Reconstructing Dispersive Scatterers With Minimal Frequency Data

Yaswanth Kalepu[ID], *Student Member, IEEE*, Yash Sanghvi[ID], and Uday K. Khankhoje[ID], *Senior Member, IEEE*

Abstract—Reconstructing the permittivity of dispersive scatterers from the measurements of scattered electromagnetic fields is a challenging problem due to the nonlinearity of the associated optimization problem. Traditionally, this has been addressed by collecting scattered field data at multiple frequencies and using lower frequency reconstructions as *a priori* information for higher frequency reconstructions. By modeling the object dispersion as a Debye medium, we propose an inversion technique that recovers the object permittivity with a minimum number of frequencies. We compare the performance of this method with our recently developed deep learning based technique (Sanghvi. *et al.*, *IEEE Trans. Comp. Imag.*, 2019) and show that given a properly trained neural network, single frequency reconstructions can be very competitive with multifrequency techniques that do not use neural networks. We quantify this performance via extensive numerical examples and comment on the hardware implications of both approaches.

Index Terms—Deep learning, dispersion, inverse scattering.

I. INTRODUCTION

WE PRESENT solutions to the problem of determining the electrical permittivity of dispersive objects by studying the electromagnetic fields scattered by them. It is well known that such problems, termed inverse scattering problems (ISPs), are nonlinear in nature [1], [2], and that nonlinearity plays a larger role as the object size, permittivity, or frequency increases.

Solutions to overcome the nonlinearity of the ISP have been extensively studied in the literature; these include iterative and algebraic schemes [1], [3]–[8], and neural network-based schemes in more recent times [9], [10] (see [11] for a review). Since ISPs are of interest in many imaging problems such as breast cancer detection [12], it is desirable to solve the problem at small wavelengths so as to obtain high-resolution reconstructions. This presents a challenge, since the nonlinearity of the ISP gets worse as the frequency is increased. One of the techniques to overcome this is to collect scattering data at multiple frequencies, and to use the lower

frequency reconstructions as *a priori* information for higher frequency results [2], [13], [14], a technique called "frequency hopping" (FH). In the absence of such information, reconstruction algorithms can often get stuck in local minima [15]. Realistic materials are dispersive in nature; as a result, it is crucial to model the dispersion into the reconstruction algorithm. In recent work [16], [17], a single pole Debye model was incorporated into the distorted Born iterative method (DBIM) for reconstructions using multiple frequencies. However, the DBIM has a large computational cost since it requires repeated solutions of the forward problem.

Our contributions are summarized as follows.

1) We present a multi frequency dispersive (MFD) subspace optimization method (SOM) to handle the problem of reconstructing dispersive media within the framework of the popular SOM [6]. The latter is chosen due to its computational advantage over the DBIM. To the best of our knowledge, this is the first such combination in the literature.

2) We compare the proposed MFD-SOM method with our recently proposed deep learning enabled SOM [10]. We show that given a properly trained neural network, this method outperforms other methods while requiring only single frequency data.

The hardware and data acquisition implications of using both methods are discussed. To form a suitable baseline for these comparisons, we use the FH method that ignores material dispersion.

The letter is organized as follows. In Section II we detail the numerical methods considered in this work, including FH, the proposed MFD-SOM, and the deep learning enabled SOM. In Section III, we present numerical results of reconstructions of various dispersive objects. Finally, we conclude in Section IV with a discussion on the various features of each approach.

II. METHODS

A. Problem Setup

The well-known electric field integral equation is used to formulate the ISP [18] for a 2-D transverse magnetic [(TM), i.e., E_z] polarization scattering problem

$$E(\vec{r}) - E^i(\vec{r}) = E^s(\vec{r}) = k^2 \iint_D G(\vec{r}, \vec{r}') \chi(\vec{r}') E(\vec{r}') d\vec{r}' \quad (1)$$

where E, E^i, and E^s are the total, incident, and scattered fields, respectively, k is the wavevector for a medium with Green's function G, and $\chi(\vec{r}) = \epsilon_r(\vec{r}) - 1$ denotes the object contrast and $\epsilon_r(\vec{r})$ is the complex relative permittivity inside the imaging domain D.

The scattered field is measured at M receiver locations with N_I transmitters illuminating one at a time with a known incident field. We numerically solve (1) by discretizing the D into an N square pixels using the method of moments [18]. A "contrast source" variable, w, is defined as the product of the contrast, x, and the internal field, d, at any point in D. The resulting discretized equations for illumination by the pth transmitter operating at frequency f_c are

$$s_p^{(c)} = G_S^{(c)} w_p^{(c)} \quad \text{(Data Equation)} \tag{2a}$$

$$d_p^{(c)} = e_p^{(c)} + G_D^{(c)} w_p^{(c)} \quad \text{(Object Equation)} \tag{2b}$$

$$w_p^{(c)} = X^{(c)}\big(e_p^{(c)} + G_D^{(c)} w_p^{(c)}\big) \tag{2c}$$

where $e_p^{(c)}, d_p^{(c)} \in \mathbb{C}^N$ are the incident and total fields in the domain of interest, respectively, and $s_p^{(c)} \in \mathbb{C}^M$ is the corresponding scattered field measured at the receivers. $G_S^{(c)} \in \mathbb{C}^{M \times N}$ and $G_D^{(c)} \in \mathbb{C}^{N \times N}$ are the matrices that result after discretization and integration of Green's function, $x^{(c)} \in \mathbb{C}^N$ is the contrast at frequency f_c and $X^{(c)} \in \mathbb{C}^{N \times N} = \text{diag}(x^{(c)})$ is its diagonal form.

The frequency dependence of contrast for a dispersive medium is given using single-pole Debye model [16]

$$\begin{aligned} x^{(c)} &= \epsilon_\infty + \frac{\Delta\epsilon}{1 + j2\pi f_c \tau} + \frac{\sigma_s}{j2\pi f_c \epsilon_0} - 1 \\ &= \epsilon_\infty - 1 + \Delta\epsilon\, g_{(c)} + \sigma_s h_{(c)} \end{aligned} \tag{3}$$

where $\epsilon_\infty, \Delta\epsilon, \sigma_s \in \mathbb{R}^N$ and $\tau \in \mathbb{R}$ are the Debye parameters.

B. Summary of the Twofold SOM

The numerical methods we present are based on a variant of the SOM, called the twofold SOM (T-SOM) [19]. The SOM works by splitting the contrast-source variable into two orthogonal vector spaces, termed the signal and noise space, respectively. This kind of split into the signal (w^s) and noise space (w^n) component is based on the spectrum of the scattering operator, G_S, i.e., with the singular value decomposition (SVD) of G_S ($= U_S \Sigma_S V_S$), we express the contrast course as

$$w = w^s + w^n. \tag{4}$$

The split could be determined heuristically [6] or based on the knowledge of the variance of noise by the Morozov principle [10]. In the T-SOM [19], an additional regularization is imposed on the contrast-source noise space components by projecting them on only the top M_0 right singular vectors from the SVD of G_D ($= U_D \Sigma_D V_D$). The algorithm determines the signal space components, w^s, from the data and then iteratively updates x and w^n until convergence is reached.

Our final objective is to estimate the contrast x at the highest frequency given measurements for various illuminations and frequencies. We now outline the three numerical methods used in the remainder of the text.

C. FH Using SOM

The baseline method which we use to compare other methods is the so-called FH that was introduced earlier in

the literature [13]. We implement the FH method using the standard T-SOM method, with an additional total variation (TV) regularization [10].

Here, the standard T-SOM cost function is minimized at each frequency [10], [19] starting from the lowest, and the solution from a lower frequency is used as the initialization point for the optimization at the higher frequency (treating the contrast as frequency independent). Some variants of the above scheme have also been discussed [13], [14], where the imaginary part of the permittivity is scaled with frequency in accordance with Ohm's law, i.e., (3) with $\Delta\epsilon = 0$. We note that as per the Debye model, depending on the relative strengths of $\Delta\epsilon$ and σ_s, it is possible that the dispersion characteristics can be quite different from Ohm's law, and contrast initializations must be done as per *a priori* information.

D. Multi Frequency Dispersive SOM

The obvious limitation of the FH method is the inability to model the medium dispersion. To compensate for this, FH requires the use of many intermediate frequencies. Instead, the Debye model, as indicated in (3), can be incorporated into the reconstruction algorithm. Previously, this was shown in conjunction with the DBIM method [16].

The central idea of the proposed method, termed the MFD SOM, is to estimate the Debye parameters by minimizing a cost function consisting of information at all frequencies (C in number) and illuminations at once as shown in (5). The benefit of the method is that reconstruction is performed with fewer frequencies. The composite cost function is

$$\begin{aligned} J\big(&\beta_1^{(1)}, \ldots, \beta_{N_I}^{(1)}, \ldots, \beta_1^{(C)}, \ldots, \beta_{N_I}^{(C)}, r\big) \\ &= \sum_{c=1}^{C} \sum_{p=1}^{N_I} \Bigg(\frac{\|s_p^{(c)} - G_S^{(c)}\big(w_p^{s(c)} + V_D^{'(c)}\beta_p^{(c)}\big)\|_2^2}{\|s_p^{(c)}\|_2^2} \\ &\quad + \frac{\|A^{(c)}\beta_p^{(c)} - b_p^{(c)}\|_2^2}{\|w_p^{s(c)}\|_2^2} \Bigg) + \gamma \, \Bigg\| \begin{bmatrix} D & 0 & 0 \\ 0 & D & 0 \\ 0 & 0 & D \end{bmatrix} r \Bigg\|_1 \end{aligned} \tag{5}$$

where

$$A^{(c)} := \big(I - X^{(c)} G_D^{(c)}\big) V_D^{'(c)}$$

$$b_p^{(c)} := X^{(c)} e_p^{(c)} - \big(I - X^{(c)} G_D^{(c)}\big) w_p^{s(c)}$$

$$V_D^{'(c)} := \big(I - V_{S+}^{(c)} V_{S+}^{(c)H}\big) V_{D+}^{(c)}$$

$$D = \big[D_v \ D_h\big]^T.$$

Here, $V_{S+}^{(c)}$ represents the basis used for representing the signal subspace of contrast-source ($w_p^{s(c)}$) at the pth illumination and cth frequency; $V_{D+}^{(c)}$ is the basis used for representing the noise space components with $\beta_i^{(c)}$ denoting the coefficient of the ith basis vector; D_v and D_h are the vertical and horizontal first-order difference matrices used to enforce TV regularization with an empirical factor, γ; and $r = \big[\epsilon_\infty \ \epsilon_\Delta \ \sigma_s\big]^T \in \mathbb{R}^{3N \times 1}$ represents the set of Debye parameters which need to be estimated.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

KALEPU *et al.*: RECONSTRUCTING DISPERSIVE SCATTERERS

3

The complete algorithm is outlined as follows:

1) Procedure for MFD SOM:

1) Solve for $x^{(1)}$ using T-SOM at the lowest frequency
2) Estimate $w_p^{s(c)}$ from $s_p^{(c)}$ $\forall c$ (Morozov's principle)
3) Estimate $w_p^{n(c)}$ using (2c) with $x^{(1)}$ at all frequencies and thereby obtain $\beta_p^{(c)}$
4) Optimize the cost function provided in (5) by iterating r and β updates using the alternating direction method of multipliers (ADMM) [20], and the conjugate gradient method, respectively (See Appendix for details of r-update).
5) Repeat step 4 till convergence.

We note that the above-described scheme is similar in essence to the so-called "hybrid" FH described in the earlier work [13].

E. Single Frequency SOM With Deep Learning

In this method, we collect the scattered fields at only the highest frequency, f_C, and estimate the contrast using our recently introduced method [10]. This method, termed as single frequency SOM with deep learning (SFDL), is a deep learning-based method, where we embed a convolution neural network in the conventional T-SOM [19].

In the conventional T-SOM approach, the contrast-source noise space components, $w_p^{n(c)}$, are initialized to zero in the reconstruction algorithm. Our key observation was that as the contrast of the object increases, most of the true solution's energy moves from the signal space to the noise space. Thus the choice of initialization of the noise space components at the origin becomes worse as the contrast increases, thereby raising the chances of the algorithm getting stuck in local minima. In our approach, the neural network is trained to learn the noise space components given the signal space components, thereby initializing the reconstruction algorithm closer to the true solution. We refer the reader to [10] and note that while the conventional T-SOM type methods were able to reconstruct standard benchmark objects (like the "Austria" profile) up to contrast 1.5, the SFDL was able to achieve reconstructions for very high contrasts up to 4.

We have performed a large number of numerical experiments on dispersive media. In the interest of clarity and conciseness, we distill the findings in the results of three distinct numerical experiments.

III. RESULTS

A. Computational Details

The computational domain is a square domain of side 2 m, illuminated by 16 transmitters/32 receivers placed uniformly on a circle of radius 6 and 4 m, respectively. The domain is discretized into 100×100 pixels for synthetic generation of scattered field measurements (corrupted with Gaussian distributed noise with a signal to noise ratio of 25 dB). For the inverse problem, the domain is discretized into 50×50 pixels. A standard "Austria" profile object (see [21] for details) is used as a benchmark object for numerical simulations. The lowest and highest frequencies used for FH and MFD are 100 and 400 MHz, respectively. Unless otherwise specified,

reconstructions are shown at 400 MHz. The standard quantitative metrics used for reconstruction are the total error, $T = (1/N) \sum_{i \in [1,N]} (|x_{t,i} - x_{r,i}|)/(|1 + x_{t,i}|)$ and the internal error, $I = (1/|S|) \sum_{i \in S} (|x_{t,i} - x_{r,i}|)/(|1 + x_{t,i}|)$, where x_r, x_t are the reconstructed and true contrast profiles, respectively, and S is the support of the object. All simulations were performed on a 3.1 GHz Intel-Xeon 4-core processor. In the case of the SFDL, each pixel in each object of the training set was assigned a random contrast in the range $[0, 7]$ (real part) and $[-5, 0]$ (imaginary part). The training was done over 30 epochs and took 36 h. Each reconstruction took approximately 300 s to converge. In the case of the MFD approach, each reconstruction involving two frequencies took approximately 2500 s to converge. The computational complexity of the SFDL, FH-SOM, and MFD-SOM methods, apart from the constant $O(N^2)$ cost of an SVD [6], are approximately: $O(I_{it} N_I N \log N)$, $O(C I_{it} N_I N \log N)$, and $O(C I_{it} N_I N \log N)$, respectively, where I_{it} is the number of iterations; the latter varies across the three methods, however, they are comparable within the same complexity class.

B. Preliminary Results

In our recent work, we have shown that a nondispersive Austria profile object with Debye-parameters: $\epsilon_\infty = 5.0$, $\Delta \epsilon = 0$, and $\sigma_s = 0$ can be reconstructed using SFDL [10, Fig. 7(c)] (but not T-SOM). For successfully reconstructing the same using the FH method, we needed scattered field data at two frequencies, 100 and 400 MHz [10, Fig. 9(a)]. However, a simplistic assumption of a nondispersive medium was made. Noting that higher values of $\Delta \epsilon$ and σ_s leads to larger dispersion, we now show the reconstruction results using the previously described methods (keeping $\tau = 20 \times 10^{-11}$ s fixed).

C. First Numerical Experiment–FH Fails

In our first experiment (see Fig. 1), we consider an Austria profile object with Debye parameters of $\epsilon_\infty = 3.0$, $\Delta \epsilon = 0.5$, and $\sigma_s = 0.02$. It is observed that the FH could not reconstruct the profile as the reconstruction at the lowest frequency is not a very good initial guess for the higher frequency due to dispersion. But in the case of MFD, we can reconstruct the solution correctly because we solve for Debye parameters (instead of contrast) using both frequencies at a time. Note that the MFD has an advantage of describing the contrast with fewer unknowns than the FH [$3N$ (3 Debye parameters per pixel) instead of $4N$ (real and imaginary parts of contrast per pixel per frequency)]. It is also observed that the SFDL reconstruction is better than MFD because the neural network helps in starting from a better initial guess in this experiment.

D. Second Numerical Experiment—MFD Needs More Frequencies

In the second experiment (see Fig. 2), we use Austria profile with Debye parameters of $\epsilon_\infty = 2.0$, $\Delta \epsilon = 1$, and $\sigma_s = 0.02$. In this case, the scatterer is more dispersive compared to the first experiment (due to a larger $\Delta \epsilon$). It can be observed that

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

4   IEEE GEOSCIENCE AND REMOTE SENSING LETTERS
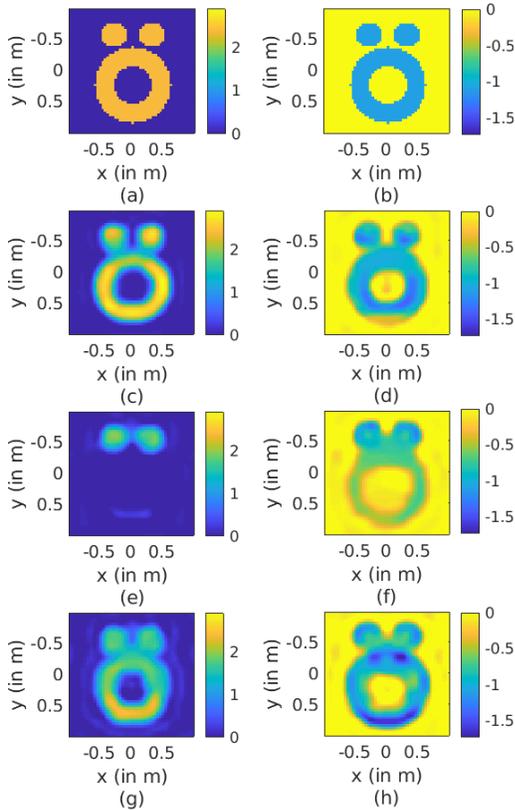

Fig. 1.   (a) Real and (b) Imaginary parts of the true scatterers. Real and Imaginary parts of reconstruction results using (c) and (d) SFDL, (e) and (f) FH, and (g) and (h) MFD. The respective internal and external errors are 16.2%, 59%, 23.6%, and 17.9%, 27.3%, 28.95%. The true Debye parameters are $\epsilon_\infty = 3.0$, $\Delta\epsilon = 0.5$, $\sigma_s = 0.02$. 100, and 400 MHz are used for FH, MFD reconstructions.
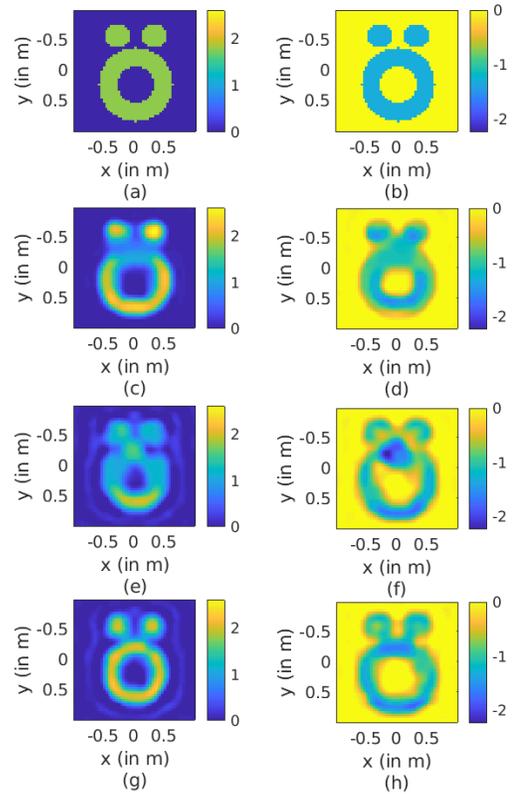

Fig. 2.   (a) Real and (b) Imaginary parts of the true scatterers. Real and Imaginary parts of reconstruction results using (c) and (d) SFDL, (e) and (f) MFD at 100, and 400 MHz, and (g) and (h) MFD at 100, 200, 300, and 400 MHz. The respective internal and external errors are 20.2%, 30.5%, 24.2%, and 17.8%, 30.3%, 24.6%. The true Debye parameters are $\epsilon_\infty = 2.0$, $\Delta\epsilon = 1.0$, and $\sigma_s = 0.02$.

SFDL reconstructed the object successfully, whereas the MFD failed at the task due to the increased dispersion. From our numerical experiments, we found that the only way in which the MFD could be made to reconstruct successfully was to give it scattered field data at more frequencies. In addition to the 100 and 400-MHz data, we also supply the algorithm with 200 and 300-MHz data, and as shown by the results in Fig. 2(g) and (h), the MFD now succeeds. We note that the quantitative reconstructions with SFDL at a single frequency are still better than the MFD with 4 frequencies.

E. Third Numerical Experiment—MFD Also Fails

In the third experiment (see Fig. 3), we consider an Austria profile object with Debye parameters of $\epsilon_\infty = 2.0$, $\Delta\epsilon = 0.5$, and $\sigma_s = 0.04$. The MFD failed to reconstruct the object even though data at 100, 200, 300, and 400 MHz was provided. We conjecture that the reason for this failure is that the skin depth at 400 MHz for this object is fairly low (0.24 m). Since the ring thickness is larger (0.30 m), some information about the interior of the object is not accessible at this frequency. It can be observed that the SFDL reconstruction is only qualitatively good, though better than MFD, aided by the fact that the neural network was trained on a data set containing lossy objects. In such situations, the recommended strategy will be to lower the highest frequency such that the entire

object can be interrogated by the incident field. However, the trade-off is that lowering the frequency also lowers the reconstruction resolution.

IV. DISCUSSION

Based on the various numerical simulations that have been reported in this study, we arrive at the following important conclusions that will inform future experiments in microwave inverse imaging.

1) A properly trained SFDL approach is superior in its reconstruction abilities as compared to multifrequency approaches. The obvious advantage from a hardware perspective is that the design of antennas becomes simpler, as only a single frequency operation needs to be optimized. The second advantage is that the data acquisition/signal processing also becomes simpler since data at multiple frequencies need not be recorded.
2) An advantage of the MFD approach is that the lengthy training times involved in the SFDL approach (detailed earlier) are avoided. A retraining of the neural network is required when the physical setup, e.g., operating frequency, estimated size of object, viewing geometry, etc., changes.

Thus, in more exploratory problems where the physical setup is yet to be optimized, it would be preferable to invest in

KALEPU *et al.*: RECONSTRUCTING DISPERSIVE SCATTERERS

5

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.
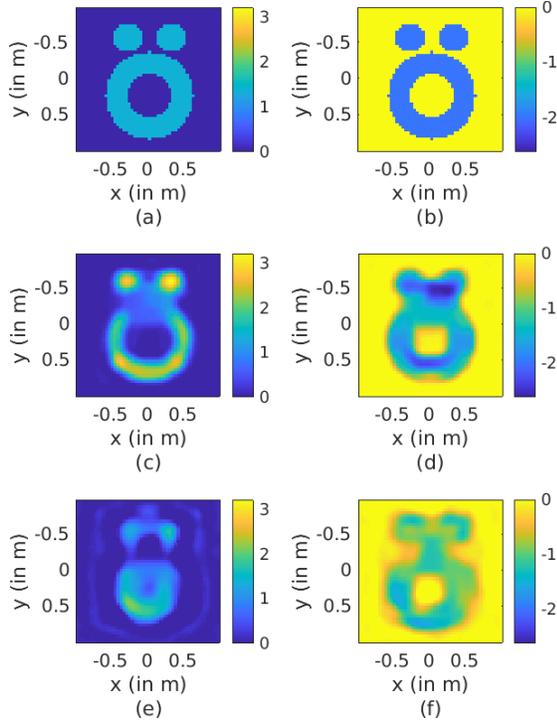

Fig. 3. (a) Real and (b) Imaginary part of the true scatterers. Real and Imaginary part of reconstruction results using (c) and (d) SFDL, (e) and (f) MFD at 100, 200, 300, and 400 MHz. The respective internal and external errors are 28.2%, 44.9%, and 23.2%, 36.1%. The true Debye parameters are $\epsilon_\infty = 2.0$, $\Delta\epsilon = 0.5$, and $\sigma_s = 0.04$.

an MFD approach. However, once the setup has been finalized, the benefits of the SFDL approach win over, as the training of the neural network is a one-time cost for a given setup. It is hoped that by future innovations, this training cost can be greatly reduced.

APPENDIX–UPDATE OF r

When (5) is rearranged by fixing β's, and by using (3) to covert contrast x to the r-variable, we get the following equation in terms of r:

$$\min_r CF = \sum_{n=1}^{N_I} \sum_{c=1}^{C} \frac{1}{2}||K_n^{(c)}r - b_n^{(c)}||_2^2 + \gamma \left\| \begin{bmatrix} D & 0 & 0 \\ 0 & D & 0 \\ 0 & 0 & D \end{bmatrix} r \right\|_1 \tag{6}$$

where

$$K_n^{(c)H} = \begin{bmatrix} \mathbb{D}\left(\text{Re}\left[\dfrac{d_n^{(c)}}{|w_n^{s(c)}|}\right]\right) & \mathbb{D}\left(\text{Im}\left[\dfrac{d_n^{(c)}}{|w_n^{s(c)}|}\right]\right) \\ \mathbb{D}\left(\text{Re}\left[\dfrac{d_n^{(c)}g_{(c)}}{|w_n^{s(c)}|}\right]\right) & \mathbb{D}\left(\text{Im}\left[\dfrac{d_n^{(c)}g_{(c)}}{|w_n^{s(c)}|}\right]\right) \\ \mathbb{D}\left(\text{Re}\left[\dfrac{d_n^{(c)}h_{s(c)}}{|w_n^{s(c)}|}\right]\right) & \mathbb{D}\left(\text{Im}\left[\dfrac{d_n^{(c)}h_{s(c)}}{|w_n^{s(c)}|}\right]\right) \end{bmatrix}$$

$$b_n^{(c)} = \begin{bmatrix} \text{Re}\left(\dfrac{w_n^{(c)} - d_n^{(c)}}{|w_n^{s(c)}|}\right) \\ \text{Im}\left(\dfrac{w_n^{(c)} - d_n^{(c)}}{|w_n^{s(c)}|}\right) \end{bmatrix}.$$

$\mathbb{D}$ converts a column vector to a diagonal matrix and the division is done element-wise. This equation is solved using the ADMM method [20].

REFERENCES

[1] W. Chew and Y. 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.

[2] W. Chew and J. Lin, "A frequency-hopping approach for microwave imaging of large inhomogeneous bodies," *IEEE Microw. Guid. Wave Lett.*, vol. 5, no. 12, pp. 439–441, 1995.

[3] P. M. Van Den Berg and R. E. Kleinman, "A contrast source inversion method," *Inverse Problems*, vol. 13, no. 6, pp. 1607–1620, Dec. 1997.

[4] T. M. Habashy and A. Abubakar, "A general framework for constraint minimization for the inversion of electromagnetic measurements," *Prog. Electromagn. Res.*, vol. 46, pp. 265–312, 2004.

[5] T. Rubek, 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. 55, no. 8, pp. 2320–2331, Aug. 2007.

[6] X. Chen, "Subspace-based optimization method for solving inverse-scattering problems," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 1, pp. 42–49, Jan. 2010.

[7] M. T. Bevacqua, L. Crocco, L. Di Donato, and T. Isernia, "An algebraic solution method for nonlinear inverse scattering," *IEEE Trans. Antennas Propag.*, vol. 63, no. 2, pp. 601–610, Feb. 2015.

[8] Y. Zhong, M. Lambert, D. Lesselier, and X. Chen, "A new integral equation method to solve highly nonlinear inverse scattering problems," *IEEE Trans. Antennas Propag.*, vol. 64, no. 5, pp. 1788–1799, May 2016.

[9] Z. Wei and X. Chen, "Deep-learning schemes for full-wave nonlinear inverse scattering problems," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 4, pp. 1849–1860, Apr. 2019.

[10] Y. Sanghvi, Y. N. G. B. Kalepu, and U. Khankhoje, "Embedding deep learning in inverse scattering problems," *IEEE Trans. Comput. Imag.*, early access. [Online]. Available: https: //github.com/sanghviyashiitb/EmbeddingDLinISP-Github, doi: 10.1109/TCI.2019.2915580.

[11] A. Massa, D. Marcantonio, X. Chen, M. Li, and M. Salucci, "DNNs as applied to electromagnetics, antennas, and propagation—A review," *IEEE Antennas Wireless Propag. Lett.*, vol. 18, no. 11, pp. 2225–2229, Nov. 2019.

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

[13] O. Bucci, L. Crocco, T. Isernia, and V. Pascazio, "Inverse scattering problems with multifrequency data: Reconstruction capabilities and solution strategies," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 4, pp. 1749–1756, Jul. 2000.

[14] Y. Liu et al., "A frequency-hopping subspace-based optimization method for reconstruction of 2-D large uniaxial anisotropic scatterers with TE illumination," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 10, pp. 6091–6099, Oct. 2016.

[15] K. Yaswanth and U. K. Khankhoje, "Two-dimensional nonlinear microwave imaging with total variation regularization," in *Proc. Progr. Electromagn. Res. Symp.-Fall (PIERS - FALL)*, Nov. 2017.

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

[17] Z. Miao and P. Kosmas, "Multiple-frequency DBIM-TwIST algorithm for microwave breast imaging," *IEEE Trans. Antennas Propag.*, vol. 65, no. 5, pp. 2507–2516, May 2017.

[18] J. Richmond, "Scattering by a dielectric cylinder of arbitrary cross section shape," *IEEE Trans. Antennas Propag.*, vol. AP-13, no. 3, pp. 334–341, May 1965.

[19] Y. Zhong and X. Chen, "Twofold subspace-based optimization method for solving inverse scattering problems," *Inverse Problems*, vol. 25, no. 8, Aug. 2009, Art. no. 085003.

[20] S. Boyd, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *FNT Mach. Learn.*, vol. 3, no. 1, pp. 1–122, 2010.

[21] K. Belkebir and A. Tijhuis, "Using multiple frequency information in the iterative solution of a two-dimensional nonlinear inverse problem," in *Proc. Progr. Electromagn. Res. Symp. (PIERS)*. Innsbruck, Austria: Univ. of Innsbruck, Jul. 1996, p. 353.