Microwave Medical Imaging Based on Sparsity and an Iterative Method With Adaptive Thresholding

Masoumeh Azghani*, Panagiotis Kosmas, and Farokh Marvasti

Abstract—We propose a new image recovery method to improve the resolution in microwave imaging applications. Scattered field data obtained from a simplified breast model with closely located targets is used to formulate an electromagnetic inverse scattering problem, which is then solved using the Distorted Born Iterative Method (DBIM). At each iteration of the DBIM method, an underdetermined set of linear equations is solved using our proposed sparse recovery algorithm, IMATCS. Our results demonstrate the ability of the proposed method to recover small targets in cases where traditional DBIM approaches fail. Furthermore, in order to regularize the sparse recovery algorithm, we propose a novel L_2-based approach and prove its convergence. The simulation results indicate that the L_2-regularized method improves the robustness of the algorithm against the ill-posed conditions of the EM inverse scattering problem. Finally, we demonstrate that the regularized IMATCS-DBIM approach leads to fast, accurate and stable reconstructions of highly dense breast compositions.

Index Terms—Adaptive thresholding, breast imaging, compressed sensing, inverse scattering, microwave tomography.

I. INTRODUCTION

MICROWAVE tomographic methods for clinical applications estimate the spatial distribution of dielectric properties in a tissue region by solving an electromagnetic (EM) inverse scattering problem [1]. Various EM inverse scattering methods have been proposed in recent years for this purpose, such as conjugate gradient techniques [2], [3] and Gauss-Newton (GN) optimization algorithms [4]–[6]. In this paper, microwave tomography is implemented by applying the Distorted Born Iterative Method (DBIM) which, as any GN approach, approximates the nonlinear inverse scattering problem with an underdetermined set of linear equations.

We propose a novel solution to the resulting linear system at every DBIM iteration based on an Iterative Method with Adaptive Thresholding for Compressed Sensing (IMATCS) [7]. In the IMATCS method, a crude reconstruction is applied succes-

sively to the linear measurements of the signal and the recovered signal at each iteration is sparsed using an adaptive thresholding function. We should note here the existence of other thresholding techniques such as iterative hard thresholding (IHT) and K-sparse algorithms [8], [9]. The performance of these IHT algorithms, however, depends highly on the value of the IHT threshold, which should be determined by trial and error. On the contrary, an important advantage of the proposed IMATCS method over algorithms such as the iterative K-sparse method is that it does not require the knowledge of the sparsity number K of the signal as *a priori* information for signal recovery. This important advantage is based on the IMATCS' adaptive thresholding approach, which enables the algorithm to pick up the significant signal entries at each iteration.

The proposed approach belongs to a wider category of sparsity regularization techniques, which are currently pursued in MicroWave Imaging (MWI). In [10], for example, an L_1 regularizer is exploited to enforce sparsity in contrast-enhanced MWI of breast tumors. The elastic net method is proposed in [11] to solve the linear problem within the DBIM, resulting in an improvement in the reconstruction of the breast interior. These methods are inspired by Compressed Sensing (CS) theory [12], [13], which can recover a sparse signal from a lower dimensional measurement vector, i.e., the number of measurements is much less than the number of signal entries. Hence, CS methods are suitable for the solution of underdetermined systems of equations using concepts of sparsity of the underlying signal.

We note that, while CS methods and the IMATCS are applicable in various domains such as sparse data acquisition or signal compression [14], their application in this work refers to the solution of an underdetermined system of linear equations based on the measurement matrix at each iteration of the DBIM algorithm. Contrary to CS applications which attempt to design the measurement matrix to satisfy appropriate conditions, the measurement matrix in this work is formed by a conventional microwave tomography setup adopted in our simulations.

The nonlinearity and ill-posedness of the EM inverse scattering problem can lead to unstable reconstructions in MWI medical applications, which involve the presence of dense or closely located scatterers in the reconstructed region. To improve robustness, regularization techniques must be applied at the expense of compromised resolution in the resulting images [15]. We hereby propose a novel L_2-regularized version of IHT which can lead to stable signal recovery, and present its convergence proof. In particular, we solve the L_2-regularized L_0-minimization problem using the Majorization-Minimization (MM) method as used in [9] for the derivation of the IHT

algorithm. Similar to IHT, the performance of the L_2-IHT relies on the proper selection of the threshold value based on *a priori* knowledge of the underlying signal. To overcome this limitation, we propose an L_2-IMATCS method, which is a combination of the IMATCS and L_2-IHT algorithms that leads to stable signal recovery in scenarios where the original IMATCS becomes unstable.

In order to demonstrate that the proposed method can improve resolution in DBIM-based MWI, we examine simplified homogeneous breast models with closely located tumor-like scaterrers of different size. This choice is motivated by ongoing research on microwave breast imaging and breast cancer detection, which is arguably the most advanced MWI medical application [16]. The IMATCS algorithm allows the DBIM method to reconstruct these closely located targets, which cannot be resolved with traditional L_2 regularization schemes. We also test the ability of the regularized version of our algorithm to image complex breast structures by considering a highly heterogenous breast distribution, which is reconstructed using multiple-frequency data.

The rest of the paper is organized as follows. A review of EM inverse scattering using the DBIM and CS theory together with the IMATCS approach is given in Section II. The development of novel L_2-IHT and L_2-IMATCS methods is presented in Section III. Simulation results for various scenarios are given in Section IV, and Section V concludes this work. Finally, we provide a mathematical proof for the convergence of the algorithm in the Appendix. A preliminary version of this work has been reported in [17].

II. COMPRESSED SENSING IN MICROWAVE IMAGING

A. EM Inverse Scattering With the DBIM

The DBIM algorithm is based on the distorted-wave Born integral equation [18]

$$\mathbf{E}^{sc}(\mathbf{r}) = \mathbf{E}^t(\mathbf{r}) - \mathbf{E}^b(\mathbf{r}) = \omega^2\mu \int_V d\mathbf{r}' \bar{\mathbf{G}}^b(\mathbf{r},\mathbf{r}')\Delta\epsilon(\mathbf{r}')\mathbf{E}^t(\mathbf{r}')$$

$$(1)$$

where $\mathbf{E}^{sc}$ is the scattered electric field, $\Delta\epsilon(\mathbf{r}') = \epsilon(\mathbf{r}') - \epsilon_b(\mathbf{r}')$ is the unknown contrast function over the volume of support V, and $\bar{\mathbf{G}}^b(\mathbf{r},\mathbf{r}')$ denotes the dyadic background Green's function, which represents propagation from the source located at $\mathbf{r}$ to the point $\mathbf{r}'$ inside V. The unknown total field $\mathbf{E}^t$ is approximated with the known background field $\mathbf{E}^b$ inside the integral in (1). The resulting linear integral equation is solved in the discrete domain at each iteration of the DBIM algorithm for the contrast function $\Delta\epsilon(\mathbf{r}')$.

This approach leads to a GN algorithm for nonlinear least-squares problems [19]. At iteration k, the algorithm solves a linear least-squares problem described by

$$\bar{J}_k^T \bar{J}_k \mathbf{p}_k^{GN} = \bar{J}_k^T \mathbf{r}_k \qquad (2)$$

where $\mathbf{p}_k$ is the desired GN direction vector corresponding to the unknown $\Delta\epsilon(\mathbf{r}')$, $\mathbf{r}_k$ is the residual vector and $\bar{J}_k$ is the Jacobian matrix of the least-squares problem. The forward model is run at each iteration to compute $\mathbf{E}^b$ and $\bar{\mathbf{G}}^b$, which are then used to calculate $\mathbf{r}_k$ and the Jacobian $\bar{J}_k$. The resulting system

(2) is solved for $\mathbf{p}_k$ the vectorized update $\Delta\epsilon(\mathbf{r})$, and the new value is added to the previous estimate of $\epsilon(\mathbf{r})$ [25].

The choice of approach for solving the linear system in (2) must consider two important factors. First, the matrix J_k is typically ill-conditioned and requires a regularization method to ensure stability. Second, the solution of this linear system for large-scale problems is computationally very demanding; therefore, a cost-efficient strategy is necessary. Previous 3-D MWI algorithms have employed a Conjugate Gradient Least Squares (CGLS) method with add-hoc stopping criteria to regularize and solve (2) efficiently [5], [20]. This work uses CS theory to solve the linear system of equations resulting from the DBIM approximation in (1) as described below.

B. Application of CS Theory

We can express the linear problem resulting from (1) at each iteration of the DBIM as

$$\mathbf{y} = \mathbf{\Phi}\mathbf{s} \qquad (3)$$

where $\mathbf{\Phi}$ is an $M \times N$ measurement matrix $(M < N)$ formed by $\bar{\mathbf{G}}^b$ and $\mathbf{E}^b$, $\mathbf{s}$ represents the N-dimensional vector of unknown dielectric contrast function update $\Delta\epsilon$, and $\mathbf{y}$ represents the residual measurement data vector $\mathbf{E}^{sc}$ [5]. Our aim is to recover $\mathbf{s}$ from $\mathbf{y}$, given that the number of measurements M is much less than the number of unknowns N. In general, CS theory solves this underdetermined problem by assuming that $\mathbf{s}$ is sparse [12], [13]. If $\mathbf{s}$ is not sparse, a transformation using appropriate basis functions which induce sparsity can be used

$$\mathbf{s} = \mathbf{\Psi}\mathbf{x} \qquad (4)$$

leading to a reformulation of (3) as

$$\mathbf{y} = \mathbf{\Phi}\mathbf{\Psi}\mathbf{x} = \mathbf{A}\mathbf{x} \qquad (5)$$

where $\mathbf{A} = \mathbf{\Phi}\mathbf{\Psi}$. To solve (5) under the constraint that $\mathbf{x}$ is sparse, we search for the most sparse solution among all the vectors satisfying the constraint $\mathbf{y} = \mathbf{A}\mathbf{x}$. We note here that sparsity is defined as the number of nonzero entries of $\mathbf{x}$, i.e., the L_0 norm of $\mathbf{x}$. Hence, the main CS problem becomes

$$\min_{\mathbf{x}} \|\mathbf{x}\|_0 \text{ subject to } \mathbf{y} = \mathbf{A}\mathbf{x}. \qquad (6)$$

Equation (6) can be solved by an exhaustive search method which is NP-hard. A tractable approach to solve the nonconvex problem (6) is to approximate it with the convex problem

$$\min_{\mathbf{x}} \|\mathbf{x}\|_1 \text{ subject to } \mathbf{y} = \mathbf{A}\mathbf{x}. \qquad (7)$$

The above problem is linked to the Basis Pursuit (BP) algorithm [21], or equivalently the Least Absolute Shrinkage and Selection Operator (LASSO) [22] algorithm that can be solved using convex optimization techniques. Different types of CS recovery methods include the family of greedy algorithms such as Matching Pursuit (MP), Orthogonal Matching Pursuit (OMP), and gradient pursuit [23], [24]. These algorithms are computationally more efficient than the L_1-norm minimization techniques at the expense of lower guaranteed recovery performance. Thresholding techniques such as IHT [8], [9] and our

proposed IMATCS [7] come in between these two groups in both recovery performance and simplicity.

Most CS recovery techniques require that special conditions are satisfied by the measurement matrix Φ in order to recover the signal successfully; this leads to the design of measurement matrices such as random Gaussian, Bernoulli, and Toeplitz matrices. In EM inverse scattering, however, the measurement matrix is imposed by the EM integral equation where straightforward manipulation of this matrix is not possible. Moreover, these EM equations are inherently nonlinear and are only approximated to a linear form shown in (1). These two factors will inevitably affect the performance of CS recovery methods in realistic EM inverse scattering applications.

In addition, the unknown vector of dielectric contrast update $\mathbf{s}$ will not be sparse at every DBIM iteration in many realistic EM inverse scattering problems, such as the imaging of heterogeneous dense breasts without *a priori* information. Finding a transformation matrix Ψ that can induce sparsity in (4) is a challenging task in these cases. This work does not address this topic but manages to reconstruct nonsparse $\mathbf{s}$ vectors by introducing the L_2-IMATCS approach presented in Section III. Therefore, we consider in (4) that Ψ is the identity matrix $\mathbf{I}$, so that $\mathbf{A} = \Phi$ is the measurement matrix defined in the standard DBIM approach.

C. The IMATCS Algorithm

As mentioned in the introduction, the IMATCS method belongs to the family of thresholding techniques. The IHT [8] solves the following optimization problem:

$$\min_{\mathbf{x}} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 + \lambda \|\mathbf{x}\|_0. \tag{8}$$

The solution of (8) can be written as

$$\mathbf{x}_{k+1} = H\left(\mathbf{x}_k + \lambda \mathbf{A}^*(\mathbf{y} - \mathbf{A}\mathbf{x}_k)\right) \tag{9}$$

where $\mathbf{A}^*$ denotes the conjugate transpose of matrix $\mathbf{A}$, λ is the relaxation parameter which controls the convergence of the algorithm, and H is a hard thresholding function which discards all the coefficients with amplitudes smaller than a predefined fixed threshold. The threshold value should be finely tuned using *a priori* information of the underlying signal. However, it is usually impossible to have *a priori* knowledge of the signal before its recovery. In order to get around this problem, IMATCS takes advantage of an adaptive thresholding procedure with a threshold initial predefined value that decreases exponentially at each iteration. This adaptive property of the thresholding function relaxes the theoretical requirement of knowing the exact value of the hard threshold.

As a result, the mathematical formulation of the IMATCS method is the same as (9) with the difference that the thresholding function H decreases with each iteration in an exponential manner given as

$$T_i = T_0 e^{-\alpha i} \tag{10}$$

where i is the iteration number, T_0 is the initial threshold value, and α indicates the threshold step. The algorithm given in (9) starts from a zero initial value, $\mathbf{x}_0 = 0$. The coefficient vector
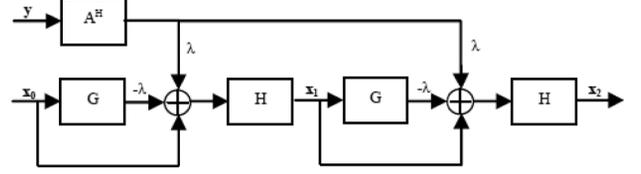

Fig. 1. Block diagram of the IMATCS method.

is recovered as $\mathbf{x}_{itermax}$, after an "*itermax*" number of iterations. The adaptivity of the threshold enables us to recover the embedding signal from its linear measurements without any knowledge of the underlying signal. The block diagram of the IMATCS method is depicted in Fig. 1, where $\mathbf{G} = \mathbf{A}^*\mathbf{A}$. The $\mathbf{G}$ operator in Fig. 1 captures the sampling (using $\mathbf{A}$) and crude reconstruction (using $\mathbf{A}^*$) processes, which are applied iteratively according to (9). At the end of each iteration, the signal is sparsed using the thresholding function H.

While the IMATCS algorithm can be applied successfully to recover the image from its compressive measurements, it must be emphasized that the set of equations in (2) is not exactly linear but approximated to be linear in MWI applications. This linear approximation can result in instability which can cause the IMATCS (or any other CS-based recovery algorithm) to diverge after some iterations. In order to come around this problem, an L_2-regularized IMATCS algorithm is proposed in the next section which incorporates the L_2-regularization strategy into the sparsity-based recovery algorithm in order to stabilize the recovery procedure.

III. IMPLEMENTATION OF REGULARIZED ADAPTIVE THRESHOLDING CS METHODS

A. Formulation

In order to deal with the ill-posedness of the EM inverse scattering problem, we can reformulate (8) to solve the following minimization problem:

$$\min_{\mathbf{x}} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 + \lambda_1 \|\mathbf{x}\|_0 + \lambda_2 \|\mathbf{x}\|_2^2. \tag{11}$$

The above cost function is an L_2-regularized L_0-minimization approach, L_0/L_2. Previous CS algorithms applied to MWI such as the elastic net [11] have minimized a cost function based on an L_2-regularized L_1-minimization, which is solved by reshaping the problem to a LASSO cost function in a larger dimensional subspace. The complexity of LASSO increases with the dimension of the underlying signal; hence, the elastic net has a great complexity and computational burden. Here, we propose an L_2-IHT method which is much simpler and more robust as a linear solver at each DBIM iteration in (2).

We propose to solve the L_0/L_2 regularized problem in (11) using the following iterative method, called L_2-IHT, given by

$$\mathbf{x}_{k+1} = \frac{1}{1+\lambda_2} H\left(\mathbf{x}_k + \lambda_1 \mathbf{A}^*(\mathbf{y} - \mathbf{A}\mathbf{x}_k)\right). \tag{12}$$

The mathematical derivation of this novel approach is given in the Appendix. The performance of the L_2-IHT algorithm is extremely sensitive to a proper selection of the threshold value, and therefore the method cannot guarantee convergence to an

acceptable solution. In order to deal with this problem, we have implemented an L_2-IMATCS method which combines the iterative thresholding approach of IMATCS with the L_2-IHT algorithm. The formulation of the L_2-IMATCS is the same as (12) with the difference that the H function is adaptively decreased according to (10).

B. Parameters Selection

The performance of the proposed L_2-IMATCS method is determined by the choice of the regularization parameters λ_1 and λ_2, and the thresholding function parameters T_0, α, and "$itermax$." The parameter λ_1 controls the convergence of the algorithm and should satisfy the following condition:

$$0 \leq \lambda_1 \leq \frac{2}{\max\, eig(\mathbf{A}^*\mathbf{A})} \tag{13}$$

where eig denotes the matrix eigenvalues, and $\mathbf{A}$ is the measurement matrix in (5), which is updated at each DBIM iteration. In the simulations of the next section, we have set λ_1 as

$$\lambda_1 = \frac{1.9}{\max\, eig(\mathbf{A}^*\mathbf{A})}. \tag{14}$$

The introduction of λ_2 controls the stability of the algorithm by promoting L_2-based solutions of the minimization problem in (11). However, selecting a large value for this parameter will reduce the impact of the other terms in (11) and can therefore increase estimation errors in the recovered image. Reconstructions with different values of λ_2 have confirmed this trade-off between the algorithm's stability and imaging accuracy.

The thresholding parameters T_0, α, and "$itermax$" are responsible for the quality and the resolution of the reconstructed image. In order to determine the thresholding parameters, we have followed the following "ad-hoc" process. First, we set a sufficiently large value for T_0, say 50, a small value for α, say 0.01, and a large integer number for "$itermax$," say 300. These initial values generate a slowly decreasing thresholding function which can capture almost all the significant components of the signal, producing a reliable estimation of the true signal. The method is simulated using these initial parameters. The recovered signal at the end of each DBIM iteration is investigated to detect the maximum signal entry. Having detected the maximum value of the underlying signal to be around 0.04 for all DBIM iterations, T_0 is set to a slightly larger value equal to 0.07. Then the optimal α is obtained by decreasing its value without altering the result obtained by the initial parameters. The resultant value for α is 0.2. Finally, "$itermax$" is selected by observing the residual error curve with respect to the iteration number. In particular, the optimal value of "$itermax$" is where the residual error has reached its minimum value (200) for the first time and does not change after that. It is important to note that this process is only run for one case and the resulting parameters are used in all other cases.

IV. RESULTS

This section presents simulation results inspired by microwave breast imaging, which illustrates the advantages of the proposed method. We first consider two closely located tumor-like scatterers in simplified breast models, where the
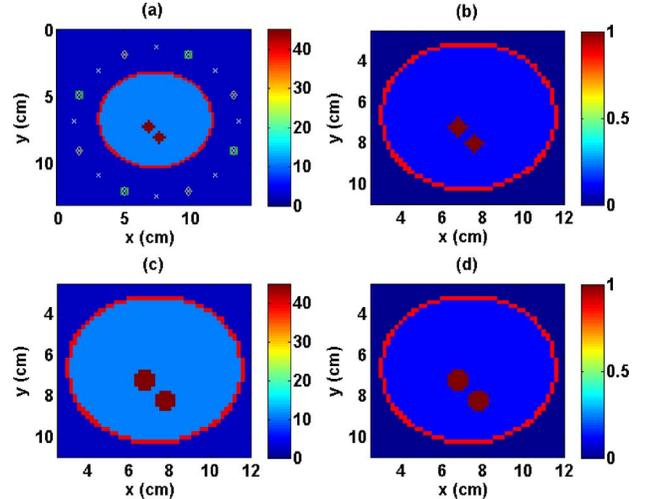

Fig. 2. Maps of (a), (c) the dielectric constant ϵ_r and (b), (d) the conductivity σ calculated at 1 GHz for the two simulation scenarios, which differ only in the size and shape of the targets. Locations of data points for a sixteen-, eight-, and four-antenna configuration are also shown in (a) with crosses, diamonds, and squares, respectively.

dielectric distribution is sparse in the reconstruction domain. Then, we demonstrate the robustness of our regularized IMATCS-DBIM algorithm in reconstructing complex breast structures. The presented breast models, simulations, and reconstructions are two-dimensional (2-D), but they can be extended in three dimensions (3-D) similar to previous work based on the same setup and methods [5]. We consider ideal dipole antennas as transceivers in our simulations, which correspond to point sources for our 2-D geometry and are arranged cylindrically to encircle the breast slice at a distance of 1–2 cm away from the skin. These point sources illuminate the breast with a wideband pulse (with a -20 dB bandwidth from 500 MHz to 3.5 GHz) sequentially, and also record the data to be used for the solution of the microwave tomographic problem. The effect of the number of these data points on the reconstructions in the presence of Gaussian noise is also considered below.

A. Simulation Testbeds

We simulate measured data using the Finite-Difference-Time-Domain (FDTD) method and a uniform grid cell size of 2.0 mm, which is also used as the forward solver in the inversion process. This "inverse crime" assumption allows us to benchmark the performance of our approach. The algorithm estimates the parameters ϵ_∞, ϵ_s, and σ_s of the Debye model for the complex relative permittivity

$$\epsilon_r(\omega) = \epsilon_\infty + \frac{\epsilon_s - \epsilon_\infty}{1 + j\omega\tau} - j\frac{\sigma_s}{\omega\epsilon_0} \tag{15}$$

where τ is assumed constant for all tissues (with a value of 17.1 ps). The background medium is assumed lossless with $\epsilon_r = 2.6$ in all our simulation testbeds, and the Debye parameters for the various tissues are extracted from the University of Wisconsin–Madison's repository data [26].

First, we have considered the testbeds depicted in Fig. 2. The testbeds resemble a two-dimensional coronal slice of a breast
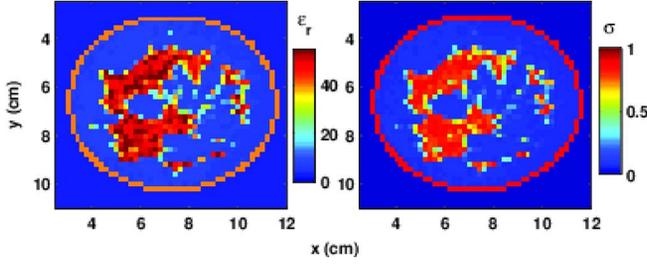
Fig. 3. Maps of the dielectric constant ϵ_r (left) and the conductivity σ (right) calculated at 1 GHz for the full-breast simulation testbed. Spatial resolution is 2 mm.

comprised of homogenous breast tissue characterized by volumetric average Debye parameters ($\epsilon_r = 11.1$ and $\sigma = 0.11\,\mathrm{S/m}$ at 1 GHz), a skin layer, and two identical closely located tumor-like targets. The dielectric constant and conductivity distributions calculated at 1 GHz are shown in Fig. 2 for the two scenarios. The two cases differ only in the size and shape of the two closely located tumors in order to test the resolution capabilities of our proposed algorithm for 16, 8, and 4 transceivers with locations shown in Fig. 2(a).

The testbeds in Fig. 2 are useful in studying the resolution abilities of our sparsity promoting algorithm in idealized scenarios, but they do not correspond to a realistic microwave breast imaging application, where the breast interior can be highly heterogeneous leading to a severely ill-posed inverse problem. To assess the feasibility of using the L_2-IMATCS algorithm for these more complicated scenarios, we have applied the previous setup to a 2-D coronal slice from a "heterogeneous breast" phantom taken from University of Wisconsin–Madison's online breast phantom repository [26]. The dielectric constant and conductivity distributions of the considered coronal slice are shown in Fig. 3.

B. Reconstructions of Two Closely Located Targets in a Homogeneous Numerical Breast Phantom

Fig. 4 presents reconstructions of the dielectric constant and conductivity distributions for the two simulation scenarios of Fig. 2. These images were obtained by the L_2-IMATCS algorithm using data from six equally spaced frequencies in the range 1.2–2.7 GHz. We used $\lambda_2 = 0.005$ in (11) for these reconstructions, as higher values resulted in lower resolution. Only the two tumor-like targets were unknown in the breast model used as the "initial guess" for the DBIM algorithm. The results in Fig. 4 demonstrate that the L_2-IMATCS approach manages to resolve the two targets and is sensitive to their size and shape. The conductivity images are less accurate than those of the dielectric constant, which agrees with previous results related to microwave tomography [4].

Importantly, traditional L_2-based methods based on CGLS or LSQR linear solvers could not resolve the two targets in the above cases. Furthermore, our simulation attempts with some other popular CS-based methods were not successful in imaging these scenarios; for example, the LASSO and OMP methods diverged after the first iteration, while the imaging performance of the IHT algorithm depended strongly on the sparsity number of the dielectric contrast vector, which is not known *a priori*.
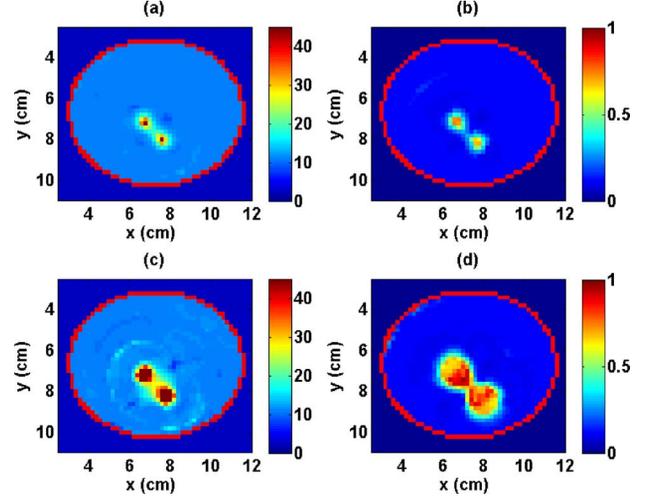

Fig. 4. (a), (c) Reconstructed dielectric constant ϵ_r and (b), (d) conductivity σ distributions calculated at 1 GHz for the L_2-IMATCS algorithm and the cases depicted in Fig. 2 with all 16 antennas depicted in Fig. 2(a).
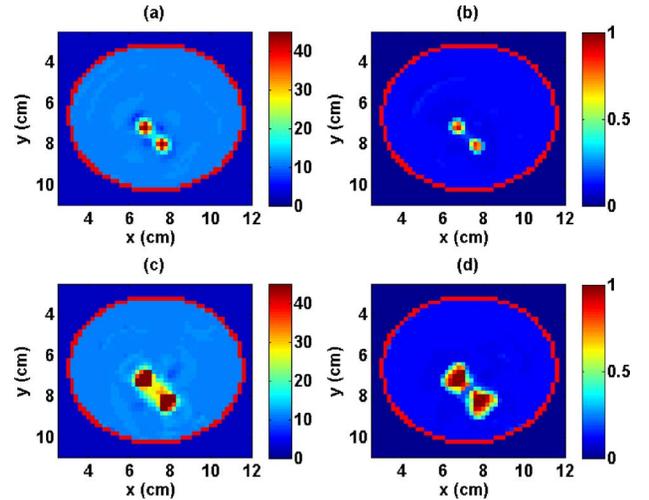

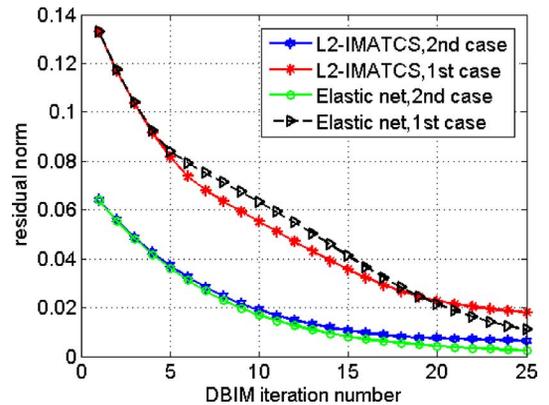Fig. 5. Same as in Fig. 4 for the elastic net algorithm [11].

Fig. 6. Norm of the residual data vector b in (2) versus DBIM iteration number for the reconstructions of Figs. 4 and 5.

We have also tested the elastic net approach of [11] for the cases in Fig. 2, and the resulting reconstructions are comparable to our proposed method as shown in Fig. 5. The similar convergence performance of the two algorithms is confirmed by Fig. 6,
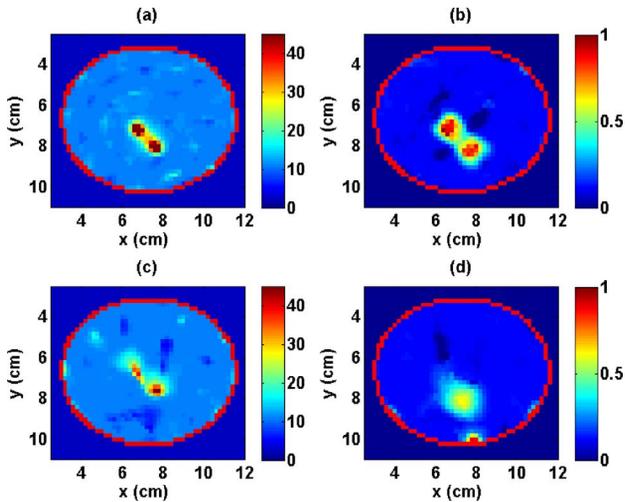
Fig. 7. (a), (c) Reconstructed dielectric constant ϵ_r and (b), (d) conductivity σ distributions calculated at 1 GHz for the L_2-IMATCS algorithm and the top case in Fig. 2, for eight (top) and four (bottom) antennas.

Fig. 8. (a), (c) Reconstructed dielectric constant ϵ_r and (b), (d) conductivity σ distributions calculated at 1 GHz for the L_2-IMATCS algorithm and the top case in Fig. 2, for SNR $=$ 60 dB (top) and SNR $=$ 30 dB (bottom), and sixteen antennas. The SNR is relative to the total received signal, which includes direct antenna contributions and skin reflections.

where the norm of the residual vector $\mathbf{r}_k$ in (2) versus the iteration k of the DBIM algorithm are plotted. Although the reconstructions of the two algorithms in Figs. 4 and 5 are of comparable quality, the computational cost of the elastic net implementation is considerably higher than that of the L_2-IMATCS algorithm. Tested on a standard PC Matlab environment, the elastic net solver took almost 10 min while the L_2-IMATCS algorithm took about 30 s for each DBIM iteration. This order-of-magnitude difference in computation times becomes a crucial factor in 3-D problems. In addition, the L_2-IMATCS algorithm proved to be more robust when the ill-posedness of the inverse problem was increased, as in the second set of simulations described in Section IV-C where the elastic net approach failed to lead to meaningful reconstructions.

We have also studied the robustness of our algorithm with respect to the number of antennas available for data acquisition and the Signal-To-Noise ratio (SNR) relative to the total recorded signal. To this end, we have first repeated the reconstruction process (keeping the same value for λ_2) for the top case of Fig. 4 using eight and four antennas shown in Fig. 2(a). The resulting permittivity and conductivity images are shown in Fig. 7. It is evident that good images can be obtained with eight antennas, but data from four antennas is not sufficient to resolve the two targets well, particularly in the estimation of their conductivity. For limited data setups, the use of more frequencies within a wider bandwidth could be used to obtain more data and improve reconstructions.

Similarly, good results are obtained when Gaussian noise is added to the simulated data even for a very low SNR of 30 dB relative to the total received signal (which includes direct antenna contributions and skin reflections). The resulting reconstructions are shown in Fig. 8 and confirm that the regularized L_2-IMATCS algorithm is robust in the presence of additive noise, allowing a flexible formulation of the $L0/L2$ minimization problem by adjusting the regularization parameter λ_2 (which for example was increased from 0.005 to 0.01 to handle the SNR $=$ 30 dB case).
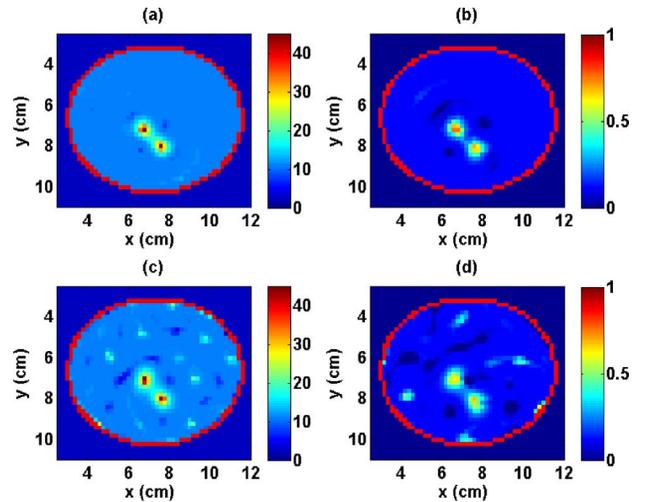
C. Reconstructions of a Heterogeneous Numerical Breast Phantom

Similar to the previous results, full breast reconstructions were obtained by using the L_2-IMATCS algorithm with data from six equally spaced frequencies in the range 1.2–2.7 GHz. This more complex imaging scenario, however, required an initial low-frequency reconstruction at 1 GHz prior to using multiple-frequency data, and a higher value of the regularization parameter $\lambda_2 = 0.2$ was used. As in [5], the algorithm used *a priori* knowledge of the breast surface and its volumetric average Debye parameters as "initial guess" for the homogeneous breast interior, but did not include any *a priori* info on the skin layer thickness or properties. As in Section IV-B, we have tested the algorithm for sixteen or eight antennas and 60 or 30 dB SNR. Comparing the images in Fig. 9 to the true distributions in Fig. 5 suggests that good reconstructions can be obtained for eight antennas and low SNRs. As before, the conductivity images are slightly less accurate than those of the dielectric constant; nonetheless, these results suggest that the proposed algorithm can reconstruct complex dielectric profiles by suppressing its sparsity-promoting minimization term and using monochromatic low-frequency data in the first iterations of the DBIM algorithm.

This latter strategy is also important to ensure that the L_2-IMATCS algorithm is robust to the choice of initial guess, which can be critical to the imaging performance of GN optimization methods. For example, using low adipose tissue Debye parameters for the initial homogenous breast interior instead of its true volumetric average parameters required a lower frequency of 700 MHz (instead of 1 GHz) as initial step in the first iterations of the DBIM algorithm. While the true volumetric average Debye parameters will not be known *a priori*, various methods can be used to estimate these average values or provide a coarse heterogeneous initial guess [3].
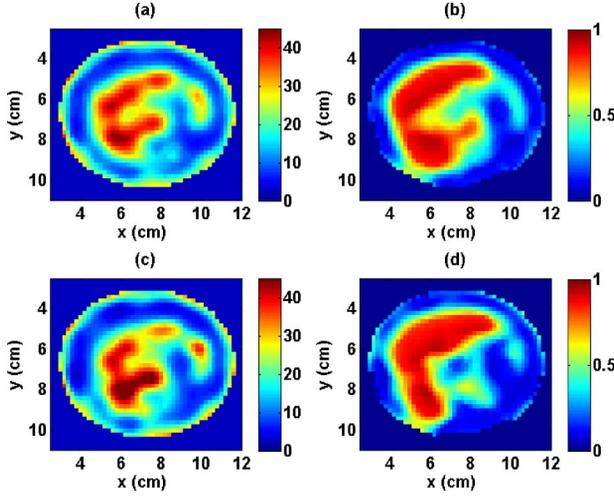
Fig. 9. (a), (c) Reconstructed dielectric constant ϵ_r and (b), (d) conductivity σ distributions calculated at 1 GHz for the L_2-IMATCS algorithm and the full breast model of Fig. 3 for eight antennas, and SNR = 60 dB (top) and 30 dB (bottom). SNR is relative to the total received signal, which includes direct antenna contributions and skin reflections.

V. CONCLUSION

We have proposed a novel MWI approach based on the DBIM algorithm and an iterative technique based on sparsity and the IMAT algorithm. Our simulation results have demonstrated the advantages and potential of this method to enhance resolution in microwave medical imaging. In particular, the proposed L_2-IMATCS algorithm can be adjusted to promote sparsity as the number of DBIM iterations increases in order to reconstruct fine details in the image. We have also presented a theoretical study of the convergence properties of this algorithm in the Appendix.

Our future work will focus on studying this approach for 3-D microwave medical imaging. This will involve computationally efficient implementations of the product $\mathbf{A}^*\mathbf{A}$ in blocks [5], and possibly the use of basis functions [3], [27] in (4) that can reduce computational complexity. These techniques can ensure that inversion is not prohibitively time consuming, while recently proposed analytic approaches [28] could replace the 3-D FDTD forward solvers thereby reducing dramatically the overall computational cost of the 3-D algorithm.

Beyond the 3-D extension, future work will focus on developing adaptive and robust strategies for the optimal choice of parameters in the L_2-IMATCS algorithm described in Section III-B, taking into account our multiple-frequency reconstruction strategy. For example, using our pre-selected values for the IMATCS parameters to reconstruct the two targets in Fig. 2(a) with low (10%) dielectric contrast relative to the homogeneous background resulted in less accurate reconstructions. This is due to the fact that the choice of the IMATCS parameters depends on the scattered signal and thus on the unknown contrast. Therefore, an automated and sound process for the selection of these parameters relative to the choice of λ_2 would increase the algorithm's robustness.

This preliminary study suggests that this novel algorithm can offer an advantageous sparsity-based approach to enhance resolution in microwave medical imaging. Our ultimate goal is to apply this algorithm to experimental systems and test whether it can lead to superior reconstructions relative to conventional L_2-based strategies.

APPENDIX
THE L_2-IHT ALGORITHM DERIVATION

The cost function to be minimized in the L_2-IHT algorithm is as follows:

$$f(x) = \|\mathbf{y} - \mathbf{A}x\|_2^2 + \lambda_1\|\mathbf{x}\|_0 + \lambda_2\|\mathbf{x}\|_2^2. \quad (16)$$

We solve this optimization problem using the Majorization Minimization scheme [8] which minimizes the surrogate of the cost function instead of its direct optimization. The surrogate function can be defined as (17) with the aid of an auxiliary variable $\mathbf{z}$

$$C(\mathbf{x}, \mathbf{z}) = \|\mathbf{y} - \mathbf{A}x\|_2^2 + \lambda_1\|\mathbf{x}\|_0 + \lambda_2\|\mathbf{x}\|_2^2 - \|\mathbf{A}x - \mathbf{A}z\|_2^2 \\ + \|\mathbf{x} - \mathbf{z}\|_2^2. \quad (17)$$

Neglecting the terms independent of $\mathbf{x}$ and decomposing the surrogate function in terms of its vector entries, we have

$$C(\mathbf{x}, \mathbf{z}) = \sum_i \mathbf{x}_i^2 - 2\mathbf{x}_i\mathbf{z}_i - 2\mathbf{x}_i\mathbf{A}_i^*\mathbf{y} + 2\mathbf{x}_i\mathbf{A}_i^*\mathbf{A}\mathbf{z} \quad (18)$$

$$+ \lambda_1|\mathbf{x}_i|_0 + \lambda_2\mathbf{x}_i^2. \quad (19)$$

If $x_i = 0$, the component $C(x_i, z_i)$ would be zero. For $x_i \neq 0$, the component $C(x_i, z_i)$ can be minimized by taking its derivative with respect to the entries, x_i. Thus, we have

$$\begin{cases} \mathbf{x}_i = 0 \rightarrow C(\mathbf{x}_i, \mathbf{z}_i) = 0 \\ \mathbf{x}_i \neq 0 \rightarrow C(\mathbf{x}_i, \mathbf{z}_i) = \lambda_1 - \frac{1}{1+\lambda_2}\left(\mathbf{z}_i + \mathbf{A}_i^*(\mathbf{y} - \mathbf{A}\mathbf{z})\right)^2 \end{cases}. \quad (20)$$

The crossing of the two curves occur in

$$|\mathbf{x}_i| = \sqrt{\frac{\lambda_1}{(1 + \lambda_2)}}. \quad (21)$$

Therefore, the iterative relation of L_2-IHT can be obtained as follows:

$$\mathbf{x}^{n+1} = \frac{1}{1 + \lambda_2}H_{\sqrt{\lambda_1/(1+\lambda_2)}}\left(\mathbf{x}^n + \mathbf{A}^*(\mathbf{y} - \mathbf{A}\mathbf{x}^n)\right) \quad (22)$$

where H is the hard thresholding operator defined as

$$H_\theta(x) = \begin{cases} 0 & \text{if } |x| \leq \theta \\ x & \text{if } |x| \geq \theta \end{cases}. \quad (23)$$

Lemma 1 (Nonincreasing): Assume $\|\mathbf{A}\|_2 \leq 1$ and let $\mathbf{x}^{n+1} = (1/(1+\lambda_2))H_{\sqrt{\lambda_1/(1+\lambda_2)}}(\mathbf{x}^n + \mathbf{A}^*(\mathbf{y} - \mathbf{A}\mathbf{x}^n))$, then the sequences $f(\mathbf{x}^n)$ and $C(\mathbf{x}^{n+1}, \mathbf{x}^n)$ are nonincreasing.
Proof 1:

$$f(\mathbf{x}^{n+1}) \leq f(\mathbf{x}^{n+1}) + \|\mathbf{x}^{n+1} - \mathbf{x}^n\|_2^2 - \|\mathbf{A}(\mathbf{x}^{n+1} - \mathbf{x}^n)\|_2^2 \\ = C(\mathbf{x}^{n+1}, \mathbf{x}^n) \leq C(\mathbf{x}^n, \mathbf{x}^n) = f(\mathbf{x}^n) \\ \leq f(\mathbf{x}^n) + \|\mathbf{x}^n - \mathbf{x}^{n-1}\|_2^2 - \|\mathbf{A}(\mathbf{x}^n - \mathbf{x}^{n-1})\|_2^2 \\ = C(\mathbf{x}^n, \mathbf{x}^{n-1})$$

where the first inequality results from the condition $\|\mathbf{A}\|_2 < 1$ and the first equality is due to the definition of the surrogate

function. The second inequality is a consequence of the fact that $\mathbf{x}^{n+1}$ is a minimizer of $C(\mathbf{x}, \mathbf{x}^n)$[8].

The iterative algorithm may be entrapped in the fixed points. Therefore, we investigate the fixed points of the recursive relation here. Let $\mathbf{x}^*$ be the fixed point of (22); in other words

$$\mathbf{x}_i^* = \frac{1}{1 + \lambda_2} H_{\sqrt{\lambda_1/(1+\lambda_2)}} \left(\mathbf{x}_i^* + \mathbf{A}_i^*(\mathbf{y} - \mathbf{A}\mathbf{x}^*) \right). \quad (24)$$

According to (23), the fixed points of the relation can be zero or nonzero depending on the absolute value of $\mathbf{x}_i^*$. We investigate the fixed points for both cases. If $\mathbf{x}_i^* = 0$, we would have

$$|\mathbf{A}_i^*(\mathbf{y} - \mathbf{A}\mathbf{x}^*)| \leq \sqrt{\frac{\lambda_1}{(1 + \lambda_2)}}. \quad (25)$$

For the case of $\mathbf{x}_i^* \neq 0$, we would have

$$\mathbf{x}_i^* = (\mathbf{x}_i^* + \mathbf{A}_i^*(\mathbf{y} - \mathbf{A}\mathbf{x}^*)) \frac{1}{1 + \lambda_2} \quad (26)$$

which results in

$$\mathbf{A}_i^*(\mathbf{y} - \mathbf{A}\mathbf{x}^*) = \mathbf{x}_i^* \lambda_2. \quad (27)$$

According to (23), $\mathbf{x}_i^* \neq 0$ implies that

$$|(\mathbf{x}_i^* + \mathbf{A}_i^*(\mathbf{y} - \mathbf{A}\mathbf{x}^*))| \geq \sqrt{\frac{\lambda_1}{(1 + \lambda_2)}}. \quad (28)$$

Combining (27) and (28), we have

$$|\mathbf{x}_i^*| \geq \frac{\sqrt{\frac{\lambda_1}{(1+\lambda_2)}}}{1 + \lambda_2} \quad (29)$$

and

$$|\mathbf{A}_i^*(\mathbf{y} - \mathbf{A}\mathbf{x}^*)| \geq \frac{\lambda_2 \sqrt{\frac{\lambda_1}{(1+\lambda_2)}}}{1 + \lambda_2}. \quad (30)$$

Thus, the fixed points of (22) satisfy the following:

$$\begin{cases} \mathbf{x}_i^* = 0 \rightarrow |\mathbf{A}_i^*(\mathbf{y} - \mathbf{A}\mathbf{x}^*)| \leq \sqrt{\frac{\lambda_1}{(1+\lambda_2)}} \\ |\mathbf{x}_i^*| \geq \frac{\sqrt{\frac{\lambda_1}{(1+\lambda_2)}}}{1+\lambda_2} \\ \rightarrow |\mathbf{A}_i^*(\mathbf{y} - \mathbf{A}\mathbf{x}^*)| \geq \frac{\lambda_2 \sqrt{\frac{\lambda_1}{(1+\lambda_2)}}}{1+\lambda_2} \end{cases} . \quad (31)$$

It should be proved that the fixed points of (22) are its local minima. Thus, being entrapped in a fixed point implies being in a local minimum. The corresponding proof is given in the following lemma.

Theorem 2 (Fixed Points-Local Minima): The fixed points of (22) are the local minima of (16).

Proof 3: we should prove that

$$f(\mathbf{x}^* + \partial\mathbf{h}) - f(\mathbf{x}^*) \geq 0 \quad \forall |\partial\mathbf{h}| \leq \varepsilon. \quad (32)$$

According to (16), we have

$$\begin{aligned} f(\mathbf{x}^* + &\partial\mathbf{h}) - f(\mathbf{x}^*) \\ &\geq \sum_i \lambda_1(|\mathbf{x}_i^* + \partial\mathbf{h}_i|_0 - |\mathbf{x}_i^*|_0) + \lambda_2(\partial\mathbf{h}_i)^2 + 2\lambda_2\mathbf{x}_i^*\partial\mathbf{h}_i \\ &+ 2\partial\mathbf{h}_i\mathbf{A}_i^*(\mathbf{A}\mathbf{x}^* - \mathbf{y}). \end{aligned} \quad (33)$$

We prove (32) for the two cases of (31), separately. First, we investigate the case where $\mathbf{x}_i^* = 0$. Then, if $\partial\mathbf{h}_i = 0$, the term is nonnegative. When $\partial\mathbf{h}_i$ becomes nonzero, we would have

$$\begin{aligned} \lambda_1(|\partial\mathbf{h}_i|_0) &+ \lambda_2(\partial\mathbf{h}_i)^2 + 2\partial\mathbf{h}_i\mathbf{A}_i^*(\mathbf{A}\mathbf{x}^* - \mathbf{y}) \\ &\geq \lambda_1(|\partial\mathbf{h}_i|_0) + \lambda_2(\partial\mathbf{h}_i)^2 \\ &+ \partial\mathbf{h}_i \left(-2\sqrt{\frac{\lambda_1}{(1 + \lambda_2)}} \right) \\ &= \left(\lambda_1 - 2\sqrt{\frac{\lambda_1}{(1 + \lambda_2)}}\partial\mathbf{h}_i \right) + \lambda_2(\partial\mathbf{h}_i)^2. \end{aligned} \quad (34)$$

Selecting $|\partial\mathbf{h}_i|$ as follows, we can have the above term greater than zero:

$$|\partial\mathbf{h}_i| \leq \frac{\sqrt{\lambda_1(1 + \lambda_2)}}{2}. \quad (35)$$

The second case is where

$$|\mathbf{x}_i^*| \geq \frac{\sqrt{\frac{\lambda_1}{(1+\lambda_2)}}}{1 + \lambda_2}. \quad (36)$$

If

$$|\partial\mathbf{h}_i| \leq \frac{\sqrt{\frac{\lambda_1}{(1+\lambda_2)}}}{1 + \lambda_2} \quad (37)$$

we can have $|\partial\mathbf{h}_i| \leq |\mathbf{x}_i|$ which results in $|\mathbf{x}_i^* + \partial\mathbf{h}_i|_0 = 1$. Thus, (34) translates to

$$\begin{aligned} \lambda_1(|\mathbf{x}_i^* + \partial\mathbf{h}_i|_0 - 1) &+ \lambda_2(\partial\mathbf{h}_i)^2 + 2\lambda_2\mathbf{x}_i^*\partial\mathbf{h}_i \\ &+ 2\partial\mathbf{h}_i\mathbf{A}_i^*(\mathbf{A}\mathbf{x}^* - \mathbf{y}) \\ &= \lambda_2(\partial\mathbf{h}_i)^2 \geq 0 \end{aligned} \quad (38)$$

where the equality is due to (27). Taking ε as $\min\left(\sqrt{\lambda_1(1 + \lambda_2)}/2, \sqrt{\lambda_1/(1 + \lambda_2)}/(1 + \lambda_2) \right)$, both (35) and (37) would be satisfied. The lemma is thus proved.

The next step is to prove the convergence of the L_2-IHT algorithm.

Theorem 4 (Convergence): If $f(\mathbf{x}^0) < \infty$ and if $\|\mathbf{A}\|_2 < 1$, then the sequence $\mathbf{x}^n$ defined by the iterative procedure in (22) converges to a local minimum of (16).

Before proving the theorem, we present the following lemma.

Lemma 5: $\forall\varepsilon > 0, \exists N$ such that

$$\forall n > N, \|\mathbf{x}^{n+1} - \mathbf{x}^n\|_2 \leq \varepsilon.$$

Proof 6: We that should prove that $\sum_{n=0}^N \|\mathbf{x}^{n+1} - \mathbf{x}^n\|_2^2$ converges by proving that it is bounded and monotonically increasing. The monotinicity is proved trivially and the boundedness results from

$$\sum_{n=0}^N \|\mathbf{x}^{n+1} - \mathbf{x}^n\|_2^2 \quad (39)$$

$$\leq \frac{1}{C} \sum_{n=0}^N \left(\|\mathbf{x}^{n+1} - \mathbf{x}^n\|_2^2 - \|\mathbf{A}(\mathbf{x}^{n+1} - \mathbf{x}^n)\|_2^2 \right) \quad (40)$$

$$\leq \frac{1}{C} \sum_{n=0}^N \left(f(\mathbf{x}^n) - f(\mathbf{x}^{n+1}) \right) \quad (41)$$

$$= \frac{1}{C} \left(f(\mathbf{x}^0) - f(\mathbf{x}^{n+1}) \right) \leq \frac{1}{C} \left(f(\mathbf{x}^0) \right) \quad (42)$$

where C is a lower bound on the spectrum of $(\mathbf{I} - \mathbf{A}^*\mathbf{A})$ and the second inequality results from the proof of Lemma 1[8].

Proof 7 (Proof of Convergence Theorem): In lemma 5, take $\varepsilon < \sqrt{\lambda_1/(1+\lambda_2)}/(1+\lambda_2)$. If $|\mathbf{x}_i^n| \geq \sqrt{\lambda_1/(1+\lambda_2)}/(1+\lambda_2)$ and $|\mathbf{x}_i^{n+1}| = 0$, then, $\|\mathbf{x}^{n+1} - \mathbf{x}^n\| \geq \sqrt{\lambda_1/(1+\lambda_2)}/(1+\lambda_2)$ which contradicts lemma 5. Thus, $|\mathbf{x}_i^m| \geq \sqrt{\lambda_1/(1+\lambda_2)}/(1+\lambda_2) \quad \forall m > n$ and the sequence $\mathbf{x}^n$ reduces to the Landweber algorithm with guaranteed convergence [29]. Similarly if $\mathbf{x}_i^n = 0$, then $\mathbf{x}_i^m = 0 \quad \forall m > n$ and the similar convergence guarantee can be given for this sequence. Therefore, the convergence of the algorithm in (22) is proved.

REFERENCES

[1] S. Semenov, "Microwave tomography: Review of the progress towards clinical applications," *Phil. Trans. R. Soc. A*, vol. 367, pp. 3021–3042, 2009.

[2] C. Gilmore, A. Abubakar, W. Hu, T. M. Habashy, and P. M. van den Berg, "Microwave biomedical data inversion using the finite-difference contrast source inversion method," *IEEE Trans. Antennas Propag.*, vol. 57, no. 5, pp. 1528–1538, May 2009.

[3] R. Scapaticci, I. Catapano, and L. Crocco, "Wavelet-based adaptive multiresolution inversion for quantitative microwave imaging of breast tissues," *IEEE Trans. Antennas Propag.*, vol. 60, no. 8, pp. 3717–3726, Aug. 2012.

[4] P. M. Meaney, M. W. Fanning, T. Raynolds, C. J. Fox, Q. Q. Fang, C. A. Kogel, S. P. Poplack, and K. D. Paulsen, "Initial clinical experience with microwave breast imaging in women with normal mammography," *Acad. Radiol.*, vol. 14, no. 2, pp. 207–218, Feb. 2007.

[5] J. D. Shea, P. Kosmas, S. C. Hagness, and B. D. Van Veen, "Three-dimensional microwave imaging of realistic numerical breast phantoms via a multiple-frequency inverse scattering technique," *Med. Phys.*, vol. 37, no. 8, pp. 4210–4226, Aug. 2010.

[6] P. Kosmas, J. D. Shea, B. D. Van Veen, and S. C. Hagness, "Three-dimensional microwave imaging of realistic breast phantoms via an inexact Gauss-Newton algorithm," in *Proc. IEEE Int. Symp. Antennas Propag.*, Jul. 2008, pp. 1–4.

[7] M. Azghani and F. Marvasti, "Iterative algorithms for random sampling and compressed sensing recovery," in *Int. Workshop Sampling Theory Appl.*, Bremen, Germany, Jul. 2013.

[8] T. Blumensath and M. E. Davies, "Iterative thresholding for sparse approximations," *J. Fourier Anal. Appl.*, vol. 14, no. 5-6, pp. 629–654, 2008.

[9] T. Blumensath and M. E. Davies, "Iterative hard thresholding for compressed sensing," *App. Computat. Harm. Analysis*, vol. 27, no. 3, pp. 265–274, 2009.

[10] F. Gao, B. Van Veen, and S. C. Hagness, "Contrast enhanced microwave imaging of breast tumors using sparsity regularization," in *Proc. IEEE Antennas Propag. Soc. Int. Symp. (APS-URSI)*, Chicago, IL, Jul. 2012, pp. 8–14.

[11] D. Winters, B. Van Veen, and S. C. Hagness, "A sparsity regularization approach to the electromagnetic inverse scattering problem," *IEEE Trans. Antennas Propag.*, vol. 58, no. 1, pp. 145–154, Jan. 2012.

[12] E. Candes and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 21–30, Mar. 2008.

[13] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[14] F. Marvasti, A. Amini, F. Haddadi, M. Soltanolkotabi, B. H. Khalaj, A. Aldroubi, S. Holm, S. Sanei, and J. Chambers, "A unified approach to sparse signal processing," *EURASIP J. Adv. Signal Process.*, 2012.

[15] P. C. Hansen, *Rank-Deficient and Discrete Ill-Posed Problems: Numerical Aspects of Linear Inversion*. Philadelphia: SIAM, 1998.

[16] N. K. Nikolova, "Microwave imaging for breast cancer," *IEEE Microw. Mag.*, vol. 12, no. 7, pp. 78–94, 2011.

[17] M. Azghani, P. Kosmas, and F. Marvasti, "Microwave imaging based on compressed sensing using adaptive thresholding," in *Proc. Eur. Conf. Antennas Propag.*, 2013.

[18] W. C. Chew, *Waves and Fields in Inhomogeneous Media*. New York: Van Nostrand Reinhold, 1990.

[19] J. Nocedal and S. J. Wright, *Numerical Optimization*, 2nd ed. New York: Springer, 2006.

[20] T. Rubæk, P. M. Meaney, P. Meincke, and K. D. Paulsen, "Nonlinear microwave imaging for breast-cancer screening using Gauss-Newton's method and the CGLS inversion algorithm," *IEEE Trans. Antennas Propag.*, vol. 54, no. 11, pp. 3517–3528, Nov. 2006.

[21] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM J. Scientif. Comput.*, vol. 20, no. 1, pp. 33–61, 1998.

[22] R. Tibshirani, "Regression shrinkage and selection via the LASSO," *J. Royal. Statist. Soc. B.*, vol. 58, no. 1, pp. 267–288.

[23] J. A. Tropp, A. C. Gilbert, and M. J. Strauss, "Algorithms for simultaneous sparse approximation: Part I: Greedy pursuit," *Signal Process.*, vol. 86, no. 3, pp. 572–588, 2006.

[24] T. Blumensath and M. E. Davies, "Gradient pursuits," *IEEE Trans. Signal Process.*, vol. 56, no. 6, pp. 2370–2382, 2008.

[25] J. D. Shea, P. Kosmas, B. Van Veen, and S. Hagness, "Contrast enhanced microwave imaging of breast tumors: A computational study using 3D realistic numerical phantoms," *Inverse Prob.*, vol. 26, p. 074009, Jul. 2010.

[26] E. Zastrow, S. K. Davis, M. Lazebnik, F. Kelcz, B. D. Van Veen, and S. C. Hagness, "Development of anatomically realistic numerical breast phantoms with accurate dielectric properties for modeling microwave interactions with the human breast," *IEEE Trans. Biomed. Eng.*, vol. 55, no. 12, pp. 2792–2800, Dec. 2008.

[27] D. Winters, J. D. Shea, P. Kosmas, B. Van Veen, and S. C. Hagness, "Three-dimensional microwave breast imaging: Dispersive dielectric properties estimation using patient-specific basis functions," *IEEE Trans. Med. Imag.*, vol. 28, no. 7, pp. 969–981, 2009.

[28] T. M. Grzegorczyk, P. M. Meaney, P. A. Kaufman, R. M. di Florio-Alexander, and K. D. Paulsen, "Fast 3-D tomographic microwave imaging for breast cancer detection," *IEEE Trans. Med. Imag.*, vol. 31, no. 8, pp. 1584–1592, Aug. 2012.

[29] L. Landweber, "An iteration formula for Fredholm integral equations of the first kind," *Amer. J. Math.*, vol. 73, pp. 615–624, 1951.