction, $s = 1, \ldots, N^P$

$$E_{t,s}^{\text{inc}}(\mathbf{r}) = E_0 e^{-jk_0 \hat{\mathbf{k}}_t \cdot \mathbf{r}} \hat{\mathbf{a}}_s.$$

Here, E_0 is the amplitude of the plane wave's electric field and $\hat{\mathbf{k}}_t$ is the unit vector in the direction of plane wave propagation. For every transmitter t, $N^P = 2$, $\hat{\mathbf{a}}_1 = \hat{\mathbf{a}}_\phi$ ($s = 1$), and $\hat{\mathbf{a}}_2 = \hat{\mathbf{a}}_\theta$ ($s = 2$), where $\hat{\mathbf{a}}_\phi$ and $\hat{\mathbf{a}}_\theta$ are the unit vectors along ϕ and θ directions in the Spherical coordinate system. The transmitters are located at points $(R^T \hat{\mathbf{a}}_r, [2m\pi/N_\phi^T]\hat{\mathbf{a}}_\phi, [n\pi/N_\theta^T]\hat{\mathbf{a}}_\theta)$, $m = 1, \ldots, N_\phi^T$ and $n = 1, \ldots, N_\theta^T$. The total number of transmitters $N^T = N_\theta^T N_\phi^T$. For a given transmitter location, the propagation unit vector $\hat{\mathbf{k}}_t$ is directed from the transmitter location toward the origin. The receivers are located at points $(R^R \hat{\mathbf{a}}_r, [2m\pi/N_\phi^R]\hat{\mathbf{a}}_\phi, [n\pi/N_\theta^R]\hat{\mathbf{a}}_\theta)$, $m = 1, \ldots, N_\phi^R$ and $n = 1, \ldots, N_\theta^R$. The total number of receivers $N^R = N_\theta^R N_\phi^R$. Let $\tau^{\text{ref}}(\mathbf{r})$ represent the actual contrast of the investigation domain and $\{\bar{t}^{\text{ref}}\}_k = \tau^{\text{ref}}(\mathbf{r}_k)$, $k = 1, \ldots N$. Measured fields stored in $\bar{E}_{t,s}^{\text{meas}}$ are synthetically generated by adding white Gaussian noise to the scattered fields, which are computed using $\bar{t}^{\text{ref}}$ and $\bar{E}_{t,s}^{\text{inc}}$ in (8). The accuracy of reconstruction is quantified using one of the following error criteria:

$$diff^{(p)} = \|\bar{t}_{(p)} - \bar{t}^{\text{ref}}\|_2$$

$$\text{err}^{(p)} = \frac{\|\bar{t}_{(p)} - \bar{t}^{\text{ref}}\|_2}{\|\bar{t}^{\text{ref}}\|_2}$$

where p represents the iteration number of the imaging algorithm used. Using error function $diff^{(p)}$ is more meaningful when there is a comparison between reconstructions of different $\bar{t}^{\text{ref}}$.

A. Parameter Selection

This section explains how the parameters of PASD are selected for its accurate and efficient execution. In all examples considered, PASD starts with a zero initial guess, i.e., $\bar{t}_{(0)} = 0$. The general rule of thumb in selecting α, γ, and υ is that their values should reduce the number of possible checks/failures of the conditions (13) while maintaining the efficiency of the algorithm. The shrinking factor υ reduces the iteration step (see Step 3.3) if the conditions in (13) are not satisfied. The value of υ should not be very close to unity since this increases the chances of another condition failure leading to an increase in the number of repetitions of the same iteration. It should not also be too small since this might reduce the iteration step leading to an unnecessary increase in execution time. Therefore, υ is selected to be 0.9. Parameters α and γ are supposed to satisfy the conditions $\alpha \geq \sup_{\bar{t} \in B_q} \|\partial_{\bar{t}} f(\bar{t})\|_2$ and $\gamma \geq 2\sup_{\{\bar{t}, \bar{h}\} \in B_q} \|\partial_{\bar{t}}^2 f(\bar{t}, \bar{h})\|_2 / \|\bar{h}\|_2^2$, respectively, as described in [38]. It is clear from these conditions that the selection of α and γ depends on the supremum of the first and the second order Frechet derivative operator computed within B_q. From (12), one can see that Frechet derivative operators are functions of the transmitter–receiver configuration (including frequency), investigation domain geometry, and discretization. Their supremum within B_q is predicted by running numerical tests using the same setup as the inverse problem before PASD is executed.

Maximum singular value $\sigma_{\max(p)}$ does not have to be computed at every iteration since $1/\sigma_{\max(p)}^2$ is used as only a normalization factor to ensure convergence. Value of $\sigma_{\max(p)}$

TABLE I
PARAMETERS OF PASD AND NIST FOR THE INVESTIGATION DOMAIN WITH TWO SPHERES

	PASD		NIST	
	$q = 0$	$q = 1$		$q = 0$	$q = 1$
l_q	225	234	λ	0.07	0.01
α	0.3	0.243			
γ	0.02	0.03			

computed at one iteration can be used at the next few iterations without sacrificing from convergence depending on how much the Frechet derivative changes from that iteration to the next few iterations. Typically, at the first few iterations of the algorithm, the value of the Frechet derivative changes significantly between iterations. However, as iterations evolve and the solution starts to converge, the variation in the Frechet derivative values is less significant. Consequently, in all numerical examples considered here, $\sigma_{\max(p)}$ is computed at every iteration for the first five iterations and every ten iterations after that.

B. Two Spheres

The investigation domain includes two dielectric spheres with radius 0.39 m, which are centered at $(0.45, 0.45, 0.45)$ and $(-0.45, -0.45, -0.45)$ m, respectively. Their relative permittivity is 2.5. The investigation domain is discretized using $N = 8000$ cubic cells with dimension $\Delta d = 0.15$ m. The sparseness level, i.e., the ratio of number of nonzero entries in $\bar{t}^{\text{ref}}$ to N, is 2.2%. The parameters of the transmitter-receiver configuration are $f = 126$ MHz, $N_\phi^T = 5$, $N_\theta^T = 4$, $N_\phi^R = 6$, and $N_\theta^R = 5$. The level of noise in $\bar{E}^{\text{meas}}$ is 25 dB. The parameters of PASD and NIST are given in Table I. Note that the sparsity relaxation explained in Section II-C is not used for this example.

Fig. 2(a) plots $\text{err}^{(p)}$, which is computed by truncated NLW, NIST, and PASD, versus p. In addition, Fig. 2(b) plots $\text{err}^{(p)}$ versus execution time of the three methods. It is clearly shown that PASD achieves significantly better accuracy within shorter time. For $q = 0$, NIST achieves $\text{err}^{(200)} = 48.51\%$ in 131.7 min, while PASD reaches the same error level in 1.8 min. For $q = 1$, NIST achieves $\text{err}^{(200)} = 47.92\%$ in 130.9 min, while PASD reaches the same error level in 15.6 min.

Fig. 2(c) shows the slices of the actual relative permittivity profile at $x = 0.375$ m and $y = -0.375$ m planes. Fig. 2(d)–(h) shows the relative permittivity profiles reconstructed by truncated NLW at $p = 200$, NIST with $q = 0$ at $p = 200$, NIST with $q = 1$ at $p = 200$, PASD with $q = 0$ at $p = 60$, and PASD with $q = 1$ at $p = 60$, respectively. As expected, reconstructions generated by PASD and NIST are sharper than that obtained by truncated NLW [for example, compare Fig. 2(d) with (g) and (h)]. In addition, the local variations in contrast levels are captured more accurately by PASD and NIST.
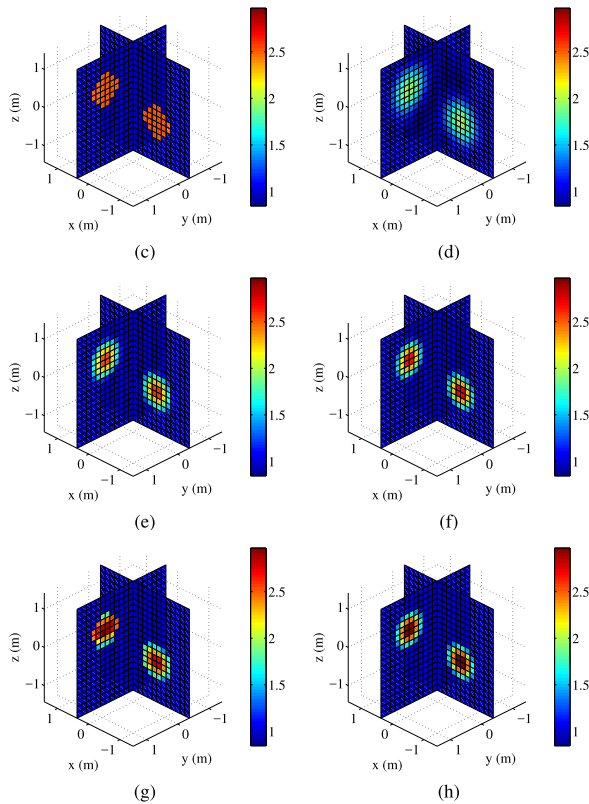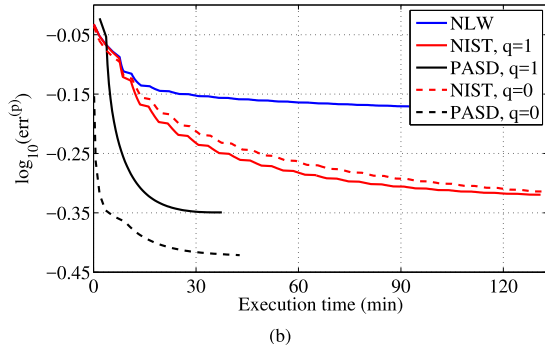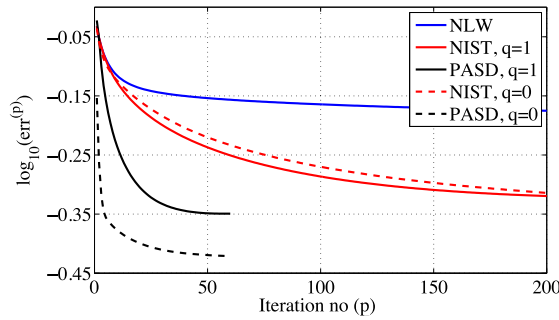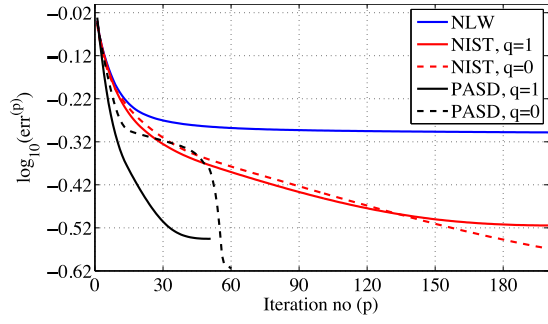
	PASD			NIST	
	$q = 0$	$q = 1$		$q = 0$	$q = 1$
l_q	-	720	λ	0.15	0.01
$l_q^{\max}$	2400	-			
$l_q^{\min}$	500	-			
α	0.17	0.16			
γ	0.03	0.02			

permittivities of the cube and layer are 1.5 and 2.5, respectively. The investigation domain is discretized using $N = 8000$ cells with dimension $\Delta d = 0.15$ m. The sparseness level is 6.4%. The parameters of the transmitter–receiver configuration are $f = 126$ MHz, $N_\phi^T = 4$, $N_\theta^T = 4$, $N_\phi^R = 4$, and $N_\theta^R = 4$. The level of noise in $\bar{E}^{\text{meas}}$ is 25 dB. The parameters of PASD and NIST are given in Table II.

Fig. 3(a) plots $\text{err}^{(p)}$, which is computed by truncated NLW, NIST, and PASD, versus p. In addition, Fig. 3(b) plots $\text{err}^{(p)}$ versus execution time of the three methods. The results show that PASD is more efficient and accurate than the other two. For $q = 0$, NIST achieves $\text{err}^{(200)} = 27.02\%$ in 358.3 min, while PASD reaches the same error level in 82.6 min. For $q = 1$, NIST achieves $\text{err}^{(200)} = 30.59\%$ in 359.38 min, while PASD reaches the same error level in 44.98 min.

Fig. 3(c) shows the slices of the actual relative permittivity profile at $x = 0$ m, $y = 0$ m, and $z = 0$ m planes. Fig. 3(d)–(h) shows the relative permittivity profiles reconstructed by truncated NLW at $p = 200$, NIST with $q = 0$ at $p = 200$, NIST with $q = 1$ at $p = 200$, PASD with $q = 0$ at $p = 60$, and PASD with $q = 1$ at $p = 51$, respectively. Comparing Fig. 3(d) with (g) and (h), one can note that in the solution reconstructed by truncated NLW (where the smoothness is promoted), the core of cube under the coating cannot be detected; both the core and coating appear as if they have the same permittivity. On the other hand, in the solutions reconstructed by NIST and PASD, the core and coating can be clearly identified.
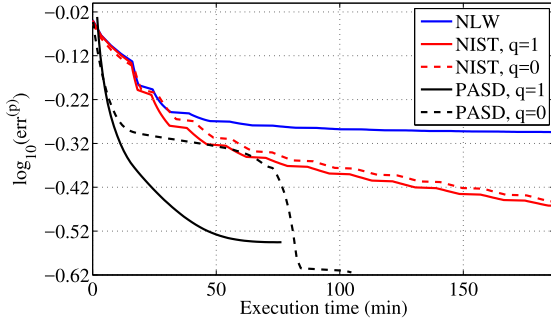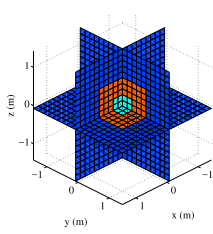
D. Austria Scatterer

The scatterer in this example is the 3-D version of the well-known 2-D Austria dielectric permittivity profile [1]. The scatterer consists of a spherical shell with inner radius 0.6 m and outer radius 1.1 m and two spheres with radius 0.35 m. The shell and spheres are centered at $(0, 0, 0)$ and $(0, 1.5, 0.95)$ and $(0, -1.5, 0.95)$ m, respectively. The relative permittivity of all structures is 2.5. The investigation domain is discretized using $N = 27\,000$ cells with dimension $\Delta d = 0.15$ m. The sparseness level is 5.4%. The parameters of the transmitter–receiver configuration are $f = 100$ MHz, $N_\phi^T = 3$, $N_\theta^T = 3$, $N_\phi^R = 5$, and $N_\theta^R = 4$. The level of noise in $\bar{E}^{\text{meas}}$ is 25 dB. The parameters of PASD and NIST are given in Table III.

Fig. 4(a) plots $\text{err}^{(p)}$, which is computed by truncated NLW, NIST, and PASD, versus p. In addition, Fig. 4(b) plots $\text{err}^{(p)}$ versus execution time of the three methods. For $q = 0$, NIST
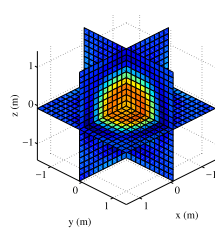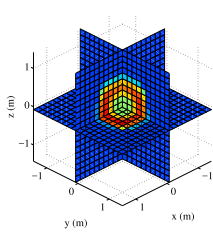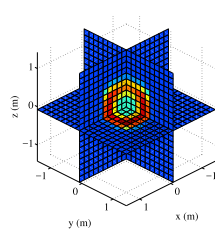
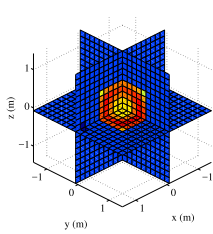Fig. 2. Investigation domain with two spheres. (a) $\text{err}^{(p)}$ versus p and (b) $\text{err}^{(p)}$ versus execution time for truncated NLW, NIST, and PASD. (c) Actual relative permittivity profile. Relative permittivity profiles reconstructed by (d) truncated NLW at $p = 200$, (e) NIST with $q = 0$ at $p = 200$, (f) NIST with $q = 1$ at $p = 200$, (g) PASD with $q = 0$ at $p = 60$, and (h) PASD with $q = 1$ at $p = 60$.

C. Coated Cube

In this example, the investigation domain includes a coated cube centered at the origin. The dimension of the cube and the thickness of the coating layer is 0.3 m. The relative
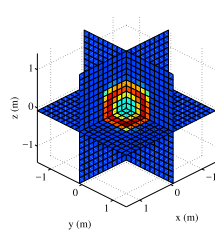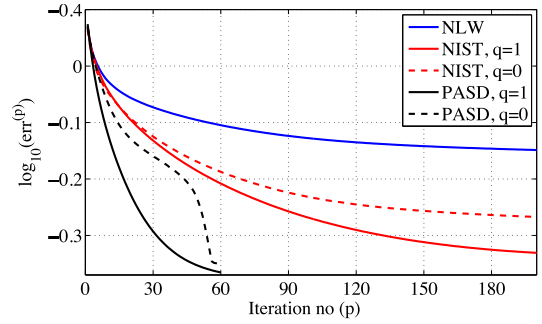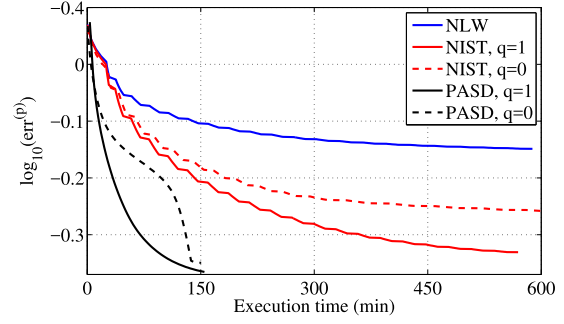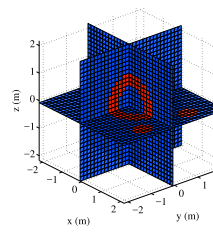
Fig. 3. Investigation domain with a coated cube. (a) err$^{(p)}$ versus p and (b) err$^{(p)}$ versus execution time for truncated NLW, NIST, and PASD. (c) Actual relative permittivity profile. Relative permittivity profiles reconstructed by (d) truncated NLW at $p = 200$, (e) NIST with $q = 0$ at $p = 200$, (f) NIST with $q = 1$ at $p = 200$, (g) PASD with $q = 0$ at $p = 60$, and (h) PASD with $q = 1$ at $p = 51$.

Fig. 4. Investigation domain with the Austria scatterer. (a) err$^{(p)}$ versus p and (b) err$^{(p)}$ versus execution time for truncated NLW, NIST, and PASD. (c) Actual relative permittivity profile. Relative permittivity profiles reconstructed by (d) truncated NLW at $p = 200$, (e) NIST with $q = 0$ at $p = 200$, (f) NIST with $q = 1$ at $p = 200$, (g) PASD with $q = 0$ at $p = 60$, and (h) PASD with $q = 1$ at $p = 60$.

achieves err$^{(200)}$ = 42.93% at 892.4 min, while PASD reaches the same error level in 107.7 min. For $q = 1$, NIST achieves err$^{(200)}$ = 37.09% in 569.5 min, while PASD reaches the same error level in 94.59 min.
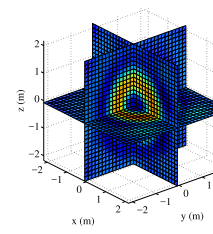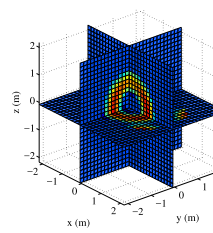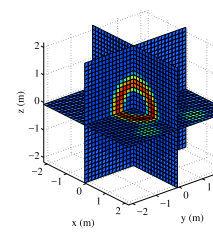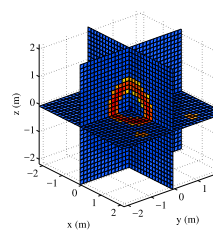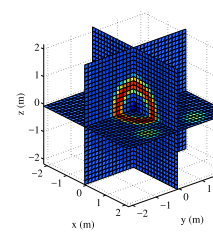
Fig. 4(c) shows the slices of the actual relative permittivity profile at $x = 0$ m, $y = 0$ m, and $z = 0$ m. Fig. 4(d)–(h) shows the relative permittivity profiles reconstructed by truncated NLW at $p = 200$, NIST with $q = 0$

TABLE III
PARAMETERS OF PASD AND NIST FOR THE INVESTIGATION
DOMAIN WITH AUSTRIA SCATTERER

	PASD			NIST	
	$q = 0$	$q = 1$		$q = 0$	$q = 1$
l_q	-	2100	λ	0.15	0.01
$l_q^{\max}$	8100	-			
$l_q^{\min}$	1600	-			
α	0.4	0.04			
γ	0.09	0.005			

TABLE IV
PARAMETERS OF PASD AND NIST FOR THE INVESTIGATION
DOMAIN WITH A HIDDEN TARGET

	PASD			NIST	
	$q = 0$	$q = 1$		$q = 0$	$q = 1$
l_q	-	2800	λ	0.1	0.01
$l_q^{\max}$	16200	-			
$l_q^{\min}$	3500	-			
α	0.06	0.06			
γ	0.01	0.09			

at $p = 200$, NIST with $q = 1$ at $p = 200$, PASD with $q = 0$ at $p = 60$, and PASD with $q = 1$ at $p = 60$, respectively. In the solution reconstructed by truncated NLW, wavelike ripples, which make the two spheres undetectable, are observed. However, these ripples are eliminated in the solutions reconstructed by NIST and PASD, allowing the detection of the spheres as two separate objects.

E. Hidden Target

The investigation domain includes a spherical shell with inner radius 1.6 m and outer radius 1.86 m and a cubic target of dimension 0.75 m. The shell and the cube are centered at $(0, 0, 0)$ and $(0.525, 0, 0.525)$ m, respectively. Note that the cube is located inside the shell. The relative permittivities of the shell and cube are 2.0 and 2.5, respectively. The investigation domain is discretized using $N = 27\,000$ cells with dimension $\Delta d = 0.15$ m. The sparseness level is 5.4%. The parameters of the transmitter–receiver configuration are $f = 100$ MHz, $N_\phi^T = 3$, $N_\theta^T = 3$, $N_\phi^R = 5$, and $N_\theta^R = 4$. The level of noise in $\tilde{E}^{\mathrm{meas}}$ is 25 dB. The parameters of PASD and NIST are given in Table IV.

Fig. 5(a) plots $\mathrm{err}^{(p)}$, which is computed by truncated NLW, NIST, and PASD, versus p. In addition, Fig. 5(b) plots $\mathrm{err}^{(p)}$ versus execution time of the three methods. For $q = 0$, NIST achieves $\mathrm{err}^{(400)} = 55.99\%$ in 1635 min, while PASD reaches the same error level in 141.5 min. For $q = 1$, NIST achieves $\mathrm{err}^{(400)} = 49.7\%$ in 1103 min, while PASD reaches the same error level in 181.7 min.

Fig. 5(c) shows the slices of the actual relative permittivity profile at $x = 0$ m, $y = 0$ m, and $z = 0$ m planes. Fig. 5(d)–(h) shows the relative permittivity profiles reconstructed by truncated NLW at $p = 400$, NIST with $q = 0$ at $p = 400$, NIST with $q = 1$ at $p = 400$, PASD with $q = 0$ at $p = 60$, and PASD with $q = 1$ at $p = 100$, respectively. It is clearly seen that the solutions reconstructed by NIST and PASD are sharper than that reconstructed by truncated NLW. In the images obtained by NIST and PASD, the target located inside the shell can clearly be identified.
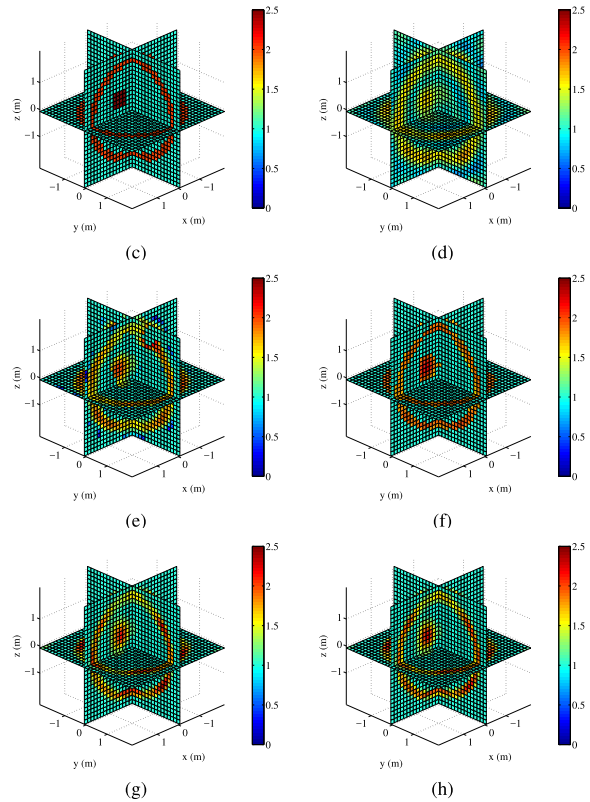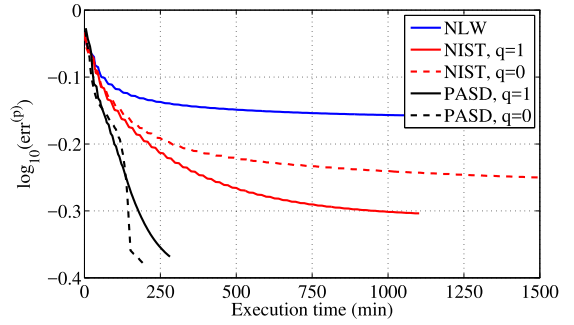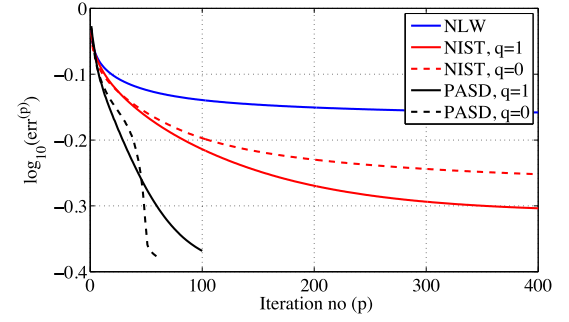

Fig. 5. Investigation domain with a hidden target. (a) $\mathrm{err}^{(p)}$ versus p and (b) $\mathrm{err}^{(p)}$ versus execution time for truncated NLW, NIST, and PASD. (c) Actual relative permittivity profile. Relative permittivity profiles reconstructed by (d) truncated NLW at $p = 400$, (e) NIST with $q = 0$ at $p = 400$, (f) NIST with $q = 1$ at $p = 400$, (g) PASD with $q = 0$ at $p = 60$, and (h) PASD with $q = 1$ at $p = 100$.

F. Range of Validity

In this section, several numerical results are presented to demonstrate how changes in receiver–transmitter configuration, contrast and sparseness levels, level of noise in
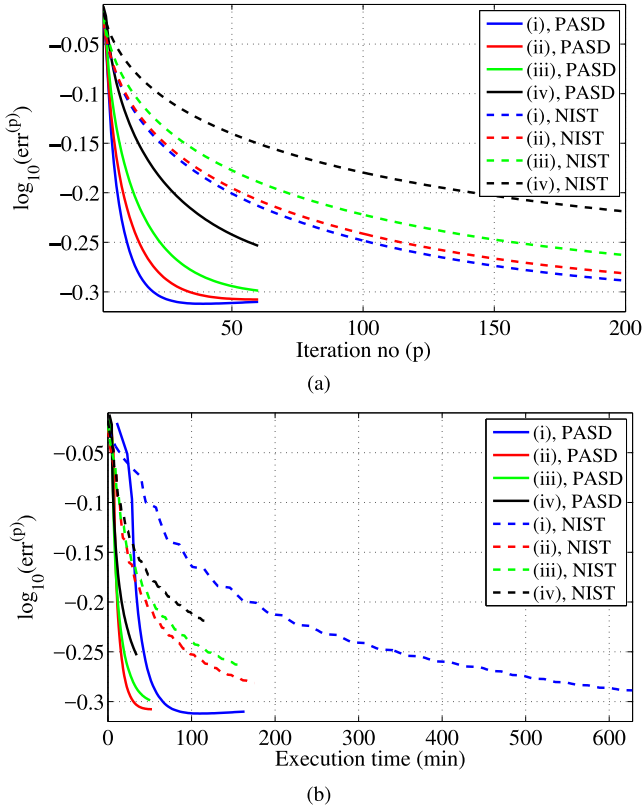
(a)

(b)

Fig. 6. Investigation domain with two spheres. (a) $err^{(p)}$ versus p and (b) $err^{(p)}$ versus execution time for PASD and NIST using four different transmitter–receiver configurations: 1) $N_\phi^T = 5$, $N_\theta^T = 4$, $N_\phi^R = 5$, and $N_\theta^R = 4$; 2) $N_\phi^T = 4$, $N_\theta^T = 3$, $N_\phi^R = 5$, and $N_\theta^R = 4$; 3) $N_\phi^T = 4$, $N_\theta^T = 3$, $N_\phi^R = 4$, and $N_\theta^R = 3$; and 4) $N_\phi^T = 3$, $N_\theta^T = 3$, $N_\phi^R = 3$, and $N_\theta^R = 3$.

Fig. 7. Investigation domain with two spheres. $diff^{(p)}$ versus execution time for PASD and NIST for five different values of permittivity: 1) 1.5; 2) 2.5; 3) 3.5; 4) 4.5; and 5) 5.5.

Fig. 8. Investigation domain with a single sphere. $diff^{(p)}$ versus execution time for PASD and NIST for four different levels of sparseness: 1) 1.1%; 2) 2%; 3) 5.4%; and 4) 11.4%.

measurements, and frequency affect the convergence of PASD.

1) Transmitter–Receiver Configuration: All parameters used in this example are the same as those in Section III-B except the transmitter–receiver configuration. In addition, the conductivity of the spheres is now 5 mS/m. Reconstruction is carried out using PASD with $q = 1$ and NIST with $q = 1$ for four different transmitter–receiver configurations: 1) $N_\phi^T = 5$, $N_\theta^T = 4$, $N_\phi^R = 5$, and $N_\theta^R = 4$; 2) $N_\phi^T = 4$, $N_\theta^T = 3$, $N_\phi^R = 5$, and $N_\theta^R = 4$; 3) $N_\phi^T = 4$, $N_\theta^T = 3$, $N_\phi^R = 4$, and $N_\theta^R = 3$; and 4) $N_\phi^T = 3$, $N_\theta^T = 3$, $N_\phi^R = 3$, and $N_\theta^R = 3$. The parameters of PASD and NIST are the same as those given in Table I.

Fig. 6(a) plots $err^{(p)}$, which is computed by PASD and NIST, versus p for all four simulations. In addition, Fig. 6(b) plots $err^{(p)}$ versus execution time of the two methods. Fig. 6(a) and (b) shows that PASD converges faster than NIST for every transmitter–receiver configuration. In addition, it is clear that decreasing the numbers of the transmitters and the receivers degrades the accuracy. But as expected this also comes with a decrease in the computation time per iteration.

2) Contrast Level: All parameters used in this example are the same as those in Section III-B except the permittivity of the spheres. In addition, the conductivity of the spheres is now 5 mS/m. Reconstruction is carried out using PASD with $q = 1$
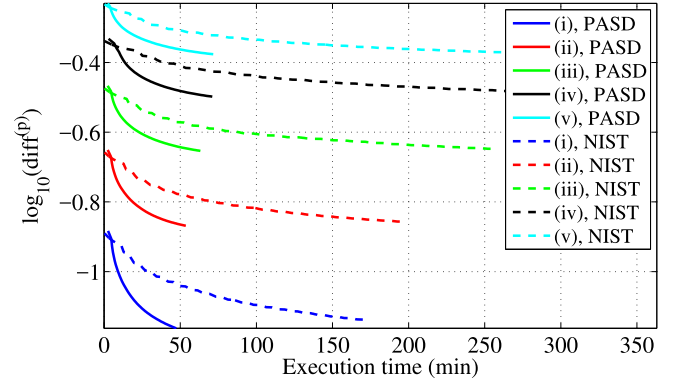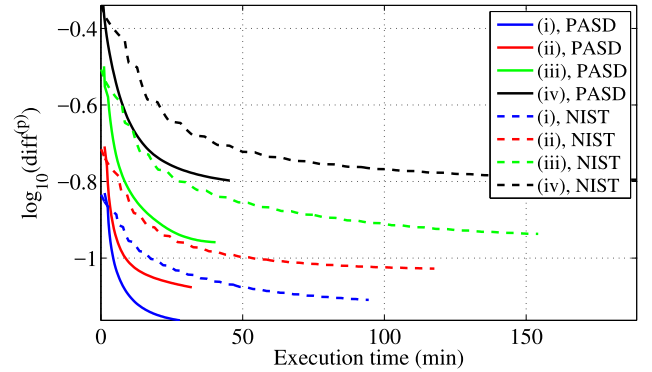
and NIST with $q = 1$ for five different values of permittivity: 1) 1.5; 2) 2.5; 3) 3.5; 4) 4.5; and 5) 5.5. The parameters of PASD and NIST are the same as those given in Table I except l_1, which is selected as 145, 235, 335, 430, and 512 for simulations 1)–5), respectively.

Fig. 7 plots $diff^{(p)}$, which is computed by PASD and NIST, versus execution time for all five simulations. Fig. 7 clearly shows that PASD does not lose its efficiency as the contrast is increased (i.e., level of nonlinearity increases).

3) Sparseness Level: In this example, the investigation domain includes a single sphere centered at the origin. Its relative permittivity and conductivity are 2.5 and 5 mS/m, respectively. The investigation domain is discretized using $N = 8000$ cubic cells with dimension $\Delta d = 0.15$ m. The parameters of the transmitter–receiver configuration are $f = 126$ MHz, $N_\phi^T = 5$, $N_\theta^T = 4$, $N_\phi^R = 6$, and $N_\theta^R = 5$. The level of noise in $\bar{E}^{\text{meas}}$ is 25 dB. Reconstruction is carried out using PASD with $q = 1$ and NIST with $q = 1$ for four different levels of sparseness: 1) 1.1%; 2) 2%; 3) 5.4%; and 4) 11.4%. It should be noted here that the varying levels of sparseness are obtained by changing the radius of the sphere. The parameters of PASD and NIST are the same as those given in Table I except l_1, which is selected as 132, 240, 648, and 1368 for simulations 1)–4), respectively.

Fig. 8 plots $diff^{(p)}$, which is computed by PASD and NIST, versus execution time for all four simulations. Fig. 8 shows
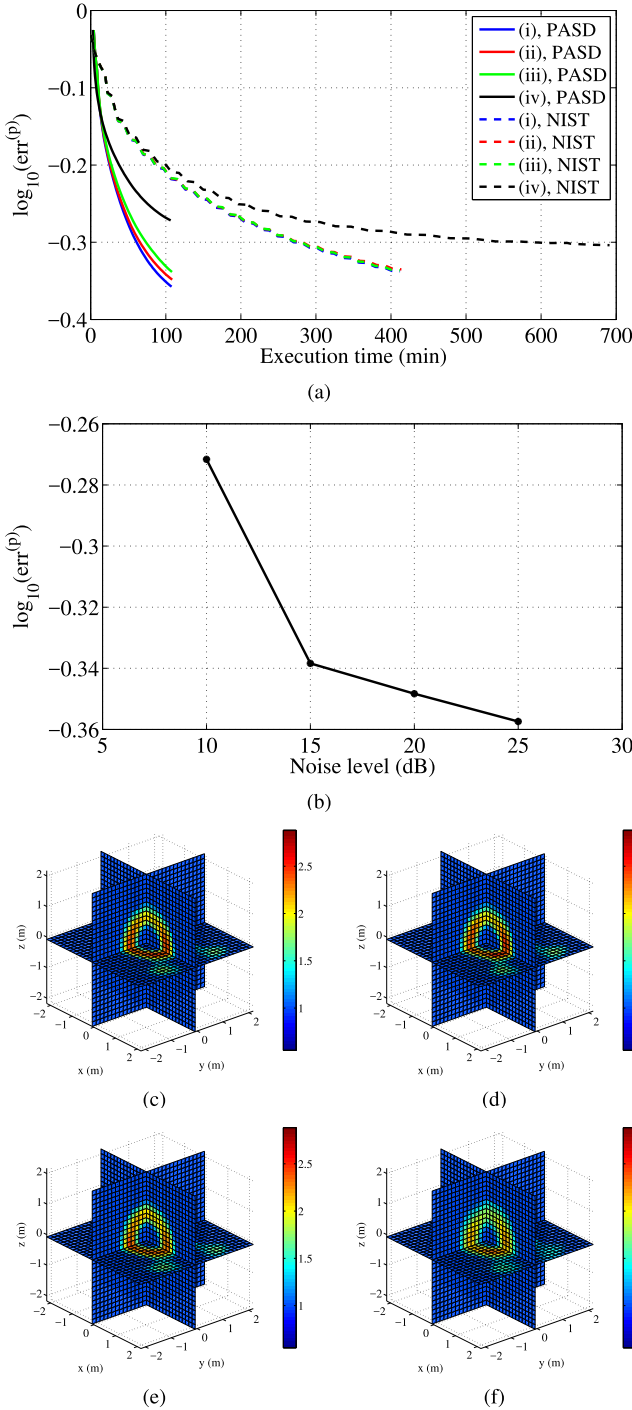
(a)

(b)

Fig. 9. Investigation domain with the Austria scatterer. (a) err$^{(p)}$ versus execution time for PASD and NIST under four different levels of noise: 1) 25; 2) 20; 3) 15; and 4) 10 dB. (b) err$^{(p)}$ at $p = 60$ versus noise level for PASD. Relative permittivity profiles reconstructed by PASD at $p = 60$ for noise level (c) 25, (d) 20, (e) 15, and (f) 10 dB.

that PASD converges much faster than NIST for all levels of sparseness considered in this example.

4) Noise Level: All parameters used in this example are the same as those in Section III-D except the level of noise in $\bar{E}^{\text{meas}}$. In addition, the conductivity of the scatterer is now 5 mS/m. Reconstruction is carried out using PASD with $q = 1$ and NIST with $q = 1$ for four different values of noise
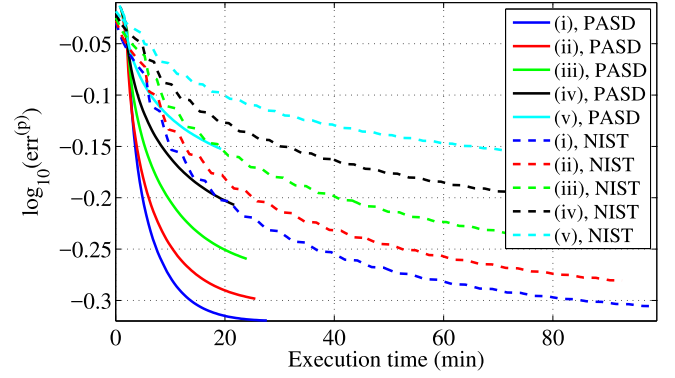

Fig. 10. Investigation domain with two spheres. err$^{(p)}$ versus execution time for PASD and NIST for five different values of frequency: 1) 125; 2) 115; 3) 105; 4) 95; and 5) 85 MHz.

level: 1) 25; 2) 20; 3) 15; and 4) 10 dB. The parameters of PASD and NIST are the same as those given in Table III.

Fig. 9(a) plots err$^{(p)}$, which is computed by PASD and NIST, versus execution time for all four simulations. Fig. 9(a) shows that PASD maintains a better convergence rate than NIST for all levels of noise considered in this example. Fig. 9(b) plots err$^{(p)}$, which is computed by PASD at $p = 60$, versus the noise level. Fig. 9(a) and (b) confirms the robustness of PASD against high levels of noise in the measurements. To demonstrate this further, Fig. 9(c)–(f) shows the slices of the relative permittivity profiles reconstructed by PASD at $p = 60$ for noise levels of 25, 20, 15, and 10 dB, respectively.

5) Frequency: All parameters used in this example are the same as those in Section III-B except the permittivity of the spheres. In addition, the conductivity of the spheres is now 5 mS/m. Reconstruction is carried out using PASD with $q = 1$ and NIST with $q = 1$ for five different values of frequency: 1) 125; 2) 115; 3) 105; 4) 95; and 5) 85 MHz. The parameters of PASD and NIST are the same as those given in Table I.

Fig. 10 plots err$^{(p)}$, which is computed by PASD and NIST, versus execution time for all five simulations. Fig. 10 shows that PASD maintains its faster convergence for different values of frequency.

G. Two Spheres: Comparison With Multiscale NIST

In this section, performance and accuracy of PASD are compared with those of multiscale NIST. Multiscaling is implemented using 3-D Gaussian blobs with changing variances [46]–[48]. To reduce the computational cost, multiscaling is only applied at NIST iterations where the maximum singular value $\sigma_{\max(p)}$ is computed, i.e., at every iteration for the first five iterations and every ten iterations after that as explained in Section III-A. The investigation domain and the parameters used in this example are the same as those in Section III-B except the numbers of transmitters and receivers. In this example, $N_\phi^T = 5$, $N_\theta^T = 4$, $N_\phi^R = 5$, and $N_\theta^R = 4$.

Fig. 11(a) plots err$^{(p)}$, which is computed by multiscale NIST with $q = 1$ and PASD with $q = 1$, versus execution time. Fig. 11(a) clearly shows that PASD has a faster convergence rate: multiscale NIST reaches err$^{(134)} = 59.42\%$
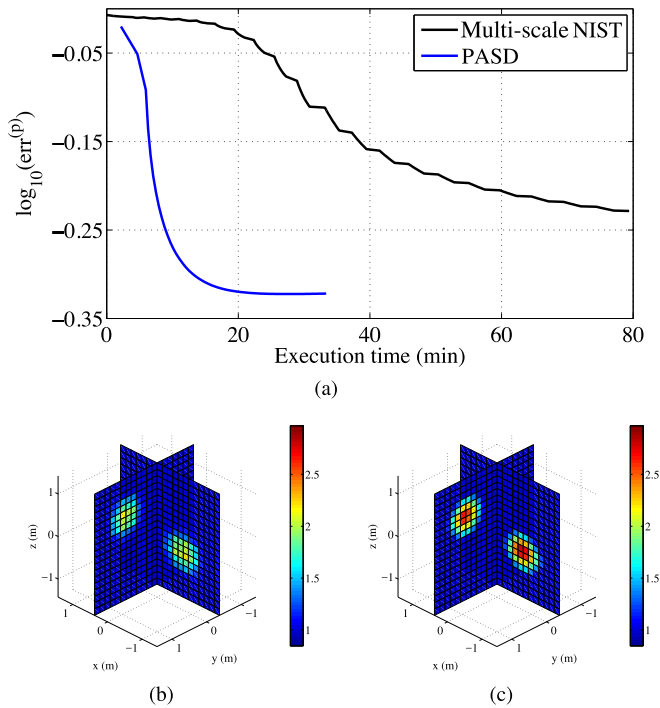
Fig. 11. Investigation domain with two spheres. (a) err$^{(p)}$ versus execution time for PASD and multiscale NIST. Relative permittivity profiles reconstructed by (b) multiscale NIST and (c) PASD.

in 79.4 min, while PASD reaches the same error level in only 8.4 min. Fig. 11(b) and (c) shows the slices of the relative permittivity profiles reconstructed by multiscale NIST at $p = 134$ and PASD at $p = 60$, respectively. It is clear from Fig. 11(b) and (c) that PASD achieves a more accurate reconstruction at convergence.

IV. CONCLUSION

A scheme for efficient and accurate nonlinear electromagnetic imaging of sparse 3-D domains is described. The scheme solves the nonlinear Tikhonov minimization problem, which is constrained with L_0/L_1-norm of the solution, using an accelerated steepest descent algorithm. A projection operator is applied at every iteration of the algorithm to enforce the sparsity constraint. The projected volume is reduced as the iterations of steepest descent algorithm proceed. This increases the convergence rate of the algorithm by avoiding the solution from getting stuck at local minima. Numerical results demonstrate the accuracy and efficiency of the scheme in reconstructing 3-D sparse dielectric profiles.

REFERENCES

[1] M. Pastorino, *Microwave Imaging*. Hoboken, NJ, USA: Wiley, 2010.
[2] D. Colton and R. Kress, *Inverse Acoustic and Electromagnetic Scattering Theory*. Berlin, Germany: Springer, 2012.
[3] R. C. Aster, B. Borchers, and C. H. Thurber, *Parameter Estimation and Inverse Problems*. New York, NY, USA: Academic, 2013.
[4] A. J. Devaney, *Mathematical Foundations of Imaging, Tomography and Wavefield Inversion*. Cambridge, U.K.: Cambridge Univ. Press, 2012.
[5] T. Takagi, J. R. Bowler, and Y. Yoshida, *Electromagnetic Nondestructive Evaluation*. Amsterdam, The Netherlands: IOS Press, 1997.
[6] E. J. Baranoski, "Through-wall imaging: Historical perspective and future directions," *J. Franklin Inst.*, vol. 345, no. 6, pp. 556–569, 2008.
[7] M. Aftanas, "Through wall imaging with UWB radar system," Ph.D. dissertation, Dept. Electron. Multimedia Commun., Tech. Univ. Košice, Košice, Slovakia, 2009.
[8] S. Caorsi, A. Massa, and M. Pastorino, "A crack identification microwave procedure based on a genetic algorithm for nondestructive testing," *IEEE Trans. Antennas Propag.*, vol. 49, no. 12, pp. 1812–1820, Dec. 2001.
[9] W. R. Scott, Jr., C. Schroeder, and J. S. Martin, "A hybrid acoustic/electromagnetic technique for locating land mines," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, vol. 1. Jul. 1998, pp. 216–218.
[10] S. Constable and L. J. Srnka, "An introduction to marine controlled-source electromagnetic methods for hydrocarbon exploration," *Geophysics*, vol. 72, no. 2, pp. WA3–WA12, 2007.
[11] T. J. Cui, W. C. Chew, A. A. Aydiner, and S. Chen, "Inverse scattering of two-dimensional dielectric objects buried in a lossy earth using the distorted Born iterative method," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 2, pp. 339–346, Feb. 2001.
[12] L. Garnero, A. Franchois, J.-P. Hugonin, C. Pichot, and N. Joachimowicz, "Microwave imaging-complex permittivity reconstruction-by simulated annealing," *IEEE Trans. Microw. Theory Techn.*, vol. 39, no. 11, pp. 1801–1807, Nov. 1991.
[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] M. Donelli, D. Franceschini, P. Rocca, and A. Massa, "Three-dimensional microwave imaging problems solved through an efficient multiscaling particle swarm optimization," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 5, pp. 1467–1481, May 2009.
[15] M. Pastorino and A. Randazzo, "Nondestructive analysis of dielectric bodies by means of an ant colony optimization method," in *Swarm Intelligence for Electric and Electronic Engineering*. Hershey, PA, USA: IGI Global, 2012, pp. 308–325.
[16] 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, p. 123003, 2009.
[17] L. C. W. Dixon and G. P. Szegö, Eds., *Towards Global Optimisation 2*. Amsterdam, The Netherlands: North-Holland, 1978.
[18] R. T. F. A. King, H. C. S. Rughooputh, and K. Deb, "Evolutionary multi-objective environmental/economic dispatch: Stochastic versus deterministic approaches," in *Evolutionary Multi-Criterion Optimization*. Berlin, Germany: Springer, 2005, pp. 677–691.
[19] S. D. Rajan and G. V. Frisk, "A comparison between the Born and Rytov approximations for the inverse backscattering problem," *Geophysics*, vol. 54, no. 7, pp. 864–871, 1989.
[20] C. Estatico, M. Pastorino, and A. Randazzo, "An inexact-Newton method for short-range microwave imaging within the second-order Born approximation," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 11, pp. 2593–2605, Nov. 2005.
[21] Y. M. Wang and W. C. Chew, "An iterative solution of the two-dimensional electromagnetic inverse scattering problem," *Int. J. Imag. Syst. Technol.*, vol. 1, no. 1, pp. 100–108, 1989.
[22] 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.
[23] G. Bozza and M. Pastorino, "An inexact Newton-based approach to microwave imaging within the contrast source formulation," *IEEE Trans. Antennas Propag.*, vol. 57, no. 4, pp. 1122–1132, Apr. 2009.
[24] A. Abubakar, P. M. van den Berg, and J. J. Mallorqui, "Imaging of biomedical data using a multiplicative regularized contrast source inversion method," *IEEE Trans. Microw. Theory Techn.*, vol. 50, no. 7, pp. 1761–1771, Jul. 2002.
[25] A. Abubakar, P. M. van den Berg, and T. M. Habashy, "Application of the multiplicative regularized contrast source inversion method on TM- and TE-polarized experimental Fresnel data," *Inverse Problems*, vol. 21, no. 6, p. S5, 2005.
[26] Y. Li and W. Yang, "Image reconstruction by nonlinear Landweber iteration for complicated distributions," *Meas. Sci. Technol.*, vol. 19, no. 9, p. 094014, 2008.
[27] M. Fornasier, Ed., *Theoretical Foundations and Numerical Methods for Sparse Recovery*, vol. 9. Berlin, Germany: Walter de Gruyter, 2010.
[28] I. Daubechies, M. Defrise, and C. De Mol, "An iterative thresholding algorithm for linear inverse problems with a sparsity constraint," *Commun. Pure Appl. Math.*, vol. 57, no. 11, pp. 1413–1457, Nov. 2004.

[29] F. Viani, L. Poli, G. Oliveri, F. Robol, and A. Massa, "Sparse scatterers imaging through approximated multitask compressive sensing strategies," *Microw. Opt. Technol. Lett.*, vol. 55, no. 7, pp. 1553–1558, 2013.

[30] L. C. Potter, E. Ertin, J. T. Parker, and M. Cetin, "Sparsity and compressed sensing in radar imaging," *Proc. IEEE*, vol. 98, no. 6, pp. 1006–1020, Jun. 2010.

[31] 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.

[32] A. Desmal and H. Bağcı, "Shrinkage-thresholding enhanced Born iterative method for solving 2D inverse electromagnetic scattering problem," *IEEE Trans. Antennas Propag.*, vol. 62, no. 7, pp. 3878–3884, Jul. 2014.

[33] A. Desmal and H. Bağcı, "A preconditioned inexact Newton method for nonlinear sparse electromagnetic imaging," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 3, pp. 532–536, Mar. 2015.

[34] A. Desmal and H. Bağcı, "Sparse electromagnetic imaging using nonlinear Landweber iterations," *Prog. Electromagn. Res.*, vol. 152, pp. 77–93, 2015.

[35] N. Anselmi, M. Salucci, G. Oliveri, and A. Massa, "Wavelet-based compressive imaging of sparse targets," *IEEE Trans. Antennas Propag.*, vol. 63, no. 11, pp. 4889–4900, Nov. 2015.

[36] A. I. Sandhu, A. Desmal, and H. Bağcı, "A sparsity-regularized Born iterative method for reconstruction of two-dimensional piecewise continuous inhomogeneous domains," in *Proc. IEEE 10th Eur. Conf. Antennas Propag. (EuCAP)*, Apr. 2016, pp. 1–3.

[37] A. Massa, P. Rocca, and G. Oliveri, "Compressive sensing in electromagnetics—A review," *IEEE Antennas Propag. Mag.*, vol. 57, no. 1, pp. 224–238, Feb. 2015.

[38] G. Teschke and C. Borries, "Accelerated projected steepest descent method for nonlinear inverse problems with sparsity constraints," *Inverse Problems*, vol. 26, no. 2, p. 025007, 2010.

[39] I. Daubechies, M. Fornasier, and I. Loris, "Accelerated projected gradient method for linear inverse problems with sparsity constraints," *J. Fourier Anal. Appl.*, vol. 14, no. 5, pp. 764–792, 2008.

[40] A. F. Peterson, S. L. Ray, and R. Mittra, *Computational Methods for Electromagnetics*, vol. 24. New York, NY, USA: IEEE Press, 1998.

[41] H. A. van der Vorst, "Bi-CGSTAB: A fast and smoothly converging variant of Bi-CG for the solution of nonsymmetric linear systems," *SIAM J. Sci. Statist. Comput.*, vol. 13, no. 2, pp. 631–644, 1992.

[42] H. Gan and W. C. Chew, "A discrete BCG-FFT algorithm for solving 3D inhomogeneous scatterer problems," *J. Electromagn. Waves Appl.*, vol. 9, no. 10, pp. 1339–1357, 1995.

[43] P. Zwamborn and P. M. van den Berg, "The three dimensional weak form of the conjugate gradient FFT method for solving scattering problems," *IEEE Trans. Microw. Theory Techn.*, vol. 40, no. 9, pp. 1757–1766, Sep. 1992.

[44] T. Blumensath and M. E. Davies, "Iterative hard thresholding for compressed sensing," *Appl. Comput. Harmon. Anal.*, vol. 27, no. 3, pp. 265–274, Nov. 2009.

[45] B. Kaltenbacher, A. Neubauer, and O. Scherzer, *Iterative Regularization Methods for Nonlinear Ill-Posed Problems*, vol. 6. Berlin, Germany: Walter de Gruyter, 2008.

[46] A. Baussard, E. L. Miller, and D. Prémel, "Adaptive *B*-spline scheme for solving an inverse scattering problem," *Inverse Problems*, vol. 20, no. 2, p. 347, 2004.

[47] F. Mokhtarian and A. K. Mackworth, "A theory of multiscale, curvature-based shape representation for planar curves," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 14, no. 8, pp. 789–805, Aug. 1992.

[48] R. Firoozabadi, E. L. Miller, C. M. Rappaport, and A. W. Morgenthaler, "Subsurface sensing of buried objects under a randomly rough surface using scattered electromagnetic field data," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 1, pp. 104–117, Jan. 2007.

Abdulla Desmal (S'11) received the B.S. degree in electronics engineering from the University of Bahrain (UOB), Isa-Town, Bahrain, in 2009, and the M.S. and Ph.D. degrees in electrical engineering from the King Abdullah University of Science and Technology (KAUST), Thuwal, Saudi Arabia, in 2010 and 2016, respectively.

From 2009 to 2016, he was a Research Assistant with the Division of Computer, Electrical, and Mathematical Science and Engineering, KAUST. Since 2016, he has been a Post-Doctoral Research Fellow at Tufts University, Medford, MA, USA. His research interests include the development of efficient forward electromagnetic solvers and sparsity-regularized inverse scattering algorithms.

Dr. Desmal was awarded the Rasheed Al-Deen Bin Al-Soori Award for excellence in academic achievement by UOB in 2015. He has authored two finalist papers in the student paper competitions at the 2014 IEEE Antennas and Propagation Society International Symposium and the 2014 Applied Computational Electromagnetics Society Conference.

Hakan Bağcı (S'98–M'07–SM'14) received the B.S. degree in electrical and electronics engineering from Bilkent University, Ankara, Turkey, in 2001, and the M.S. and Ph.D. degrees in electrical and computer engineering from the University of Illinois at Urbana–Champaign (UIUC), Urbana, IL, USA, in 2003 and 2007, respectively.

From 1999 to 2001, he was an Undergraduate Researcher with the Computational Electromagnetics Group, Bilkent University. From 2001 to 2006, he was a Research Assistant with the Center for Computational Electromagnetics and Electromagnetics Laboratory, UIUC. From 2007 to 2009, he was a Research Fellow with the Radiation Laboratory, University of Michigan, Ann Arbor, MI, USA. In 2009, he joined the Division of Computer, Electrical, and Mathematical Science and Engineering, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia, as an Assistant Professor of Electrical Engineering and became an Associate Professor of Electrical Engineering in 2015. His research interests include various aspects of computational electromagnetics with an emphasis on time-domain integral equations and their fast marching-on-in-time-based solutions, well-conditioned integral-equation formulations, and the development of fast hybrid methods for analyzing statistical EMC/EMI phenomena on complex and fully loaded platforms.

Dr. Bağcı was a recipient of the 2008 International Union of Radio Scientists Young Scientist Award and the 2004–2005 Interdisciplinary Graduate Fellowship from the Computational Science and Engineering Department, UIUC. His paper titled "Fast and Rigorous Analysis of EMC/EMI Phenomena on Electrically Large and Complex Structures Loaded With Coaxial Cables" was one of the three finalists (with honorable mention) for the 2008 Richard B. Schulz Best TRANSACTIONS Paper Award given by the IEEE Electromagnetic Compatibility Society. He has authored or co-authored 13 finalist papers in the student paper competitions at the 2005, 2008, 2010, 2014, 2015, and 2016 IEEE Antennas and Propagation Society International Symposiums and the 2013, 2014, and 2016 Applied Computational Electromagnetics Society Conferences.