An Improved Attributed Scattering Model Optimized by Incremental Sparse Bayesian Learning

Zenghui Li, *Student Member, IEEE*, Kan Jin, Bin Xu, Wei Zhou, and Jian Yang, *Senior Member, IEEE*

***Abstract*—In this paper, we propose an improved attributed scattering model to mathematically unify the scattering models of several canonical primitives. These primitives include not only point- and line-segment-scatterers, such as trihedral, cylinder, dihedral, and rectangular plane, but also arc scatterers, such as sphere and top-hat. The estimation of the model parameters can be posed as an ill-posed linear inverse problem. To overcome the ill-posedness, we employ the incremental sparse Bayesian learning method to realize the sparsity-driven continuous parameter estimation. Inverse scattering experiments demonstrate that the proposed methodology not only provides desirable sparse representations of the target scattering response but is also able to capture richer geometrical information than the existing models.**

***Index Terms*—Attributed scattering model (ASM), improved attributed scattering model (IASM), incremental sparse Bayesian learning (ISBL), inverse scattering, sparse Bayesian learning (SBL).**

I. Introduction

FROM the geometrical theory of diffraction (GTD), the high-frequency scattering response from a complex object can be well approximated as a superposition of responses from a series of individual scatterers [1]. Parametric modeling for isolated radar scattering responses, developed from this theory, plays an essential role in scattering mechanism analysis and feature-based target discrimination and identification.

Many researchers have made great efforts in developing a parsimonious physics-based parametric model [2]–[5]. Potter *et al.* [4] proposed an attributed scattering center model for dominant scattering mechanisms, based on the assumption that the scattering responses of both localized and distributed objects originate from a set of discrete scattering centers. Halman *et al.* [6] represented the frequency dependence of the scattering center as a second-order polynomial to enhance the modeling of the mixed scattering mechanisms. Gerry *et al.* [5] disputed the accuracy of modeling the distributed scattering responses by point-scattering centers and developed a new parametric attributed scattering model (ASM) which characterized the frequency and aspect angle dependence for both localized and distributed scattering mechanisms. The parameter set of this model provides a physically relevant description of the scatterers, so that it can offer rich information about the object geometries. These parametric scattering models have been successfully applied in synthetic aperture radar (SAR) feature extraction and automatic target recognition [4], [7]–[13].

Model parameter estimation is an important problem that accompanies with parametric modeling because it is crucial for both improving the accuracy of extracted features and detailing the physical description of targets. To solve the problem of high-dimensional parameter optimization, several algorithms have been proposed, such as the approximate maximum likelihood algorithm [5], the modified RELAX algorithm [11], and the improved orthogonal matching pursuit algorithm [14], [15]. Essentially, the task of estimating the model parameters is equivalent to solving a linear inverse problem, i.e., employing the parametric model to linearly approximate the measured data. Thus, besides the aforementioned methods, some other sparse approximation methods [16] are also applicable since they can also overcome the ill-posedness property of inverse scattering problems and improve the solution quality. Among them, sparse Bayesian learning (SBL) [17]–[22] has recently attracted much attention in the research community for its advantages in parameter selection. In addition, incremental optimization strategy [23] further facilitates the application of this method since it can avoid overcomplete dictionary matrices and directly optimize continuous parameters [24]–[26].

In this paper, we propose an improved ASM (IASM) and apply the incremental SBL (ISBL) method to model optimization. There are two major innovative contributions in our work. On one hand, the developed IASM mathematically coincides with all of the canonical scattering models [27], [28], making the model versatile to point- and line-segment-scatters as well as arc ones. On the other hand, the ISBL method for real-valued atoms is adapted to complex-valued atoms. Based on this, continuous optimization with the proposed scattering model toward sparsity is successfully applied to the involved inverse scattering problems for the first time.

The rest of this paper is organized as follows. In Section II, we propose IASM and evaluate the consistency with canonical scattering models. In Section III, we introduce the SBL method and extend incremental optimization to the case of complex-valued atoms. In Section IV, we discuss the inverse scattering experiments under different conditions, including different geometrical shapes, polarizations, signal-to-noise ratios (SNRs), and scattering models. Finally, the conclusion is given in Section V.

Manuscript received December 1, 2014; revised June 12, 2015, October 7, 2015, and November 23, 2015; accepted December 14, 2015. This work was supported in part by the National Natural Science Foundation of China (NSFC) under Grant 41171317, by the Key Project of the NSFC under Grant 61132008, by the Major Research Plan of the NSFC under Grant 61490693, and by the Research Foundation of Tsinghua University.

The authors are with the Department of Electronic Engineering, Tsinghua University, Beijing 100084, China (e-mail: sunshinenudt@gmail.com; kevincabbage@sina.com; xubin07161@gmail.com; zhouw06@mails.tsinghua.edu.cn; yangjian_ee@mail.tsinghua.edu.cn).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TGRS.2015.2509539

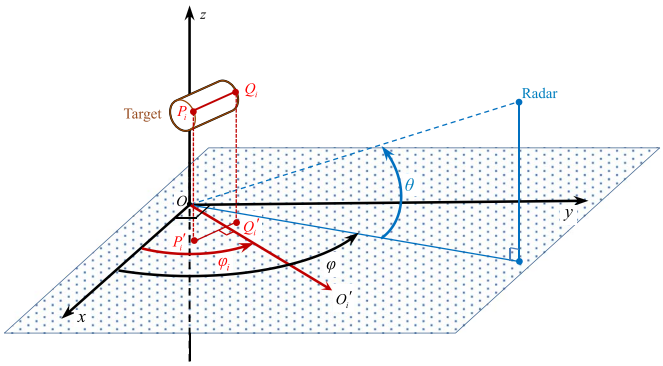


Fig. 1. Geometrical relationship between the radar and the target. The angle pair (φ, θ) describes the location of a radar in azimuth and elevation. The cylinder P_iQ_i, parallel to the image plane XOY, is a part of the target which is assumed to be located in the point O. The length and the aspect angle of the cylinder are L_i and φ_i, respectively.

II. Parametric Modeling of Scattering Response

From the GTD theory of the far-field pattern, the backscattered field $S(\kappa, \varphi)$ of a target can be approximated by a sum of M individual normalized scattering terms $\{\phi_i(\kappa, \varphi)\}$ $(i = 1, 2, \ldots, M)$ which are commonly modeled as the functions of the wavenumber κ and the aspect angle φ

$$S(\kappa, \varphi) = \sum_{i=1}^{M} A_i \phi_i(\kappa, \varphi) + \varepsilon(\kappa, \varphi) \tag{1}$$

where A_i is the amplitude of the ith scattering term and $\varepsilon(\kappa, \varphi)$ represents the residual error. To simplify the analysis and identify the structural features, an important consideration is to develop a unified and compact form for $\phi_i(\kappa, \varphi)$. This unified model should be able to mathematically approximate various localized and distributed scattering mechanisms.

A. ASM

For convenience, we rewrite the ith normalized ASM [5] as

$$\phi_i^{(\mathrm{ASM})}(\kappa, \varphi; x_i, y_i, \lambda_i, \gamma_i, L_i, \varphi_i) = \left(\frac{j\kappa}{\kappa_c}\right)^{\lambda_i} \exp(-\gamma_i \kappa \sin\varphi) \times \exp\{-j2\kappa(x_i \cos\varphi + y_i \sin\varphi)\} \operatorname{sinc}(L_i \Delta_i) \tag{2}$$

where

$$\Delta_i = \kappa \sin(\varphi - \varphi_i), \kappa = \frac{2\pi f}{c}, \kappa_c = \frac{2\pi f_c}{c} \tag{3}$$

f is the frequency, f_c is the center frequency, c is the propagation velocity, λ_i characterizes the associated frequency dependence, and γ_i has no direct physical interpretation. As shown in Fig. 1, (x_i, y_i) stands for the position of the ith scatterer as projected onto the imaging plane XOY, and L_i and φ_i are the length and the orientation of the ith distributed scatterer, respectively. Obviously, the model parameters provide physically relevant descriptions of targets, and these descriptions objectively reveal the geometrical structures. However, ASM can efficiently approximate point- and line-segment-scattering mechanisms rather than arc-scattering mechanisms. Moreover, its damped exponential function which models the amplitude dependence on aspect angle is not consistent with canonical scattering models in mathematical forms [27], [28]. As a result, the dependence factor has no direct physical interpretation [5]. These shortcomings motivate us to propose an improvement for ASM.

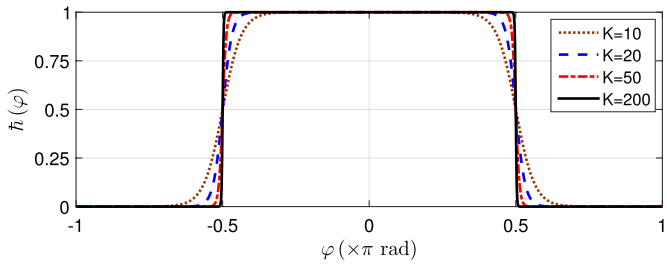


Fig. 2. Curve approximations for different steepness parameters.

B. IASM

We modify ASM to obtain higher precision and accommodate more sophisticated scattering behaviors. The developed model is named as IASM, written as

$$\begin{aligned}\phi_i^{(\mathrm{IASM})}&\left(\kappa, \varphi; x_i, y_i, r_i, \lambda_i, \tilde{\gamma}_i, L_i, \varphi_i, \varphi_i^{(1)}, \varphi_i^{(2)}\right) \\ &= \left(\frac{j\kappa}{\kappa_c}\right)^{\lambda_i} \cos^{\tilde{\gamma}_i}(\varphi - \varphi_i) \operatorname{sinc}(L_i \Delta_i) \\ &\quad \times \exp\{-j2\kappa(x_i \cos\varphi + y_i \sin\varphi - r_i)\}\, \hbar\left(\varphi; \varphi_i^{(1)}, \varphi_i^{(2)}\right)\end{aligned} \tag{4}$$

where the parameters $\kappa, \varphi, \lambda_i, L_i, \varphi_i, \Delta_i$ have the same physical meanings as those of ASM. By contrast, the exponential function $\exp(-\gamma_i \kappa \sin\varphi)$ in (2) is replaced by the cosine function $\cos^{\tilde{\gamma}_i}(\varphi - \varphi_i)$. The exponent $\tilde{\gamma}_i$ definitely characterizes the aspect dependence factor, which describes the amplitude dependence of the scattering response on aspect angle. In addition, referring to [28], the parameter r_i is defined to determine the migrating displacement of the scatterer, which describes the migration phenomenon of the scatterer. The rectangular window function $\hbar(\varphi; \varphi_i^{(1)}, \varphi_i^{(2)})$ is defined as

$$\hbar\left(\varphi; \varphi_i^{(1)}, \varphi_i^{(2)}\right) = \begin{cases} 1 & \varphi \in \left[\varphi_i^{(1)}, \varphi_i^{(2)}\right] \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

while a derivable approximation of this function can be given by the difference between two translated logistic functions $\lambda(\varphi - \varphi_i^{(1)}), \lambda(\varphi - \varphi_i^{(2)})$ [29]

$$\begin{aligned}\hbar\left(\varphi; \varphi_i^{(1)}, \varphi_i^{(2)}\right) &\approx \lambda\left(\varphi - \varphi_i^{(1)}\right) - \lambda\left(\varphi - \varphi_i^{(2)}\right) \\ &= \left\{1 + \exp\left[-K\left(\varphi - \varphi_i^{(1)}\right)\right]\right\}^{-1} \\ &\quad - \left\{1 + \exp\left[-K\left(\varphi - \varphi_i^{(2)}\right)\right]\right\}^{-1}\end{aligned} \tag{6}$$

where K affects the steepness of the curve (see Fig. 2). Therefore, it is usually chosen as a sufficiently large positive number, and in this study, $K = 200$.

$$S_{\mathrm{tri}} = A_{\mathrm{tri}}\phi^{(\mathrm{IASM})}\left(\kappa,\varphi;x_i,y_i,0,1,1,0,\varphi_{\mathrm{tri}}+\frac{\pi}{4},\varphi_{\mathrm{tri}}-\frac{\pi}{4},\varphi_{\mathrm{tri}}\right) + A_{\mathrm{tri}}\phi^{(\mathrm{IASM})}\left(\kappa,\varphi;x_i,y_i,0,1,1,0,\varphi_{\mathrm{tri}}-\frac{\pi}{4},\varphi_{\mathrm{tri}},\varphi_{\mathrm{tri}}+\frac{\pi}{4}\right) \tag{13}$$

$$S_{\mathrm{pla}} = A_{\mathrm{pla}}\phi^{(\mathrm{IASM})}\left(\kappa,\varphi;x_i,y_i,\frac{1}{2}H_{\mathrm{pla}}\sin\theta,0,0,L_{\mathrm{pla}}\cos\theta,\varphi_i,\varphi_i-\frac{\pi}{2},\varphi_i+\frac{\pi}{2}\right) - A_{\mathrm{pla}}\phi^{(\mathrm{IASM})}\left(\kappa,\varphi;x_i,y_i,-\frac{1}{2}H_{\mathrm{pla}}\sin\theta,0,0,L_{\mathrm{pla}}\cos\theta,\varphi_i,\varphi_i-\frac{\pi}{2},\varphi_i+\frac{\pi}{2}\right) \tag{14}$$

C. Relationship Between IASM and Each Canonical Scattering Model

For the sake of model evaluation, we use the proposed IASM to represent six canonical shapes: sphere, top-hat, cylinder, dihedral, trihedral, and rectangular plate. In particular, the physical correlations between geometrical parameters and model parameters are explicitly given. According to the 3-D and bistatic parametric models of these canonical scattering responses [28], the corresponding 2-D monostatic counterparts can be rewritten as functions of the wavenumber κ, the aspect angle φ, and the elevation angle θ. Moreover, both the rotation around the z-axis and the translation in 3-D space are introduced to form the backscattering models:

1) Sphere:

$$\begin{aligned} S_{\mathrm{sph}} &= A_{\mathrm{sph}}\exp\left\{-j2\kappa(x_i\cos\varphi+y_i\sin\varphi-r_i)\right\}\\ A_{\mathrm{sph}} &= \sqrt{\pi}r_i\\ r_i &= r_{\mathrm{sph}}. \end{aligned} \tag{7}$$

2) Top-hat:

$$\begin{aligned} S_{\mathrm{top}} &= A_{\mathrm{top}}\sqrt{\frac{j\kappa}{\kappa_c}}\exp\left\{-j2\kappa(x_i\cos\varphi+y_i\sin\varphi-r_i)\right\}\\ A_{\mathrm{top}} &= \sqrt{4\sqrt{2}\kappa_c r_i}H_{\mathrm{top}}\begin{cases}\sin\theta & \theta\in\left[0,\frac{\pi}{4}\right]\\ \cos\theta & \theta\in\left[\frac{\pi}{4},\frac{\pi}{2}\right]\end{cases}\\ r_i &= r_{\mathrm{top}}\cos\theta. \end{aligned} \tag{8}$$

3) Cylinder:

$$\begin{aligned} S_{\mathrm{cyl}} &= A_{\mathrm{cyl}}\left(\frac{j\kappa}{\kappa_c}\right)^{\frac{1}{2}}\cos^{\frac{1}{2}}(\varphi-\varphi_i)\\ &\quad\times\exp\left\{-j2\kappa(x_i\cos\varphi+y_i\sin\varphi)\right\}\mathrm{sinc}(L_i\Delta_i)\\ A_{\mathrm{cyl}} &= L_{\mathrm{cyl}}\sqrt{\kappa_c r_{\mathrm{cyl}}}\\ L_i &= L_{\mathrm{cyl}}\cos\theta\\ \varphi_i &= \varphi_{\mathrm{cyl}}. \end{aligned} \tag{9}$$

4) Dihedral:

$$\begin{aligned} S_{\mathrm{dih}} &= A_{\mathrm{dih}}\left(\frac{j\kappa}{\kappa_c}\right)\exp\left\{-j2\kappa(x_i\cos\varphi+y_i\sin\varphi)\right\}\\ &\quad\times\mathrm{sinc}(L_i\Delta_i)\\ A_{\mathrm{dih}} &= \frac{2L_{\mathrm{dih}}H_{\mathrm{dih}}\kappa_c}{\sqrt{\pi}}\begin{cases}\sin\theta & \theta\in\left[0,\frac{\pi}{4}\right]\\ \cos\theta & \theta\in\left[\frac{\pi}{4},\frac{\pi}{2}\right]\end{cases}\\ L_i &= L_{\mathrm{dih}}\cos\theta\\ \varphi_i &= \varphi_{\mathrm{dih}}. \end{aligned} \tag{10}$$

5) Trihedral:

$$\begin{aligned} S_{\mathrm{tri}} &= A_{\mathrm{tri}}\left(\frac{j\kappa}{\kappa_c}\right)\cos(\varphi-\varphi_i)\\ &\quad\times\exp\left\{-j2\kappa(x_i\cos\varphi+y_i\sin\varphi)\right\}\\ A_{\mathrm{tri}} &= -\frac{2\sqrt{3}H_{\mathrm{tri}}^2\kappa_c}{\sqrt{\pi}}\begin{cases}\sin\left(\theta+\frac{\pi}{4}-\theta_{\mathrm{tri}}\right) & \theta\in[0,\theta_{\mathrm{tri}}]\\ \cos\left(\theta+\frac{\pi}{4}-\theta_{\mathrm{tri}}\right) & \theta\in\left[\theta_{\mathrm{tri}},\frac{\pi}{2}\right]\end{cases}\\ \varphi_i &= \begin{cases}\varphi_{\mathrm{tri}}+\frac{\pi}{4} & \varphi-\varphi_{\mathrm{tri}}\in\left[-\frac{\pi}{4},0\right]\\ \varphi_{\mathrm{tri}}-\frac{\pi}{4} & \varphi-\varphi_{\mathrm{tri}}\in\left[0,\frac{\pi}{4}\right]\end{cases}\\ \theta_{\mathrm{tri}} &= \tan^{-1}\left(2^{-\frac{1}{2}}\right). \end{aligned} \tag{11}$$

6) Rectangular plate:

$$\begin{aligned} S_{\mathrm{pla}} &= A_{\mathrm{pla}}\mathrm{sinc}(L_i\Delta_i)\\ &\quad\times\left\{\exp\left[-j2\kappa\left(x_i\cos\varphi+y_i\sin\varphi-r_i^{(1)}\right)\right]\right.\\ &\quad\left.-\exp\left[-j2\kappa\left(x_i\cos\varphi+y_i\sin\varphi-r_i^{(2)}\right)\right]\right\}\\ A_{\mathrm{pla}} &= \frac{L_{\mathrm{pla}}}{2\sqrt{\pi}\sin\theta}\\ L_i &= L_{\mathrm{pla}}\cos\theta\\ r_i^{(1)} &= -r_i^{(2)} = \frac{1}{2}H_{\mathrm{pla}}\sin\theta\\ \varphi_i &= \varphi_{\mathrm{pla}}. \end{aligned} \tag{12}$$

Fig. 3 shows the definitions of the parameters r_{sph}, r_{cyl}, r_{top}, L_{cyl}, L_{dih}, L_{pla}, H_{top}, H_{dih}, H_{tri}, and H_{pla}. (x_i,y_i) are the 2-D coordinates of the point Q under XOY coordination system. The orientations for cylinder, dihedral, trihedral, and rectangular plate are all defined as each ray $\overrightarrow{\mathrm{QU}}$. The aspect angles of these rays under XOY coordination system are $\varphi_{\mathrm{cyl}},\varphi_{\mathrm{dih}},\varphi_{\mathrm{tri}},\varphi_{\mathrm{pla}}$, respectively.

Obviously, canonical scattering responses can be represented by IASM. It is worth noting that the scattering responses of both the trihedral and the rectangular plate are the sum of two IASMs (refer to (13) and (14), shown at the top of the page).

GTD-inferred values of canonical scattering primitives are summarized in Table I. The comparisons between the inferred and measured values will be discussed in Section IV-A.

D. Theoretical Analysis for Imaging Behaviors of IASM

When preforming image reconstruction for the scattering responses that are generated by IASM, there exist four cases.

1) $L_i = 0, r_i = 0$
   IASM is degenerated into a point-scattering mechanism, e.g., trihedral.

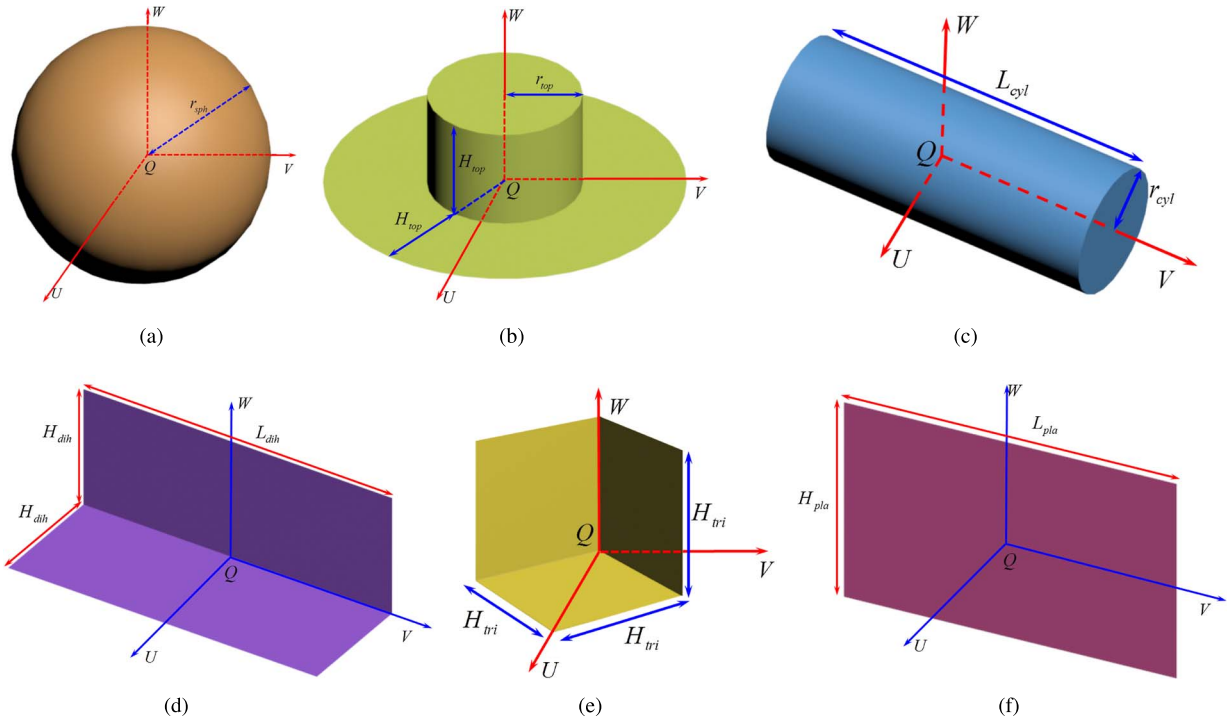


Fig. 3. Canonical scattering primitives and the relevant parameter definitions. (a) Sphere. (b) Top-hat. (c) Cylinder. (d) Dihedral. (e) Trihedral. (f) Rectangular plate.

TABLE I
GTD-INFERRED VALUES OF THE PARAMETERS FOR CANONICAL SCATTERING MECHANISMS

Canonical Shape	λ_i	γ_i	r_i	L_i	φ_i	$\varphi_i^{(1)}$	$\varphi_i^{(2)}$
Sphere	0	0	$r_{\rm sph}$	0	$\neq \pm\frac{\pi}{2}$	$-\pi$	π
Top-hat	0.5	0	$r_{\rm top}\cos\theta$	0	$\neq \pm\frac{\pi}{2}$	$-\pi$	π
Cylinder	0.5	0.5	0	$L_{\rm cyl}\cos\theta$	$\varphi_{\rm cyl}$	$\varphi_{\rm cyl}-\frac{\pi}{2}$	$\varphi_{\rm cyl}+\frac{\pi}{2}$
Dihedral	1	0	0	$L_{dih}\cos\theta$	φ_{dih}	$\varphi_{dih}-\frac{\pi}{2}$	$\varphi_{dih}+\frac{\pi}{2}$
Trihedral	1	1	0	0	$\varphi_{\rm tri}+\frac{\pi}{4}$ $\varphi_{\rm tri}-\frac{\pi}{4}$	$\varphi_{\rm tri}-\frac{\pi}{4}$ $\varphi_{\rm tri}$	$\varphi_{\rm tri}$ $\varphi_{\rm tri}+\frac{\pi}{4}$
Rectangular Plate	0	0	$\frac{1}{2}H_{\rm pla}\sin\theta$ $-\frac{1}{2}H_{\rm pla}\sin\theta$	$L_{\rm pla}\cos\theta$	$\varphi_{\rm pla}$	$\varphi_{\rm pla}-\frac{\pi}{2}$	$\varphi_{\rm pla}+\frac{\pi}{2}$

2) $L_i = 0, r_i \neq 0$
IASM exhibits an arc-scattering mechanism, e.g., sphere and top-hat.

3) $L_i \neq 0, r_i = 0$
If the incident wave of the radar sweeps over the normal direction of the scatterer, IASM stands for a line-segment scatterer whose endpoints are located at $(x_i - (L_i/2)\sin\varphi_i, y_i + (L_i/2)\cos\varphi_i)$ and $(x_i + (L_i/2)\sin\varphi_i, y_i - (L_i/2)\cos\varphi_i)$, respectively. Otherwise, IASM should be decomposed into two parts [refer to (15), shown at the bottom of the page]. Thus, IASM corresponds to two point scatterers which are located at the two endpoints, respectively.

4) $L_i \neq 0, r_i \neq 0$
Likewise, if the incident wave of the radar sweeps over the normal direction, IASM still exhibits a line-segment-scattering mechanism. The only difference is that the parameter r_i makes the line-segment-scatterer be translated along the direction φ_i. Then, the coordinates become $(x_i - r_i\cos\varphi_i - (L_i/2)\sin\varphi_i, y_i - r_i\sin\varphi_i + (L_i/2)$

$$
\begin{aligned}
&\phi^{(\mathrm{IASM})}\left(\kappa,\varphi;x_i,y_i,0,\lambda_i,\gamma_i,L_i,\varphi_i,\varphi_i^{(1)},\varphi_i^{(2)}\right)\\
&\quad=\frac{\kappa_c}{2}L_i^{-1}\sin^{-1}(\varphi-\varphi_i)\left\{\phi^{(\mathrm{IASM})}\left(\kappa,\varphi;x_i+\frac{L_i}{2}\sin\varphi_i,y_i-\frac{L_i}{2}\cos\varphi_i,0,\lambda_i-1,\gamma_i,0,\varphi_i,\varphi_i^{(1)},\varphi_i^{(2)}\right)\right.\\
&\qquad\left.-\phi^{(\mathrm{IASM})}\left(\kappa,\varphi;x_i-\frac{L_i}{2}\sin\varphi_i,y_i+\frac{L_i}{2}\cos\varphi_i,0,\lambda_i-1,\gamma_i,0,\varphi_i,\varphi_i^{(1)},\varphi_i^{(2)}\right)\right\} \qquad (15)
\end{aligned}
$$

$$\begin{aligned}&\phi^{(\mathrm{IASM})}\left(\kappa,\varphi;x_i,y_i,0,\lambda_i,\gamma_i,L_i,\varphi_i,\varphi_i^{(1)},\varphi_i^{(2)}\right)\\&\quad=\frac{\kappa_c}{2}L_i^{-1}\sin^{-1}(\varphi-\varphi_i)\left\{\phi^{(\mathrm{IASM})}\left(\kappa,\varphi;x_i+\frac{L_i}{2}\sin\varphi_i,y_i-\frac{L_i}{2}\cos\varphi_i,r_i,\lambda_i-1,\gamma_i,0,\varphi_i,\varphi_i^{(1)},\varphi_i^{(2)}\right)\right.\\&\qquad\left.-\phi^{(\mathrm{IASM})}\left(\kappa,\varphi;x_i-\frac{L_i}{2}\sin\varphi_i,y_i+\frac{L_i}{2}\cos\varphi_i,r_i,\lambda_i-1,\gamma_i,0,\varphi_i,\varphi_i^{(1)},\varphi_i^{(2)}\right)\right\}\end{aligned}\tag{16}$$

$\cos\varphi_i)$ and $(x_i-r_i\cos\varphi_i+(L_i/2)\sin\varphi_i, y_i-r_i\sin\varphi_i-(L_i/2)\cos\varphi_i)$, respectively. Otherwise, two circular arc scatterers are obtained by (16), shown at the top of the page, where their associated circular centers are located at the two endpoints.

III. Parameter Optimization With ISBL

A. Problem Description

Suppose that $S(\kappa_n,\varphi_n)(n=1,2,\ldots,N)$ stands for the collected scattering responses, where each pair (κ_n,φ_n) corresponds to a single observation. For convenience, the following column vectors are defined for $S(\kappa_n,\varphi_n)$, the noise term $\varepsilon(\kappa_n,\varphi_n)$ and the ith scattering term $\phi_i(\kappa_n,\varphi_n)$ as

$$\begin{aligned}\mathbf{t}&=[\ldots,S(\kappa_n,\varphi_n),\ldots]^{\mathrm{T}}\in\mathbb{C}^N\\\boldsymbol{\varepsilon}&=[\ldots,\varepsilon(\kappa_n,\varphi_n),\ldots]^{\mathrm{T}}\in\mathbb{C}^N\end{aligned}\tag{17}$$

$$\boldsymbol{\phi}_i=\boldsymbol{\phi}(\mathbf{x};\boldsymbol{\vartheta}_i)=[\ldots,\phi_i(\kappa_n,\varphi_n),\ldots]^{\mathrm{T}}\in\mathbb{C}^N\tag{18}$$

where the superscript "T" stands for the transpose

$$\mathbf{x}=[\boldsymbol{\kappa},\boldsymbol{\varphi}]\in\mathbb{R}^{N\times 2}$$

$$\boldsymbol{\vartheta}_i=\begin{cases}[x_i,y_i,\lambda_i,\gamma_i,L_i,\varphi_i]^{\mathrm{T}} & \text{for ASM}\\ \left[x_i,y_i,\lambda_i,\tilde{\gamma}_i,L_i,\varphi_i,\varphi_i^{(1)},\varphi_i^{(2)}\right]^{\mathrm{T}} & \text{for IASM}\end{cases}\tag{19}$$

$$\begin{aligned}\boldsymbol{\kappa}&=[\ldots,\kappa_n,\ldots]^{\mathrm{T}}\in\mathbb{C}^N\\\boldsymbol{\varphi}&=[\ldots,\varphi_n,\ldots]^{\mathrm{T}}\in\mathbb{C}^N.\end{aligned}\tag{20}$$

Then, the observation model in (1) can be rewritten as

$$\mathbf{t}=\boldsymbol{\Phi}\mathbf{w}+\boldsymbol{\varepsilon}\tag{21}$$

where $\boldsymbol{\varepsilon}$ is the noise vector; $\mathbf{w}=[A_1,A_2,\ldots,A_M]^{\mathrm{T}}\in\mathbb{C}^M$ is the coefficient vector; the matrix $\boldsymbol{\Phi}$ is composed of $\boldsymbol{\phi}_i$, i.e., $\boldsymbol{\Phi}=[\ldots,\boldsymbol{\phi}_i,\ldots]$ $(i=1,2,\ldots,M)$; each vector $\boldsymbol{\phi}_i$ is generated by a basis function $\boldsymbol{\phi}(\mathbf{x};\boldsymbol{\vartheta})$ with a specific parameter set $\boldsymbol{\vartheta}_i$; and the set is associated with a specific scattering component.

Referring to (21), the task of modeling the target response with a scattering model is a linear inverse problem, so we must introduce proper prior knowledge and sophistical optimization algorithms to cope with its ill-posedness. Generally, the number of the scattering components is considered to be far smaller than the number of the measured samples. Thus, it is reasonable to introduce sparsity constraints to regularize the inverse problem. The commonly used sparse representation methods rely on overcomplete dictionary matrices, resulting in inevitable grid errors and intensive computations introduced by the multiplication and inversion operations of the matrices. More important, in most cases, such computation is unacceptable. In this paper, we employ the ISBL method to overcome these difficulties. Hereinafter, we concisely deduce the complex-valued ISBL from its real-valued version [23] for the inverse problem.

B. SBL

SBL assigns hierarchical priors to both the coefficient vector $\mathbf{w}$ and the noise vector $\boldsymbol{\varepsilon}$, wherein each element follows a conditional Gaussian distribution, i.e.,

$$p(\mathbf{w}|\boldsymbol{\alpha})=\prod_{i=1}^{M}\mathcal{CN}\left(w_i|0,\alpha_i^{-1}\right)=\mathcal{CN}(\mathbf{0},\boldsymbol{\Gamma}^{-1})\tag{22}$$

$$p(\boldsymbol{\varepsilon}|\beta)=\mathcal{CN}(0,\beta^{-1}\mathbf{I})\tag{23}$$

where $\mathcal{CN}$ denotes that the vector is both circularly symmetric and complex [30], $\boldsymbol{\alpha}=[\alpha_1,\alpha_2,\ldots,\alpha_M]^{\mathrm{T}}\in\mathbb{R}^M$ is a hyperparameter vector, α_i denotes the precision of the weight w_i, $\boldsymbol{\Gamma}=\mathrm{diag}(\boldsymbol{\alpha})$ denotes a diagonal matrix whose diagonal entries are the elements of the vector $\boldsymbol{\alpha}$, and the hyperparameter $\beta\in\mathbb{R}$ denotes the precision of the observation noise.

Referring to [23], we assume an uninformative prior for β

$$p(\beta)=\text{const}\tag{24}$$

and a sparsity prior for $\boldsymbol{\alpha}$

$$p(\boldsymbol{\alpha}|\beta)\propto\exp\left(-\rho\,\mathrm{trace}(\mathbf{F})\right)\tag{25}$$

respectively, where $\mathbf{F}=\beta\boldsymbol{\Phi}\mathbf{C}_{\mathbf{w}}\boldsymbol{\Phi}^{\mathrm{H}}$ is the smoothing matrix, the superscript "H" is the conjugate transpose, the sparsity parameter ρ controls the amount of desired sparsity, and $\mathrm{trace}(\,)$ is the trace of a matrix. In this paper, the parameter ρ is set to be $\log(N)/2$, which corresponds to the Bayesian information criterion [23].

By inference from these prior assumptions, the *maximum a posteriori* estimate of $\mathbf{w}$ turns out to be

$$\hat{\mathbf{w}}=\beta\mathbf{C}_{\mathbf{w}}\boldsymbol{\Phi}^{\mathrm{H}}\mathbf{t}\tag{26}$$

where

$$\mathbf{C}_{\mathbf{w}}=(\beta\boldsymbol{\Phi}^{\mathrm{H}}\boldsymbol{\Phi}+\boldsymbol{\Gamma}).\tag{27}$$

Benefiting from the hierarchical prior, the necessary parameters can be automatically learned during the model optimization

$$\hat{\alpha}_i = \left\{ \sqrt{\frac{1}{4}\left[\bar{w}_i^2 - (\rho-1)[\mathbf{C_w}]_{ii}\right]^2 + \rho[\mathbf{C_w}]_{ii}^2} + \frac{1}{2}\left(\bar{w}_i^2 - (\rho-1)[\mathbf{C_w}]_{ii}\right) \right\}^{-1}$$
$$\hat{\beta} = 2N \left\{ \sqrt{\left[\|\mathbf{t} - \mathbf{\Phi}\bar{\mathbf{w}}\|_2^2 + \mathrm{trace}(\mathbf{C_w}\mathbf{\Phi}^{\mathrm{H}}\mathbf{\Phi})\right]^2 + 4N\rho\,\mathrm{trace}(\mathbf{C_w}\mathbf{\Phi}^{\mathrm{H}}\mathbf{\Phi})} + \|\mathbf{t} - \mathbf{\Phi}\bar{\mathbf{w}}\|_2^2 + \mathrm{trace}\left(\mathbf{C_w}\mathbf{\Phi}^{\mathrm{H}}\mathbf{\Phi}\right) \right\}^{-1} \quad (33)$$

by maximizing the following logarithm of the marginal likelihood:

$$(\hat{\boldsymbol{\alpha}}, \hat{\beta}) = \arg\max_{(\boldsymbol{\alpha},\beta)} \log p(\mathbf{t}, \boldsymbol{\alpha}|\beta) \quad (28)$$

$$\begin{aligned} \mathbf{\Lambda}^s(\beta, \boldsymbol{\alpha}) &= \log p(\mathbf{t}, \boldsymbol{\alpha}|\beta) \\ &= \log \mathcal{CN}(\mathbf{0}, \mathbf{C}) - \rho\,\mathrm{trace}(\mathbf{F}) \\ &= -N\log(\pi) - \log|\mathbf{C}| - \mathbf{t}^{\mathrm{H}}\mathbf{C}^{-1}\mathbf{t} \\ &\quad - \rho\left(M - \sum_{i=1}^{M} \alpha_i [\mathbf{C_w}]_{ii}\right) \end{aligned} \quad (29)$$

where [23]

$$\mathbf{C} = (\beta^{-1}\mathbf{I} + \mathbf{\Phi}\mathbf{\Gamma}^{-1}\mathbf{\Phi}^{\mathrm{H}}) \quad (30)$$

and $[\,]_{ii}$ denotes the ith diagonal element of a matrix.

Then, the partial derivatives of $\mathbf{\Lambda}^s(\beta, \boldsymbol{\alpha})$ with respect to $\log\alpha_i$ and $\log\beta$ [23] are

$$\begin{aligned} \frac{\partial \mathbf{\Lambda}^s(\beta, \boldsymbol{\alpha})}{\partial \log\alpha_i} &= \left(1 - \alpha_i[\mathbf{C_w}]_{ii} - \alpha_i\bar{w}_i^2\right) \\ &\quad + \rho\left(1 - \alpha_i[\mathbf{C_w}]_{ii}\right)\alpha_i[\mathbf{C_w}]_{ii} \end{aligned} \quad (31)$$

$$\begin{aligned} \frac{\partial \mathbf{\Lambda}^s(\beta, \boldsymbol{\alpha})}{\partial \log\beta} &= \left[\frac{N}{\beta} - \|\mathbf{t} - \mathbf{\Phi}\bar{\mathbf{w}}\|_2^2 - \mathrm{trace}(\mathbf{C_w}\mathbf{\Phi}^{\mathrm{H}}\mathbf{\Phi})\right] \\ &\quad - \beta\rho\,\mathrm{trace}(\mathbf{C_w}\mathbf{\Phi}^{\mathrm{H}}\mathbf{\Phi}) \end{aligned} \quad (32)$$

respectively, and by setting them to zeros, we can obtain the update formula as (33), shown at the top of the page, where $\|\ \|_2$ denotes the ℓ_2-norm of a vector.

C. ISBL

The implementation of SBL demands an overcomplete dictionary matrix, and each atom is generated by executing a basis function with respect to the preset discrete values of the continuous parameters. This optimization is therefore performed on a discrete subset of the feasible region. In addition, the choice of the quantized interval of each parameter has to balance the grid error and computational complexity introduced by the operations of the matrix. In particular, optimizing multidimensional parameters must bear such computational burden to better the estimation precision. Incremental optimization method for the SBL [23], [31], namely ISBL, can overcome these shortcomings.

Its basic idea is to learn individual atom parameters in an incremental manner before the global optimization. To isolate the dependence of the objective function on a single atom, we decompose the matrix $\mathbf{C}$ in (30) as [31]

$$\mathbf{C} = \mathbf{C}_{-i} + \alpha_i^{-1}\boldsymbol{\phi}_i\boldsymbol{\phi}_i^{\mathrm{H}} \quad (34)$$

where $\mathbf{C}_{-i}$ contains the contribution of all of the atoms except the ith atom. To simplify derivations, define the sparsity factor s_i, the quality factor q_i, and its absolute square p_i as

$$s_i = \boldsymbol{\phi}_i^{\mathrm{H}}\mathbf{C}_{-i}^{-1}\boldsymbol{\phi}_i, q_i = \boldsymbol{\phi}_i^{\mathrm{H}}\mathbf{C}_{-i}^{-1}\mathbf{t}, p_i = |q_i|^2. \quad (35)$$

Then, referring to [23], we obtain the dependence of the objective function $\mathbf{\Lambda}^s(\beta, \boldsymbol{\alpha})$ on a single hyperparameter α_i as

$$\begin{aligned} \ell^s(\alpha_i) &= \mathbf{\Lambda}^s(\beta, \boldsymbol{\alpha}) - \mathbf{\Lambda}^s(\beta, \boldsymbol{\alpha}_{-i}) \\ &= \log\frac{\alpha_i}{\alpha_i + s_i} + \frac{p_i - \rho s_i}{\alpha_i + s_i} \end{aligned} \quad (36)$$

where the vector $\boldsymbol{\alpha}_{-i}$ is $\boldsymbol{\alpha}$ with α_i omitted. It has been proved in [23] that $\ell^s(\alpha_i)$ has a unique maximum at

$$\alpha_i = \begin{cases} \frac{s_i^2}{p_i - (\rho+1)s_i} & \text{if } p_i > (\rho+1)s_i \\ \infty & \text{otherwise.} \end{cases} \quad (37)$$

Obviously, only the atoms that meet $p_i > (\rho+1)s_i$ should be introduced into the model. For the atoms with continuous parameters, the optimal parameter value can be obtained by maximizing

$$\ell^s\left(\frac{s_i^2}{p_i - (\rho+1)s_i}\right) = \left[\left(\frac{p_i}{s_i} - \rho\right) - \log\left(\frac{p_i}{s_i} - \rho\right) + 1\right] \quad (38)$$

with respect to the parameter vector $\boldsymbol{\vartheta}_i$. Suppose that

$$u_i = \frac{p_i}{s_i} = \frac{\boldsymbol{\phi}_i^{\mathrm{H}}\mathbf{C}_{-i}^{-1}\mathbf{t}\mathbf{t}^{\mathrm{H}}\mathbf{C}_{-i}^{-1}\boldsymbol{\phi}_i}{\boldsymbol{\phi}_i^{\mathrm{H}}\mathbf{C}_{-i}^{-1}\boldsymbol{\phi}_i}. \quad (39)$$

Then the derivative of ℓ^s on u_i can be given by

$$\frac{d\ell^s}{du_i} = \left(1 - \frac{1}{u_i - \rho}\right) > 0 \text{ for } u_i > \rho + 1. \quad (40)$$

ℓ^s is a strictly monotonically increasing function about u_i, and the optimization can be therefore simplified as

$$\arg\max_{\boldsymbol{\vartheta}_i} \ell^s = \arg\max_{\boldsymbol{\vartheta}_i} u_i. \quad (41)$$

At each iteration, only the atom that increases the marginal likelihood the most is added to the model. After that, the objective function $\mathbf{\Lambda}^s$ is globally optimized with respect to the whole parameter set in the model. The continuous maximization of ℓ^s and $\mathbf{\Lambda}^s$ is performed by the quasi-Newton BFGS method [23]. The required derivatives are given in Appendix A.

As a result, the inclusion of continuous optimization effectively eliminates the grid error caused by discrete optimization, and the omission of large dictionary matrices greatly decreases the computational complexity caused by large matrix operations. Meanwhile, this incremental optimization manner also facilitates the hybrid atom optimization that examines

many types of basis functions simultaneously during each local optimization.

D. Algorithm

Inverse scattering with ISBL consists of alternating iterations among four steps: candidate atom parameter set generation, inactive candidate atom optimization and selection, active atom optimization, and unnecessary active atom removal.

1) Candidate Atom Parameter Set Generation: If the candidate atom parameter set $\boldsymbol{\Theta}^{(\mathrm{iact})}$ is empty at the ith iteration, the image $\mathbf{I}_i$ of the residual response $(\mathbf{t}-\boldsymbol{\Phi}\hat{\mathbf{w}})$ is reconstructed by the convolution backprojection (CBP) method [32]

$$\mathbf{I}_i = \mathrm{CBP}(\mathbf{t}-\boldsymbol{\Phi}\hat{\mathbf{w}}). \tag{42}$$

Afterward, we detect all possible scatterers from $\mathbf{I}_i$, and the implementation details will be discussed in Section III-E. Finally, the initial values of the geometrical parameters, including $x_{i,l}$, $y_{i,l}$, $r_{i,l}$, $L_{i,l}$, $\varphi_{i,l}$, $\varphi_{i,l}^{(1)}$, and $\varphi_{i,l}^{(2)}$, $(l=1,2,\ldots,P)$, are roughly estimated from the detected scatterers. The dependence parameters are selected from their individual preset sets, such as

$$\lambda_{i,l} \in \{-1,-0.5,0,0.5,1\} \triangleq \Psi_\lambda \tag{43}$$

and for ASM

$$\gamma_{i,l} \in \left\{-\kappa_c^{-1},-(2\kappa_c)^{-1},0,(2\kappa_c)^{-1},\kappa_c^{-1}\right\} \triangleq \Psi_\gamma \tag{44}$$

for IASM

$$\tilde{\gamma}_{i,l} \in \{0,0.5,1\} \triangleq \Psi_{\tilde{\gamma}}. \tag{45}$$

All possible dependence parameter values are separately combined with geometrical parameter values to form the initial set $\boldsymbol{\Theta}^{(\mathrm{iact})}$, which contains Q vectors and

$$Q = \begin{cases} P\times 5\times 5 = 25P & \text{for ASM} \\ P\times 5\times 3 = 15P & \text{for IASM.} \end{cases} \tag{46}$$

2) Inactive Atom Optimization and Selection: Separately maximize the objective function u_i using the quasi-Newton BFGS method with each initial parameter vector in $\boldsymbol{\Theta}^{(\mathrm{iact})}$. Denote the maximum value and the associated optimum parameter vector as $u_{i,l}^{(\max)}$ and $\boldsymbol{\vartheta}_{i,l}^{(\max)}(l=1,2,\ldots,Q)$, respectively. If the optimum parameter vector $\boldsymbol{\vartheta}_i^{(\max)}$, which increases u_i the most, meets $u_i^{(\max)} > (\rho+1)$, it can be selected into the active parameter set $\boldsymbol{\Theta}$.

3) Active Atom Optimization: Globally maximize $\boldsymbol{\Lambda}^s$ using the quasi-Newton BFGS method with the currently active parameter set $\boldsymbol{\Theta}$, while the required derivatives can be efficiently computed by (A2) and (A3).

4) Unnecessary Active Atom Removal: For each parameter vector in $\boldsymbol{\Theta}$, recompute $u_l(l=1,2,\ldots,i)$, and exclude all of the parameter vectors that meet $u_l < (\rho+1)$.

5) Stop Criterion: We assume that the algorithm has converged when step 1) cannot give any meaningful atom parameter vectors. That is, the newly generated atom parameter set satisfies $u_i^{(\max)} \le (\rho+1)$.

The complete procedure of this algorithm is summarized in Algorithm 1.

Algorithm 1 The ISBL Algorithm for Inverse Scattering Problems.

Inputs: $\mathbf{t}\in\mathbb{C}^{N\times 1}$, $\mathbf{x}=[\boldsymbol{\kappa},\boldsymbol{\varphi}]\in\mathbb{C}^{N\times 2}$.
Parameters: Set $\rho=\log N/2$.
Initialization: Set $\boldsymbol{\Theta}\leftarrow\varnothing$, $\boldsymbol{\Theta}^{\mathrm{iact}}\leftarrow\varnothing$, $i\leftarrow 0$, $\boldsymbol{\Phi}\leftarrow[]$, $\boldsymbol{\Gamma}\leftarrow[]$, $\mathbf{C}_{\mathbf{w}}\leftarrow[]$, $\hat{\mathbf{w}}\leftarrow[]$.
while 1 **do**
  Detect scatterers and roughly estimate $x_{i,l}, y_{i,l}, r_{i,l}, L_{i,l}, \varphi_{i,l}, \varphi_{i,l}^{(1)}, \varphi_{i,l}^{(2)} (l=1,2,\ldots,P)$ from $\mathbf{I}_i$ (see (42)).
  for $l=1,2,\ldots,P_i$ **do**
    for $\lambda_{i,l}\in\Psi_\lambda$ and $\gamma_{i,l}\in\Psi_\gamma$ (for ASM) or $\Psi_{\tilde{\gamma}}$ (for IASM) **do**
      $\boldsymbol{\vartheta}_{i,l}^{(0)} = \begin{cases} [x_{i,l}, y_{i,l}, \lambda_{i,l}, \gamma_{i,l}, L_{i,l}, \varphi_{i,l}]^{\mathrm{T}} \\ \quad \text{for ASM} \\ [x_{i,l}, y_{i,l}, r_{i,l}, \lambda_{i,l}, \tilde{\gamma}_{i,l}, L_{i,l}, \varphi_{i,l}, \varphi_{i,l}^{(1)}, \varphi_{i,l}^{(2)}]^{\mathrm{T}} \\ \quad \text{for IASM.} \end{cases}$
      $\boldsymbol{\Theta}^{(\mathrm{iact})} \leftarrow \boldsymbol{\Theta}^{(\mathrm{iact})} \cup \{\boldsymbol{\vartheta}_{i,l}^{(0)}\}$.
    end for
  end for
  Compute Q by (46) and set $R\leftarrow 0$.
  while $Q\neq 0$ **do**
    Set $i\leftarrow i+1$.
    if $R=0$ **then**
      for $l=1,2,\ldots,Q$ **do**
        Maximize u_i using the BFGS method with $\boldsymbol{\vartheta}_{i,l}^{(0)}\in\boldsymbol{\Theta}^{(\mathrm{iact})}$.
        Set $u_{i,l}^{(\max)}\leftarrow\max\limits_{\boldsymbol{\vartheta}_{i,l}} u_i$ and $\boldsymbol{\vartheta}_{i,l}^{(\max)}\leftarrow\arg\max\limits_{\boldsymbol{\vartheta}_{i,l}} u_i$.
      end for
    else
      Set $\boldsymbol{\vartheta}_{i,l}^{(\max)}\leftarrow\boldsymbol{\vartheta}_{i,l}^{(0)}$ for $\boldsymbol{\vartheta}_{i,l}^{(0)}\in\boldsymbol{\Theta}^{(\mathrm{iact})}$ and compute $u_{i,l}^{(\max)}$ by (39).
    end if
    Set $l^{(\max)}\leftarrow\arg\max\limits_{l} u_{i,l}^{(\max)}$, $u_i^{(\max)}\leftarrow u_{i,l^{(\max)}}^{(\max)}$ and $\boldsymbol{\vartheta}_i^{(\max)}\leftarrow\boldsymbol{\vartheta}_{i,l^{(\max)}}^{(\max)}$.
    if $R=0$ and $u_i^{(\max)}\le(\rho+1)$ **then go to** 37
    else if $u_i^{(\max)}\le(\rho+1)$ **then** set $i\leftarrow i-1$ and **go to** 5
    else
      Set $R\leftarrow 1$.
    end if
    Set $\boldsymbol{\Theta}\leftarrow\boldsymbol{\Theta}\cup\{\boldsymbol{\vartheta}_{i^{(\max)}}^{(\max)}\}$, $\boldsymbol{\Theta}^{(\mathrm{iact})}\leftarrow\boldsymbol{\Theta}^{(\mathrm{iact})}/\{\boldsymbol{\vartheta}_{i^{(\max)}}^{(\max)}\}$ and $Q\leftarrow Q-1$.
    Compute s_i, q_i, p_i, α_i by (35) and (37), and update $\boldsymbol{\Phi}$, $\boldsymbol{\Gamma}$ and $\mathbf{C}_{\mathbf{w}}$.
    Maximize $\boldsymbol{\Lambda}^s$ using the BFGS method with $\boldsymbol{\Theta}$.
    Compute $\boldsymbol{\alpha}_l(l=1,2,\ldots,i)$ and β by (33), and recompute all u_l by (39).
    Remove $\boldsymbol{\vartheta}_{i,l}$ satisfying $u_l<\rho+1$ from $\boldsymbol{\Theta}^{(\mathrm{iact})}$.
    Subtract the number of the removed vectors from i, and update $\boldsymbol{\Phi}$, $\boldsymbol{\Gamma}$, $\mathbf{C}_{\mathbf{w}}$ and $\hat{\mathbf{w}}$.
  end while
end while
Outputs: $\boldsymbol{\Theta}$, $\boldsymbol{\Phi}$, β and $\hat{\mathbf{w}}$.

The most complicated computation for each ISBL iteration occurs in the global active atom optimization with the quasi-Newton BFGS method. We therefore mainly examine the computational cost of this step. For convenience, we assume that no atom is removed from the active atom set during the whole iteration, and then, the active atom set contains i atoms or $n_i = 9i$ parameters at ith ISBL iteration. Let $\mathcal{C}(\Lambda^s)$ be the cost of one function evaluation of the objective function Λ^s. Its matrix inversion and determinant both take $\mathcal{O}(i^3)$ flops, while its matrix multiplication takes $\mathcal{O}(Ni^2)$ flops [33]. Since $N \gg i$, $\mathcal{C}(\Lambda^s) = \mathcal{O}(Ni^2)$. Let $n_{i,k}$ be the function evaluation times at the kth BFGS iteration. The computational cost per BFGS iteration is $\mathcal{O}(n_i^2) + \mathcal{O}(n_{i,k}\mathcal{C}(\Lambda^s))$ [34]. Provided that N_i BFGS iterations are performed for ith ISBL iteration, it has a complexity of $\mathcal{O}(N_i n_i^2) + \mathcal{O}(Ni^2 \sum_{k=1}^{N_i} n_{i,k})$.

E. Scatterer Detection and Parameter Selection

First of all, we compute a threshold (denoted as T_{res}) for the image $\mathbf{I}_i$. For high-SNR images, we use Otsu's method [35] to accomplish the task, while for low-SNR images (SNR < 0 dB), we utilize the constant false alarm rate detector to directly compute the threshold at a specific false alarm rate P_{fa}. Since the background noise of the amplitude image follows the Rayleigh distribution, the maximum likelihood estimate of the Rayleigh parameter b can be represented as [36]

$$\hat{b} = \sqrt{\frac{1}{2N_i}\sum_{n_i}[\mathbf{I}_i]_{n_i}^2} \tag{47}$$

where $[\mathbf{I}_i]_{n_i}$ denotes the n_ith element of the vectorized matrix. It implies that, for convenience, we use the noisy image as the background noise, which will certainly cause a little impact on ISBL optimization, but it is not serious. In particular, when the SNR further reduces, the impact may become negligible. We empirically set $P_{\text{fa}} = 0.01$ in this paper and deduce the threshold from the inverse cumulative distribution function [36]

$$T_{\text{res}} = \sqrt{-2\hat{b}^2 \log P_{\text{fa}}}. \tag{48}$$

Then, all possible point scatterers are detected by searching for local extreme points with intensities larger than T_{res}. By thresholding with T_{res}, we convert the image $\mathbf{I}_i$ to a binary image. Then, we perform a morphological operation [37] on the binary image to produce the skeleton image, with which both candidate line-segment- and circular arc scatterers are extracted by Hough-transform-based methods [38], [39].

In principle, we select the necessary parameters so conservatively as to proper improve the detection rate, which inevitably introduces a number of false alarms. Nonetheless, these false scatterers will be subsequently filtered by the local optimization of ISBL. This filtering leads to a bit extra computation that is well worth committing.

IV. Numerical Experiments

Three data sets are selected for numerical experiments. First, individually simulated electromagnetic scattering data for six canonical primitives are used to testify the inferred values in Table I. Second, the simulated electromagnetic scattering data for a scene of six canonical primitives are applied to evaluate the modeling performance. The scattering responses in the former two groups of experiments are predicted by the electromagnetic simulation software "CST Microwave Studio 2014" [40]. Third, the publicly released "Backhoe Data Dome, Version 1.0" by the Air Force Research Laboratory are selected to demonstrate the effectiveness when applying the propose model to realistic complicated targets.

A. Inverse Scattering for Individual Canonical Scatterers

Numerical inverse scattering experiments are carried out separately for six canonical scattering objects based on IASM and ISBL. The used electromagnetic scattering data are generated by CST with the method-of-moments (MoM) or the shooting and bouncing ray (SBR) methods. Radar parameters are as follows: the frequency ranges from 6 to 12 GHz, the elevation angle is fixed to 15°, and the azimuth angle goes from $-40°$ to 40° for the trihedral and from $-85°$ to $85°$ for the other objects. To avoid loss of generality, three experiments with different parameters are specified for each shape. All parameters and estimates are summarized in Table II. The columns of $\lambda_i, \tilde{\gamma}_i, r_i, L_i, \varphi_i$ contain the GTD-inferred values, while the columns of $\hat{\lambda}_i^{(\text{HH})}$, $\hat{\tilde{\gamma}}_i^{(\text{HH})}, \hat{r}_i^{(\text{HH})}, \hat{L}_i^{(\text{HH})}, \hat{\varphi}_i^{(\text{HH})}$ and $\hat{\lambda}_i^{(\text{VV})}, \hat{\tilde{\gamma}}_i^{(\text{VV})}, \hat{r}_i^{(\text{VV})}, \hat{L}_i^{(\text{VV})}$, $\hat{\varphi}_i^{(\text{VV})}$ contain their estimates from HH and VV polarized responses, respectively.

By contrast, all estimates, except the results in boldface, conform closely to the corresponding GTD-inferred values. Those discrepancies can be divided into four categories. First, it is considered that the scattering response of a trihedral theoretically originates from two IASM atoms, while the experiment shows that this response can be well approximated by only one IASM atom. In that case, φ_i represents the azimuth angle of the trihedral, and its true value is 0, which well agrees with the estimates. Therefore, the estimates of φ_i agree with the true values. Second, after repeated comparisons between the objective function of the GTD-inferred value and the estimate for cylinders with different sizes, we can confirm that the optimum solution of $\tilde{\gamma}_i$ is 0. That is to say, the inferred value of $\tilde{\gamma}_i$ should be indeed 0 but 0.5. Third, although the physical sizes are not always much greater than one wavelength that the high-frequency approximation is inapplicable, most estimates (see Table II) still agree with the inferred values, except that the $\tilde{\gamma}_i$ estimates for dihedrals, trihedrals, and rectangular plates have relatively large values and/or variances. Thus, when discriminating among these objects, other parameters, such as λ_i and L_i, may be more reliable.

From GTD theory, only dominant scattering components are considered into the theoretical models, whereas the realistic scattering responses contain several other components. It is an important error source for the theoretical models. As shown

TABLE II
GTD-INFERRED AND MEASURED VALUES OF THE PARAMETERS

Shape φ (°)	Size unit(cm)	Simu. method	λ_i	$\hat{\lambda}_i^{(HH)}$	$\hat{\lambda}_i^{(VV)}$	γ_i	$\hat{\gamma}_i^{(HH)}$	$\hat{\gamma}_i^{(VV)}$	r_i	$\hat{r}_i^{(HH)}$ unit(cm)	$\hat{r}_i^{(VV)}$	l	$\hat{l}_i^{(HH)}$ unit(cm)	$\hat{l}_i^{(VV)}$	φ_i	$\hat{\varphi}_i^{(HH)}$ unit($\times\pi$ rad)	$\hat{\varphi}_i^{(VV)}$
Sphere $[-85, 85]$	$r=5$	MoM	0	0.04	0.04	0	0.00	0.00	5.00	4.97	4.97	0	0.00	0.00	$\neq\pm0.5$	-0.66	-0.17
	$r=20$	SBR	0	0.00	0.00	0	0.00	0.00	20.0	20.0	20.0	0	0.00	0.00	$\neq\pm0.5$	-0.00	-0.00
	$r=50$	SBR	0	0.00	0.00	0	0.00	0.00	50.0	50.0	50.0	0	0.00	0.00	$\neq\pm0.5$	0.25	-0.25
Top-hat $[-85, 85]$	$r=5, H=40$	SBR	0.5	0.49	0.53	0	0.00	0.00	4.83	4.84	4.82	0	0.00	0.00	$\neq\pm0.5$	-0.20	0.18
	$r=10, H=30$	SBR	0.5	0.49	0.46	0	0.00	0.00	9.66	9.70	9.64	0	0.00	0.00	$\neq\pm0.5$	-0.11	-0.24
	$r=20, H=40$	SBR	0.5	0.49	0.50	0	0.00	0.00	19.3	19.4	19.3	0	0.00	0.00	$\neq\pm0.5$	0.02	0.08
Cylinder $[-85, 85]$	$r=5, L=20$	MoM	0.5	0.49	0.51	0.5	**0.00**	**0.00**	0	**0.69**	**0.71**	19.3	19.4	19.5	0	-0.00	0.00
	$r=5, L=50$	SBR	0.5	0.48	0.48	0.5	**0.00**	**0.00**	0	**0.66**	**0.73**	48.3	48.3	48.2	0	0.00	-0.00
	$r=10, L=50$	SBR	0.5	0.49	0.49	0.5	**0.00**	**0.00**	0	**1.22**	**1.37**	48.3	48.4	48.1	0	-0.00	-0.00
Dihedral $[-85, 85]$	$H=10, L=20$	SBR	1	0.87	0.98	0	**5.64**	**13.5**	0	**-2.35**	**-3.26**	19.3	18.6	18.6	0	-0.00	-0.00
	$H=10, L=30$	SBR	1	0.88	0.99	0	**5.46**	**13.1**	0	**-2.34**	**-3.48**	29.0	28.3	28.3	0	-0.00	0.00
	$H=25, L=50$	SBR	1	0.97	0.99	0	**14.5**	**20.6**	0	**-6.19**	**-7.03**	48.3	47.0	47.0	0	0.00	0.00
Trihedral $[-40, 40]$	$H=20$	MoM	1	0.87	1.01	1	**7.10**	**6.52**	0	0.14	-0.38	0	0.00	0.00	±0.25	**0.00**	**-0.00**
	$H=30$	SBR	1	0.99	1.00	1	**7.04**	**7.05**	0	0.14	-0.14	0	0.00	0.00	±0.25	**0.00**	**0.00**
	$H=50$	SBR	1	1.00	1.00	1	**7.06**	**7.05**	0	0.09	-0.09	0	0.00	0.00	±0.25	**0.00**	**0.00**
Rectangu -lar Plate $[-85, 85]$	$H=20, L=30$	MoM	0	0.12	0.07	0	0.00	0.00	-2.59	-1.53	-2.22	29.0	28.6	28.6	0	-0.00	-0.00
			0	0.03	0.09	0	0.00	0.00	2.59	2.52	2.53	29.0	29.0	29.0	0	0.00	-0.00
	$H=30, L=40$	SBR	0	0.00	0.00	0	0.00	0.00	-3.88	-4.01	-3.84	38.6	38.4	38.7	0	0.00	0.00
			0	0.00	0.00	0	**1.00**	0.00	3.88	3.73	3.75	38.6	38.5	38.8	0	0.00	0.00
	$H=30, L=60$	SBR	0	0.00	0.00	0	0.00	0.00	-3.88	-3.97	-3.84	58.0	57.9	58.0	0	0.00	0.00
			0	0.00	0.00	0	**0.58**	0.00	3.88	3.64	4.01	58.0	57.8	57.8	0	0.00	-0.00

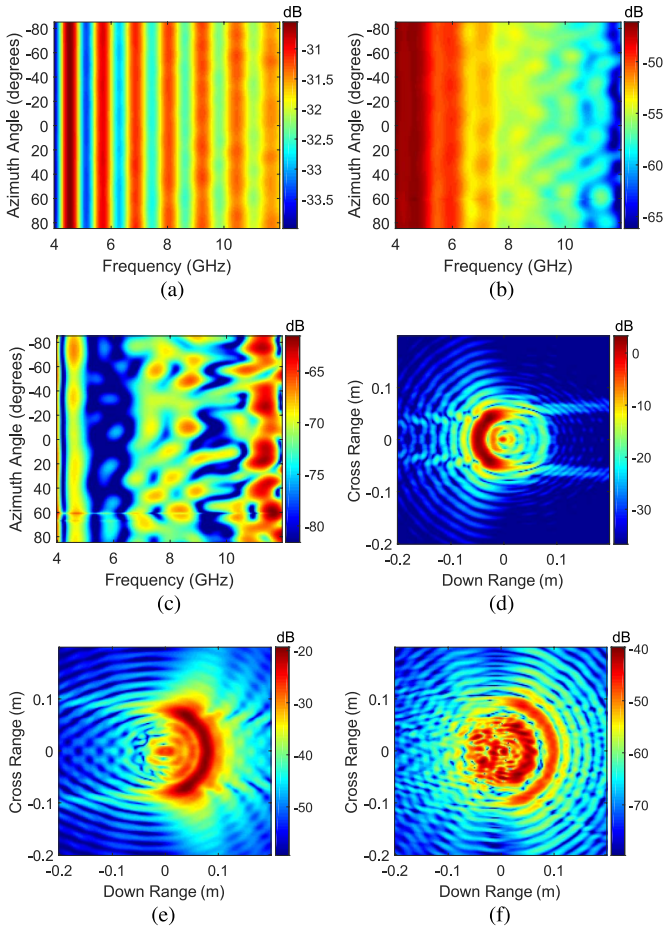


Fig. 4. Backscattering response and residual responses after inverse scattering for a sphere, and their CBP images. (a) Backscattering response. (b) Residual response after one component extracted. (c) Residual response after two components extracted. (d)–(f) CBP images of (a)–(c), respectively.

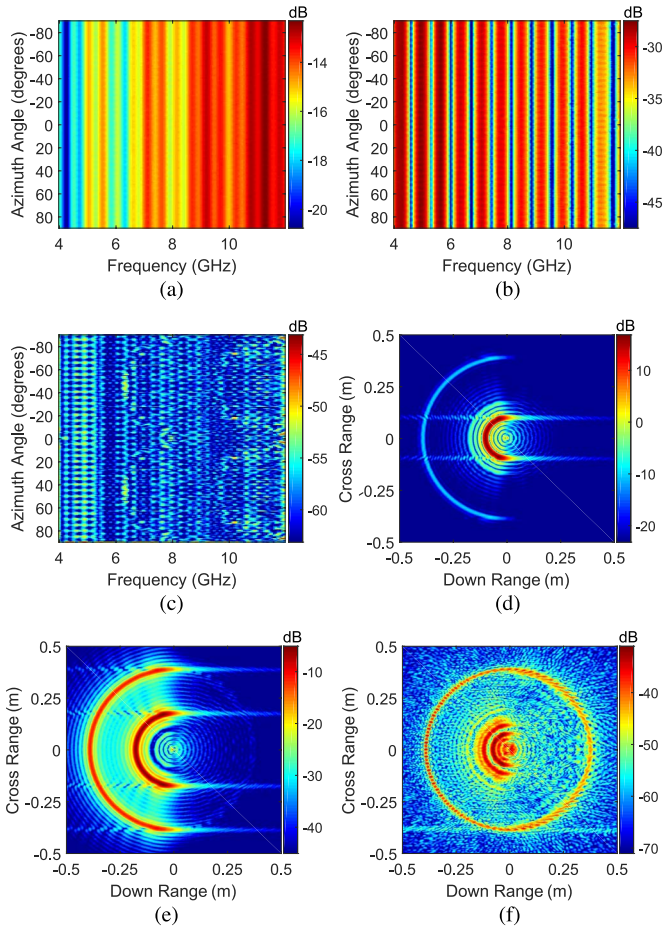


Fig. 5. Backscattering response and residual responses after inverse scattering for a top-hat, and their CBP images. (a) Backscattering response. (b) Residual response after one component extracted. (c) Residual response after three components extracted. (d)–(f) CBP images of (a)–(c), respectively.

in Figs. 4(e), 5(e), 6(d), 7(e), 8(d), and 9(e), after the extraction of the dominant energy, physical scattering components obviously remain. For sphere, the residual scattering component originates from a diffracted arc scatterer (see Fig. 4). For top-hat, the residual scattering components consist of two arc scattering responses, and one of them can be assured from the edge of the bottom plate (see Fig. 5). For cylinder, the edges and the surface of the cylinder create the residual scattering

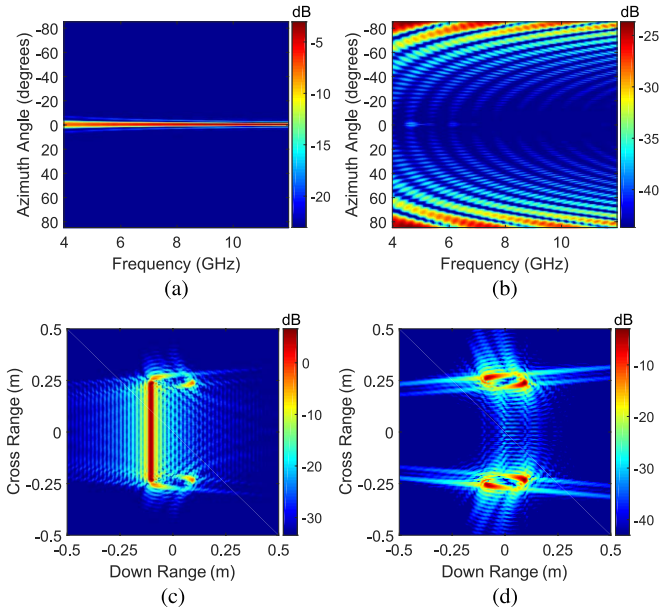


Fig. 6. Backscattering response and residual response after inverse scattering for a cylinder, and their CBP images. (a) Backscattering response. (b) Residual response after one dominant component extracted. (c) and (d) CBP images of (a) and (b), respectively.

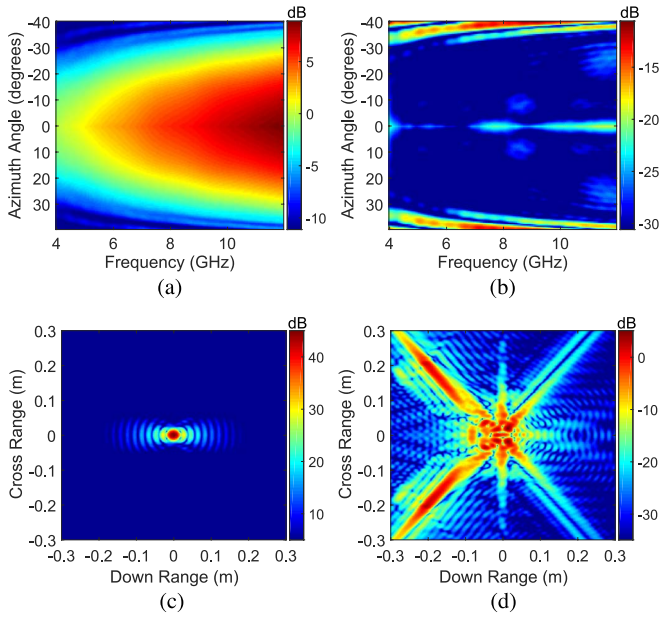


Fig. 8. Backscattering response and residual response after inverse scattering for a trihedral, and their CBP images. (a) Backscattering response. (b) Residual response after one dominant component extracted. (c) and (d) CBP images of (a) and (b), respectively.

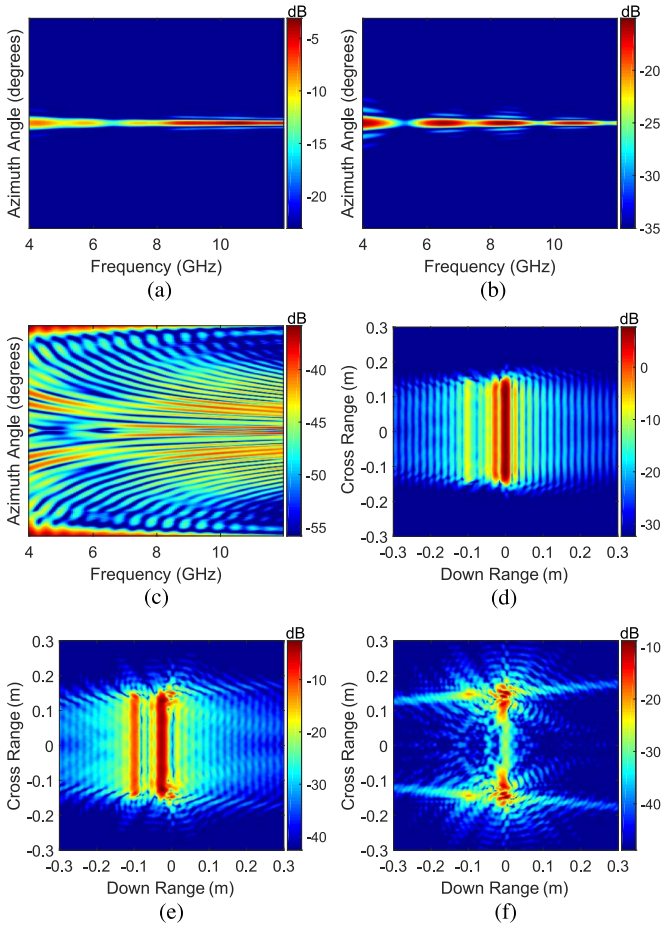


Fig. 7. Backscattering response and residual responses after inverse scattering for a dihedral, and their CBP images. (a) Backscattering response. (b) Residual response after one component extracted. (c) Residual response after three components extracted. (d)–(f) CBP images of (a)–(c), respectively.

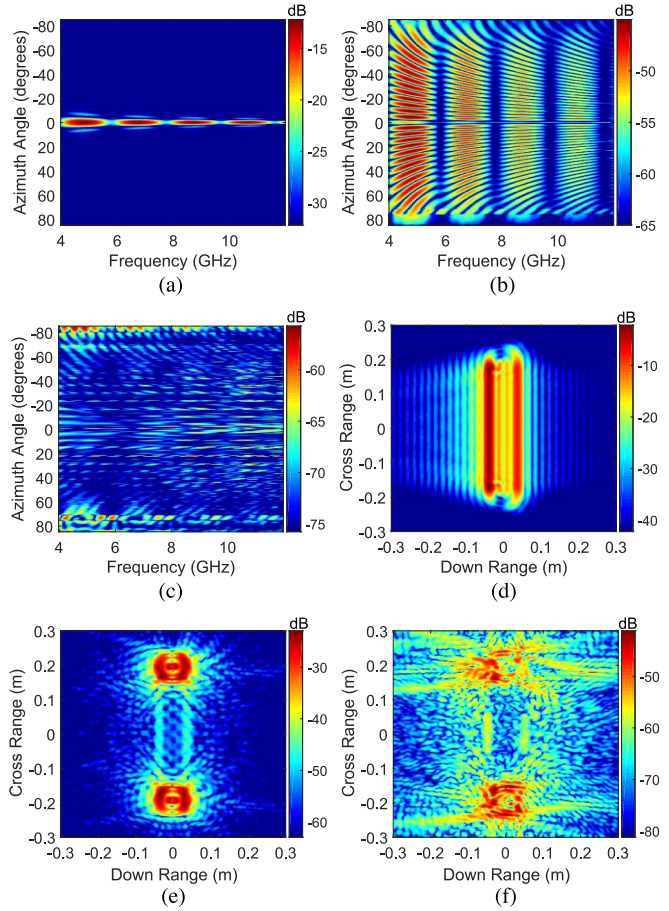


Fig. 9. Backscattering response and residual responses after inverse scattering for a rectangular plate, and their CBP images. (a) Backscattering response. (b) Residual response after two dominant components extracted. (c) Residual response after two dominant components and four weak components extracted. (d)–(f) CBP images of (a)–(c), respectively.

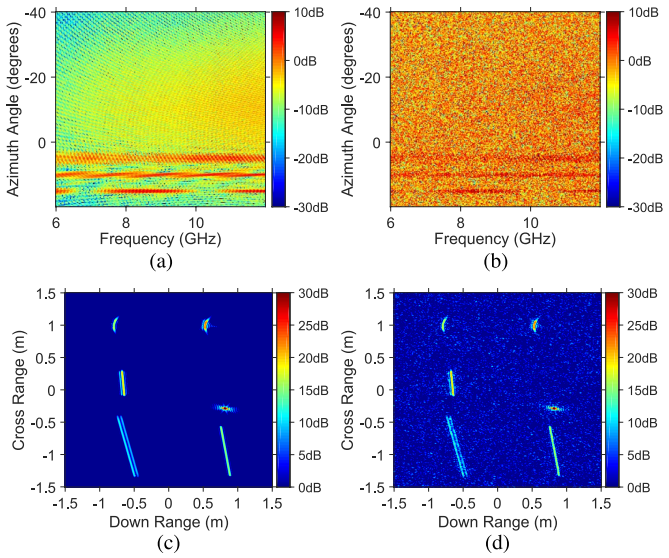


Fig. 10. The noisy frequency-aspect magnitude responses with (a) SNR $=$ 20 dB and (b) SNR $= -5$ dB, and their associated CBP images with (c) SNR $=$ 20 dB and (d) SNR $= -5$ dB.

components, which cannot be accurately and compactly described by IASM (see Fig. 6). For dihedral, the residual energy is generated by the sides of the plates (see Fig. 7). For trihedral, the residual energy is insignificant as comparing with the dominant energy, and it comes from two dihedrals, each of which consists of a vertical plate and the bottom plate (see Fig. 8). For rectangular plate, the residual components are four arc scattering responses, which verifies the imaging behavior of (16) (see Fig. 9). The efficient extraction for the residual scattering energy further demonstrates the effectiveness of IASM.

B. Inverse Scattering for a Simple Scene of Canonical Scatterers

Referring to the numerical experiments in [9], we create simple scenes of canonical scatterers to analyze the performance of the proposed model. The backscattering predictions are produced using CST software with the SBR technique. The radar frequency ranges from 6 to 12 GHz. The elevation angle θ is fixed at 15°, and the azimuth angle φ sweeps over 60° from $-40°$ to 20°.

First, we analyze the feature extraction with IASM at different SNRs. The scene contains all six canonical scatterers. Two kinds of Gaussian noise with different variances (related to SNRs) are separately added to the complex phase history data to produce two groups of noisy EM data, and their SNRs are 20 and -5 dB, respectively. The noisy frequency-aspect magnitude responses and their associated CBP images are shown in Fig. 10. The resultant parameter estimates through inverse scattering are listed in Table III. For each SNR, we perform 50 times of inverse scattering for different noisy data to estimate the means and the standard deviations. Obviously, the estimates of the geometrical parameters, including the radii $(\hat{R}_{\rm sph}, \hat{R}_{\rm top})$, the lengths $(\hat{L}_{\rm cyl}, \hat{L}_{\rm dih}, \hat{L}_{\rm pla})$, and the orientations $(\hat{\varphi}_{\rm cyl}, \hat{\varphi}_{\rm dih}, \hat{\varphi}_{\rm tri}, \hat{\varphi}_{\rm pla})$, are very close to the truth values. In particular, the -5-dB SNR does not cause great impacts on the estimation. By contrast, even when the EM data are free of noise, the estimates of several frequency dependence factors deviate appreciably from the GTD-inferred values, as well as the estimates in the inverse scattering experiment of individual canonical scatterers. This phenomenon is probably caused by the situation that the scatterers in a scene are affected by other scatterers around them. For example, the scattering response of the cylinder is partly distorted due to the existence of shadowing and diffraction effects, which lead to the estimation bias of λ_i. The estimations of a portion of aspect dependence factors have relatively large variances (see the bold numbers in Table III), which conforms to the estimations in the inverse scattering experiment of individual canonical scatterers. Nonetheless, in most cases, these biases bring no risks to the discrimination of the properties of the scatterers.

Second, we compare the feature extraction ability of IASM with that of ASM. Since ASM cannot efficiently describe arc scatterers, we consider a scene of only point- and line-segment-scatterers, including dihedral, trihedral, cylinder, and rectangular plate. The simulated scattering response is separately modeled by ASM and IASM. The original magnitude response, the residual magnitude responses, and their CBP images are shown in Fig. 11. We can conclude that IASM not only helps to capture more energy from the scene response but also provides a sparser description about the scene with less atoms than ASM.

Third, we compare IASM with the recently proposed model—physical and polynomial basis function (PPBF) [6]. PPBF confines to describe point scattering, and therefore, the trihedral scattering response is used to make a comparison. We model the scattering response in Fig. 8(a) with one PPBF atom. The reconstructed and residual scattering responses are shown in Fig. 12(a) and (b), respectively. PPBF introduces no aspect dependence into the model so that it cannot efficiently match the trihedral response. Although the frequency dependence factor can be accurately estimated, the residual energy is much greater than that obtained by IASM [see Fig. 8(b)]. The peak value of the IASM residual CBP image is 25 dB lower than that of the PPBF residual image. Thus, it is evident that PPBF is inadequate for modeling wide-angle scattering responses.

C. Inverse Scattering for a Realistic Target

The selected data for the inverse scattering experiment are the VV polarized scattering response of a backhoe with $0°$-elevation, $-50° \sim 50°$-azimuth aperture, and a bandwidth of 4 GHz. The data contain 346×1524 samples, i.e., $N = 346 \times 1524$. As discussed in Section III-D, the computational cost is proportional to N. We generate the candidate parameter set with all N samples, whereas to speed up the implementation, we perform the subsequent ISBL optimization on downsampled data which are obtained by randomly selecting $N/10$ samples with distinct indices from the data set. Hereinafter, we discuss the results from three aspects.

First, the ISBL method motivates a sufficient sparse solution. To avoid excessive iterations, we force to terminate the process when the number of atom reaches 70. The dominant energy is concentrated into these atoms (see Fig. 13). The original

TABLE III
MEANS AND STANDARD DEVIATIONS OF THE PARAMETER ESTIMATES AT DIFFERENT SNRS

parameter (in -ferred value)	noiseless	SNR=20dB		SNR=−5dB		parameter (truth value)	noiseless	SNR=20dB		SNR=−5dB	
		$\hat{m}$	$\hat{\sigma}\times 100$	$\hat{m}$	$\hat{\sigma}\times 10$			$\hat{m}$	$\hat{\sigma}\times 100$	$\hat{m}$	$\hat{\sigma}\times 10$
$\hat{\lambda}_{\text{sph}}\,(0)$	0.01	0.00	0.79	0.01	1.23	$\hat{R}_{\text{sph}}\,(0.2\text{m})$	0.20	0.20	0.01	0.20	0.02
$\hat{\lambda}_{\text{top}}\,(0.5)$	0.41	0.55	0.45	0.55	0.86	$\hat{R}_{\text{top}}\,(0.15\text{m})$	0.15	0.15	0.01	0.15	0.01
$\hat{\lambda}_{\text{cyl}}\,(0.5)$	0.64	0.64	0.50	0.65	0.83	$\hat{L}_{\text{cyl}}\,(0.8\text{m})$	0.80	0.80	0.01	0.80	0.05
$\hat{\lambda}_{\text{dih}}\,(1)$	0.97	0.85	0.42	0.86	0.84	$\hat{L}_{\text{dih}}\,(0.4\text{m})$	0.39	0.39	0.01	0.39	0.03
$\hat{\lambda}_{\text{tri}}\,(1)$	1.00	0.97	0.18	0.98	0.48	$\hat{L}_{\text{pla}}\,(1\text{m})$	1.00	1.00	0.06	0.99	0.13
$\hat{\lambda}_{\text{pla}}\,(0)$	-0.01	-0.01	0.81	0.03	1.22		1.00	1.00	0.16	0.99	0.14
	0.02	0.02	0.93	0.02	1.51	$\hat{\varphi}_{\text{cyl}}\,(10^\circ)$	10.04	10.04	0.07	10.04	0.10
$\hat{\tilde{\gamma}}_{\text{sph}}\,(0)$	0.00	0.01	2.03	0.16	3.52	$\hat{\varphi}_{\text{dih}}\,(5^\circ)$	5.00	5.00	0.12	5.00	0.20
$\hat{\tilde{\gamma}}_{\text{top}}\,(0)$	0.00	0.00	0.00	0.08	2.00	$\hat{\varphi}_{\text{tri}}\,(-10^\circ)$	-10.01	-10.00	1.19	-9.94	2.00
$\hat{\tilde{\gamma}}_{\text{cyl}}\,(0)$	0.00	0.00	0.00	**0.35**	**8.07**	$\hat{\varphi}_{\text{pla}}\,(15^\circ)$	14.99	15.00	0.07	15.00	0.16
$\hat{\tilde{\gamma}}_{\text{dih}}\,(0)$	**9.51**	**9.48**	**31.0**	**7.22**	**56.6**		14.99	15.00	0.09	15.00	0.13
$\hat{\tilde{\gamma}}_{\text{tri}}\,(1)$	**7.05**	**7.04**	**1.20**	**6.89**	**5.08**						
$\hat{\tilde{\gamma}}_{\text{pla}}\,(0)$	**2.00**	**2.02**	**6.31**	**0.22**	**6.13**						
	2.00	**2.01**	**4.20**	**0.67**	**9.24**						

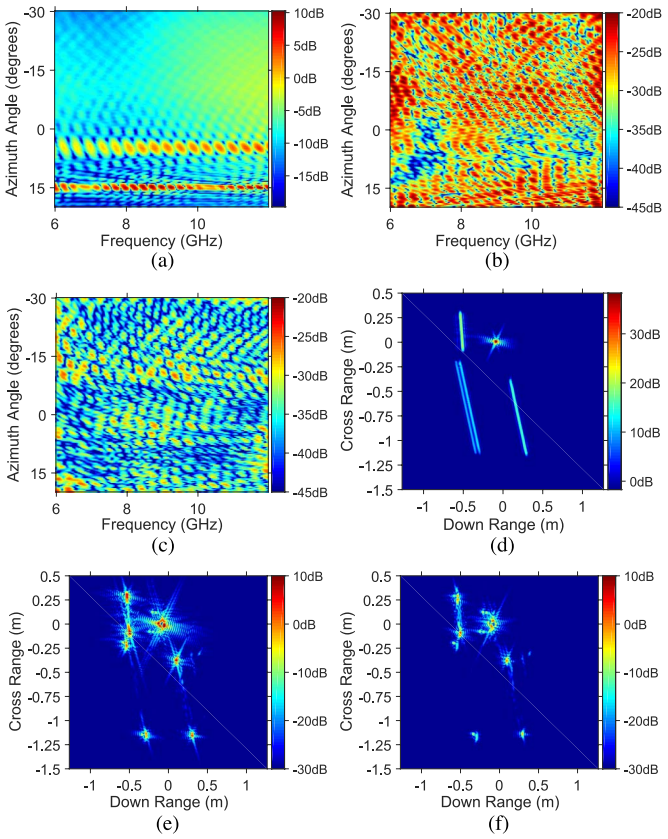


Fig. 11. Inverse scattering comparison between IASM and ASM. (a) Backscattering response of point- and line-segment-scattering reflectors, including one dihedral, one trihedral, one cylinder, and one rectangular plate. (b) and (c) Residual responses after the dominant energy extracted by 11 ASM atoms and 10 IASM atoms, respectively. (d)–(f) CBP images of (a)–(c), respectively.

scattering response, the residual response, and their CBP images are shown in Fig. 14. More importantly, the extracted features comprise an efficient description for the backhoe. The power of the original response is 26.35 dB, while the residual

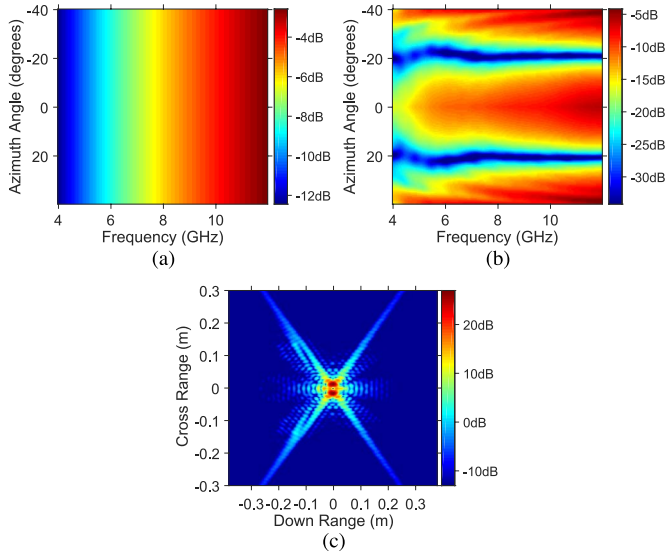


Fig. 12. Extracted and residual response after inverse scattering for a trihedral, and the residual CBP image. (a) Extracted response. (b) Residual response after dominant components extracted by PBF. (c) residual CBP image.

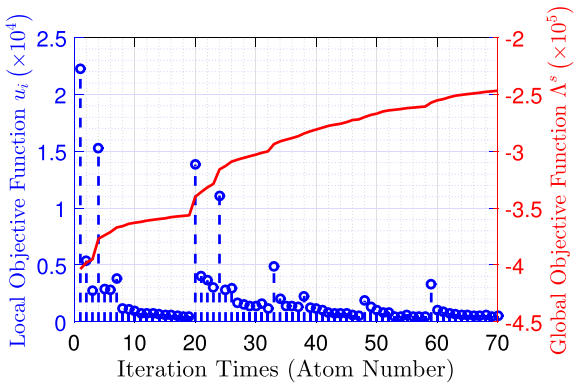


Fig. 13. Local and global objective functions versus iteration times or atom number.

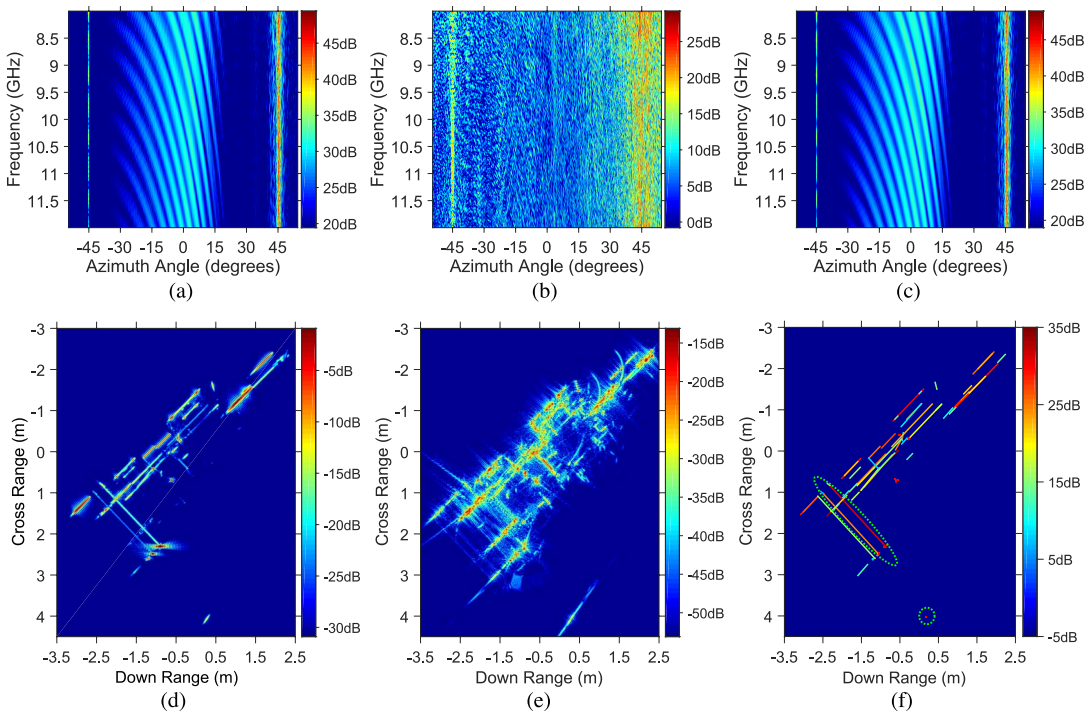


Fig. 14. Backscattering response and residual response after inverse scattering for the backhoe, and their CBP images. (a) Backscattering response. (b) Residual response after dominant components extracted by IASM. (c) Extracted response. (d) and (e) CBP images of (a) and (b), respectively. (f) Illustration of the extracted features, which all associate with some certain geometrical components, except those features circled by green dot lines.

response is 11.17 dB, and that is, more than 97% of energy is modeled by IASM. In fact, we can subsequently analyze the weak scatterers with the residual responses. Since the modeling of the dominant energy with IASM eliminates the major strong interferences on the weak scattering components, we can do the analysis for the weak scattering components more effectively. In addition, we compare our results in Fig. 14 with the "composite, point-enhanced" imaging result in [41]. Our result embodies many finer and clearer geometrical structures but discards some weaker scatterers and a few of the scattering responses from irregular geometrical structures. The point-enhanced method used in [41] is equivalent to calculating a sparse approximation with the overcomplete dictionary whose columns are generated by the ideal point-scattering model. Essentially, this point model cannot motivate a sufficient sparse solution as our model. Furthermore, the employed optimization method in [41] usually encounters a high computational burden.

Second, the extracted features by IASM are useful for shape interpretation. As depicted in Fig. 15, most features accurately reflect some geometrical structures. In particular, many indistinguishable structures in the CBP image have been resolved into true geometrical structures. However, three edges inside the dot ellipses and circles have no corresponding structures. It is believed that the blade attached in front of the backhoe has a special groove that will cause multiple scattering. This additional scattering may lead to the artificial scatterers outside the blade.

Third, the computational cost for ISBL is affordable. The personal computer used in the experiment has an Intel E5 processor with 2.10-GHz main frequency and 16.00-GB memory. The BFGS iteration times, the function evaluation times, and the running time taken for each ISBL iteration are individually recorded to evaluate the complexity with the conclusion in Section III-D. As shown in Fig. 16, the evaluated complexity and the real running time have similar change trends. In fact, if the parallel computation is applied for the function evaluation, the computational efficiency can be further greatly improved.

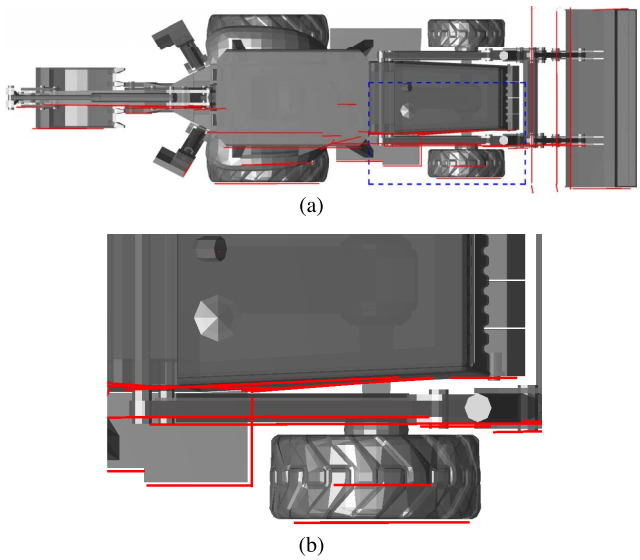


Fig. 15. Mapping the extracted 2-D features (painted red) by IASM into the 3-D geometrical components (painted gray). (a) Full view. (b) Partial enlarged view of the partitioned region in (a) with blue dashed lines.

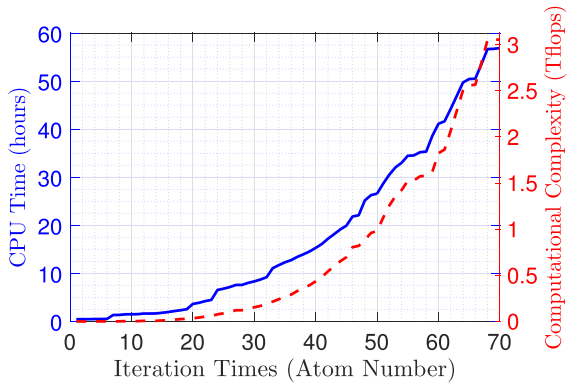


Fig. 16. Running time and computational complexity versus iteration times or atom number.

V. Conclusion

In this paper, we have proposed an improved ASM to mathematically unify the benchmarking canonical scattering models. The developed model utilizes a cosine function to replace the original aspect dependence function, introduces a physical parameter to describe the migrating displacement phenomenon, and provides an approximate rectangular window to characterize the shadowing effects. In addition, the complex-valued ISBL has been employed to perform sparsity-driven inverse scattering with IASM at an affordable computational cost. The numerical experiments were carried out on the simulated canonical scattering responses and the public-released wide-angle SAR data of a backhoe. The results have revealed that IASM has excellent compatibility with not only canonical primitives but also realistic targets, and the extracted features can provide accurate geometrical information about the target. The GTD-inferred values of canonical primitives have been experimentally evaluated by the proposed methodology, and several discrepancies have been initially discussed based on the experimental results.

Future works will focus on twofold: first, considering that IASM is merely appropriate for point-, line-segment-, and arc scatterers, there is still a need for research on more general models for arbitrary curve- and polyline-scatterers; second, the quasi-Newton iterative algorithm employed in ISBL optimization cannot guarantee global optimal solutions, so we will further study on other global optimization algorithms in future.

Appendix A
Partial Derivatives of the Local and Global Objective Functions on the Parameters

According to [42], the partial derivatives of the local and global objective functions on the parameters can be given by

$$\frac{\partial u_i}{\partial \boldsymbol{\vartheta}_i^{\mathrm{T}}} = \frac{2}{s_i}\Re\left(q_i\left(\mathbf{t}^{\mathrm{H}}\mathbf{C}_{-i}^{-1}\frac{\partial \boldsymbol{\phi}_i}{\partial \boldsymbol{\vartheta}_i^{\mathrm{T}}}\right)\right) - \frac{2|q_i|^2}{s_i^2}\Re\left(\boldsymbol{\phi}_i^{\mathrm{H}}\mathbf{C}_{-i}^{-1}\frac{\partial \boldsymbol{\phi}_i}{\partial \boldsymbol{\vartheta}_i^{\mathrm{T}}}\right) \tag{A1}$$

$$\frac{\partial \Lambda^s}{\partial \boldsymbol{\vartheta}_i^{\mathrm{T}}} = 2\Re\left(\frac{\partial \Lambda^s}{\partial \boldsymbol{\phi}_i^{\mathrm{T}}}\frac{\partial \boldsymbol{\phi}_i}{\partial \boldsymbol{\vartheta}_i^{\mathrm{T}}}\right)$$

$$\frac{\partial \Lambda^s}{\partial \boldsymbol{\phi}_i} = \frac{q_i}{(\alpha_i+s_i)}\left(\mathbf{t}^{\mathrm{H}}\mathbf{C}_{-i}^{-1}\right)^{\mathrm{T}} - \left[\frac{1}{\alpha_i+s_i} + \frac{|q_i|^2+\rho\alpha_i}{(\alpha_i+s_i)^2}\right]\left(\boldsymbol{\phi}_i^{\mathrm{H}}\mathbf{C}_{-i}^{-1}\right)^{\mathrm{T}} \tag{A2}$$

where $\Re(\,)$ denotes the real part of a complex-valued variable. To greatly reduce the computational complexity of matrix inversion in (A1), we derive the following formulas:

$$\begin{aligned}\boldsymbol{\phi}_i^{\mathrm{H}}\mathbf{C}_{-i}^{-1}\frac{\partial \boldsymbol{\phi}_i}{\partial \boldsymbol{\vartheta}_i^{\mathrm{T}}} &= \frac{\alpha_i+s_i}{\alpha_i}\boldsymbol{\phi}_i^{\mathrm{H}}\mathbf{C}^{-1}\frac{\partial \boldsymbol{\phi}_i}{\partial \boldsymbol{\vartheta}_i^{\mathrm{T}}} \\ &= \frac{\alpha_i+s_i}{\alpha_i}\left\{\beta\boldsymbol{\phi}_i^{\mathrm{H}}\frac{\partial \boldsymbol{\phi}_i}{\partial \boldsymbol{\vartheta}_i^{\mathrm{T}}} - \beta^2\left(\boldsymbol{\Phi}^{\mathrm{H}}\boldsymbol{\phi}_i\right)^{\mathrm{H}}\mathbf{C}_{\mathbf{w}}\left(\boldsymbol{\Phi}^{\mathrm{H}}\frac{\partial \boldsymbol{\phi}_i}{\partial \boldsymbol{\vartheta}_i^{\mathrm{T}}}\right)\right\}\end{aligned}$$

$$\begin{aligned}\mathbf{t}^{\mathrm{H}}\mathbf{C}_{-i}^{-1}\frac{\partial \boldsymbol{\phi}_i}{\partial \boldsymbol{\vartheta}_i^{\mathrm{T}}} &= \beta\mathbf{t}^{\mathrm{H}}\frac{\partial \boldsymbol{\phi}_i}{\partial \boldsymbol{\vartheta}_i^{\mathrm{T}}} - \beta^2\left(\boldsymbol{\Phi}^{\mathrm{H}}\mathbf{t}\right)^{\mathrm{H}}\mathbf{C}_{\mathbf{w}}\left(\boldsymbol{\Phi}^{\mathrm{H}}\frac{\partial \boldsymbol{\phi}_i}{\partial \boldsymbol{\vartheta}_i^{\mathrm{T}}}\right) \\ &\quad + \frac{q_i^*}{\alpha_i+s_i}\left(\boldsymbol{\phi}_i^{\mathrm{H}}\mathbf{C}_{-i}^{-1}\frac{\partial \boldsymbol{\phi}_i}{\partial \boldsymbol{\vartheta}_i^{\mathrm{T}}}\right)\end{aligned} \tag{A3}$$

where the computation formulas for s_i and q_i are given in [23].

Acknowledgment

The authors would like to thank the Air Force Research Laboratory for free use of public released data "Backhoe Data Dome" and the reviewers for their helpful comments.

References

[1] J. B. Keller, "Geometrical theory of diffraction," *J. Opt. Soc. Amer.*, vol. 52, no. 2, pp. 116–130, Jan. 1962.
[2] L. C. Potter, D. M. Chiang, R. Carrière, and M. J. Gerry, "A GTD-based parametric model for radar scattering," *IEEE Trans. Antennas Propag.*, vol. 43, no. 10, pp. 1058–1067, Oct. 1995.
[3] N. Akhter, "Far Zone Electromagnetic Scattering From Complex Shapes Using Geometrical Theory of Diffraction," Ph.D. dissertation, Ohio State Univ., Columbus, OH, USA, 1993.
[4] L. C. Potter and R. L. Moses, "Attributed scattering centers for SAR ATR," *IEEE Trans. Image Process.*, vol. 6, no. 1, pp. 79–91, Jan. 1997.
[5] M. J. Gerry, L. C. Potter, I. J. Gupta, and A. v. d.Merwe, "A parametric model for synthetic aperture radar measurements," *IEEE Trans. Antennas Propag.*, vol. 47, no. 7, pp. 1179–1188, Jul. 1999.
[6] J. Halman and R. J. Burkholder, "Sparse expansions using physical and polynomial basis functions for compressed sensing of frequency domain EM scattering," *IEEE Antennas Wireless Propag. Lett.*, vol. 14, pp. 1048–1051, May 2015.
[7] M. Koets, "Automated Algorithms for Extraction of Physically Relevant Features From Synthetic Aperture Radar Imagery," M.S. thesis, Ohio State Univ., Columbus, OH, USA, 2000.
[8] Y. Akyildiz, "Feature Extraction From Synthetic Aperture Radar Imagery," M.S. thesis, Ohio State Univ., Columbus, OH, USA, 2000.
[9] J. A. Jackson and R. L. Moses, "Feature extraction algorithm for 3D scene modeling and visualization using monostatic SAR," in *Proc. SPIE Algorithms Synthetic Aperture Radar Imagery XIII*, E. G. Zelnio and F. D. Garber Eds., 2006, vol. 6237, pp. 1–12.
[10] H. Chiang, R. L. Moses, and L. C. Potter, "Model-based classification of radar images," *IEEE Trans. Inf. Theory*, vol. 46, no. 5, pp. 1842–1854, Aug. 2000.
[11] J. Duan, L. Zhang, M. Xing, Y. Wu, and M. Wu, "Polarimetric target decomposition based on attributed scattering center model for synthetic aperture radar targets," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 12, pp. 2095–2099, Dec. 2014.
[12] E. Ertin and R. L. Moses, "Through-the-wall SAR attributed scattering center feature estimation," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 5, pp. 1338–1348, May 2009.
[13] D. F. Fuller and M. A. Saville, "A high-frequency multipeak model for wide-angle SAR imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 7, pp. 4279–4291, Jul. 2013.
[14] R. Xi, "Super resolution processing of SAR images by matching pursuit method based on genetic algorithm," in *Proc. 3rd Int. Congr. Image Signal Process.*, Yantai, China, Oct. 2010, pp. 2066–2070.

[15] M. Wu, M. Xing, L. Zhang, J. Duan, and G. Xu, "Super-resolution imaging algorithm based on attributed scattering center model," in *Proc. IEEE China Summit Int. Conf. Signal Inf. Process.*, Xi'an, China, Jul. 2014, pp. 271–275.

[16] J. A. Tropp and S. J. Wright, "Computational methods for sparse solution of linear inverse problems," *Proc. IEEE*, vol. 98, no. 6, pp. 948–958, Jun. 2010.

[17] M. E. Tipping, "Sparse Bayesian learning and the relevance vector machine," *J. Mach. Learn. Res.*, vol. 1, pp. 211–244, 2001.

[18] C. M. Bishop, *Pattern Recognition and Machine Learning*. New York, NY, USA: Springer-Verlag, Aug. 2006.

[19] B. Krishnapuram, A. J. Hartemink, and M. A. T. Figueiredo, "A Bayesian approach to joint feature selection and classifier design," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 9, pp. 1105–1111, Sep. 2004.

[20] A. Schmolck and R. Everson, "Smooth relevance vector machine: A smoothness prior extension of the RVM," *Mach. Learn.*, vol. 68, no. 2, pp. 107–135, Aug. 2007.

[21] A. C. Faul and M. E. Tipping, "Analysis of sparse Bayesian learning," in *Advances in Neural Information Processing Systems*. Cambridge, MA, USA: MIT Press, 2001, pp. 383–389.

[22] H. Liu, B. Jiu, H. Liu, and Z. Bao, "Superresolution ISAR imaging based on sparse Bayesian learning," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 8, pp. 5005–5013, Aug. 2014.

[23] D. G. Tzikas, A. C. Likas, and N. P. Galatsanos, "Sparse Bayesian modeling with adaptive kernel learning," *IEEE Trans. Neural Netw.*, vol. 20, no. 6, pp. 926–937, Jun. 2009.

[24] S. Lee and M. Eskenazi, "Incremental sparse Bayesian method for online dialog strategy learning," *IEEE J. Sel. Topics Signal Process.*, vol. 6, no. 8, pp. 903–916, Dec. 2012.

[25] D. Shutin, W. Wang, and T. Jost, "Incremental sparse Bayesian learning for parameter estimation of superimposed signals," in *Proc. 10th Int. Conf. SampTA*, Jul. 2013, pp. 513–516.

[26] T. L. Hansen, M. A. Badiu, B. H. Fleury, and B. D. Rao, "A sparse Bayesian learning algorithm with dictionary parameter estimation," in *Proc. IEEE 8th SAM Signal Process. Workshop*, 2014, pp. 385–388.

[27] J. A. Jackson, B. D. Rigling, and R. L. Moses, "Parametric scattering models for bistatic synthetic aperture radar," in *Proc. IEEE Radar Conf.*, May 2008, pp. 1–5.

[28] J. A. Jackson, B. D. Rigling, and R. L. Moses, "Canonical scattering feature models for 3D and bistatic SAR," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 46, no. 2, pp. 525–541, Apr. 2010.

[29] Logistic Function. [Online]. Available: http://en.wikipedia.org/wiki/Logistic_function#Double_logistic_function

[30] R. G. Gallager, Circularly-Symmetric Gaussian Random Vectors, 2008. [Online]. Available: http://www.rle.mit.edu/rgallager/documents/CircSymGauss.pdf

[31] M. Tipping and A. Faul, "Fast marginal likelihood maximisation for sparse Bayesian models," in *Proc. 9th Int. Workshop Artif. Intell. Stat.*, Key West, FL, USA, 2003, pp. 3–6.

[32] L. A. Gorham and L. J. Moore, "SAR image formation toolbox for MATLAB," in *Proc. SPIE Algorithms Synthetic Aperture Radar Imagery XVII*, Apr. 2010, vol. 7699, pp. 1–13.

[33] Computational Complexity of Mathematical Operations. [Online]. Available: http://en.wikipedia.org/wiki/Computational_complexity_of_mathematical_operations#cite_note-Schonhage-1

[34] J. E. Dennis, Jr. and J. J. Moré, "Quasi-Newton methods, motivation and theory," *SIAM Rev.*, vol. 19, no. 1, pp. 46–89, Jan. 1977.

[35] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-9, no. 1, pp. 62–66, Jan. 1979.

[36] N. L. Johnson, A. W. Kemp, and S. Kotz, *Univariate Discrete Distributions*, 3rd ed. Hoboken, NJ, USA: Wiley, 2005.

[37] P. Soille, *Morphological Image Analysis: Principles and Applications*, 2nd ed. Berlin, Germany: Springer-Verlag, 2004.

[38] R. C. Gonzalez, R. E. Woods, and S. L. Eddins, *Digital Image Processing Using MATLAB*. Upper Saddle River, NJ, USA: Prentice-Hall, 2003.

[39] H. K. Yuen, J. Princen, J. Illingworth, and J. Kittler, "Comparative study of Hough transform methods for circle finding," *Image Vis. Comput.*, vol. 8, no. 1, pp. 71–77, Feb. 1990.

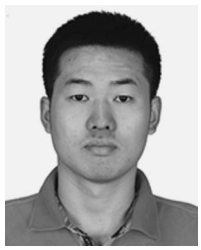
[40] CST Microwave Studio Technique Specification, 2014. [Online]. Available: https://www.cst.com/Products/CSTMWS/DownloadPDF?specificationId=80

[41] M. Çetin and R. L. Moses, "SAR imaging from partial-aperture data with frequency-band omissions," in *Proc. SPIE Defense Security Symp., Algorithms Synthetic Aperture Radar Imagery XII*, Orlando, FL, USA, Mar. 2005, vol. 5808, pp. 1–12.

[42] A. Hjørungnes and D. Gesbert, "Complex-valued matrix differentiation: Techniques and key results," *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 2740–2746, Jun. 2007.

Zenghui Li (S'15) received the B.S. degree in electronic and engineering from Beihang University, Beijing, China, in 2005 and the M.S. degree in information and communication technology from the National University of Defense Technology, Changsha, China, in 2007. He is currently working toward the Ph.D. degree in the Department of Electronic Engineering, Tsinghua University, Beijing.

His research work focuses on polarimetric inverse scattering.

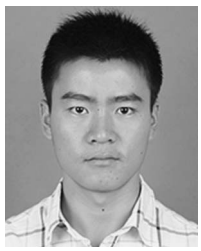


Kan Jin received the B.S. and M.S. degrees from Tsinghua University, Beijing, China, in 2011 and 2014, respectively, where he is currently working toward the Ph.D. degree in the Department of Electronic Engineering.

His research work focuses on electromagnetic scattering computation and polarimetric SAR image processing.

Bin Xu received the B.S. degree from Tsinghua University, Beijing, China, in 2011, where he is currently working toward the Ph.D. degree in the Department of Electronic Engineering.

His research work focuses on SAR image processing and polarimetric SAR image processing.



Wei Zhou received the B.S. degree from Tsinghua University, Beijing, China, in 2010, where he is currently working toward the Ph.D. degree in the Department of Electronic Engineering.

His research work focuses on radar target imaging and recognition, target motion compensation, and radar signal processing.

Jian Yang (M'98–SM'02) received the B.S. and M.S. degrees from Northwestern Polytechnical University, Xian, China, in 1985 and 1990, respectively, and the Ph.D. degree from Niigata University, Niigata, Japan, in 1999.

In 1985, he was with the Department of Applied Mathematics, Northwestern Polytechnical University. From 1999 to 2000, he was an Assistant Professor with Niigata University. In April 2000, he was with the Department of Electronic Engineering, Tsinghua University, Beijing, China, where he is currently a Professor. His interests include radar polarimetry, remote sensing, mathematical modeling, optimization in engineering, and fuzzy theory.

Dr. Yang is an Associate Editor of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, the Vice Chairman of the IEEE Aerospace and Electronic Systems Beijing chapter, and the former Chairman of the Institute of Electronics, Information and Communication Engineers in the Beijing area.