Research Article

Sparse Bayesian Perspective for Radar Coincidence Imaging with Model Errors

Bo Fan [ID],[1] **Xiaoli Zhou** [ID],[2] **Shuo Chen,**[1] **Zhijie Jiang,**[1] and **Yongqiang Cheng** [ID][2]

[1]*National Innovation Institute of Defense Technology, Academy of Military Science, Beijing 100071, China*
[2]*College of Electronic Science and Technology, National University of Defense Technology, Changsha 410073, China*

Correspondence should be addressed to Yongqiang Cheng; nudtyqcheng@gmail.com

Sparsity-driven methods are commonly applied to reconstruct targets in radar coincidence imaging (RCI), where the reference matrix needs to be computed precisely and the prior knowledge of the accurate imaging model is essential. Unfortunately, the existence of model errors in practical RCI applications is common, which defocuses the reconstructed image considerably. Accordingly, this paper aims to formulate a unified framework for sparsity-driven RCI with model errors based on the sparse Bayesian approach. Firstly, a parametric joint sparse reconstruction model is built to describe the RCI when perturbed by model errors. The structured sparse Bayesian prior is then assigned to this model, after which the structured sparse Bayesian autofocus (SSBA) algorithm is proposed in the variational Bayesian expectation maximization (VBEM) framework; this solution jointly realizes sparse imaging and model error calibration. Simulation results demonstrate that the proposed algorithm can both calibrate the model errors and obtain a well-focused target image with high reconstruction accuracy.

1. Introduction

Imaging radar is an outstanding form of remote sensing equipment with many advantages, including its robust performance under all-weather and all-day circumstances, long-distance capabilities, and high probability of target identification [1]. Generally speaking, there are two main types of imaging radars, which are defined in terms of the validity of the aperture [2], namely, real aperture radar (RAR) [3] and virtual synthetic aperture (SA) radar [4–7]. While RAR can realize staring imaging for targets or regions of interest, the azimuth resolution is limited by the practical antenna aperture; therefore, it is difficult to apply RAR to high-resolution imaging applications. For its part, the virtual synthetic aperture radar (SAR) [8, 9], which encompasses SAR and inverse SAR (ISAR), exhibits high-resolution ability based on the range-Doppler (RD) principle, where the high resolution generally depends on large signal bandwidth and angle variation. However, the dependence on the target's relative motion leads to difficulties in motion compensation, forward-looking, or staring imaging.

Based on optical coincidence imaging, radar coincidence imaging (RCI) [1, 10, 11] is a novel high-resolution staring imaging technique that can operate under forward-looking or staring observing geometry and has significant potential for resolution enhancement, instantaneous imaging, interference, and jamming suppression. RCI forms a temporal-spatial stochastic radiation field by random-modulating the wavefront of the detecting signals. Preliminary theoretical analysis shows that stochastic radiation can lead to super-resolution capability, i.e., breaking through the Rayleigh resolution limit of the antenna [12]. Accordingly, RCI can be regarded as a complement of conventional imaging methods such as RAR and SAR/ISAR and can further be employed in some important applications including high-resolution Earth observation, oceanic monitoring, and military reconnaissance.

In theory, RCI reconstructs the target image by means of coincidence processing between the target scattering echo and the stochastic radiation field. Thus, the reference matrix, which characterizes the radiation field, needs to be computed precisely according to the imaging model and preset

parameters. However, a variety of model errors (e.g., gain-phase error [13], off-grid error [14], array position error [15], and target motion-induced error [16, 17]) exist in most application scenarios. In the presence of these errors, the reference matrix is severely perturbed, meaning that the reconstructed image will be considerably defocused.

Recently, some attention has been paid to RCI with certain types of model error [13, 17], and corresponding autofocus imaging algorithms have been proposed to jointly calibrate the model error and reconstruct the image. It is worth noting that autofocus techniques have been investigated widely for SAR/ISAR imaging [5, 7] in a sparse reconstruction or compressive sensing (CS) framework [18, 19]. However, while the abovementioned studies only focused on a certain type of model error, many types of model error tend to coexist in reality; hence, conventional algorithms cannot be directly applied in practical RCI applications, and an autofocus technique is required to calibrate the model error by considering the universal reference matrix perturbation. To the best of the author's knowledge, few studies have focused on RCI with fully perturbed reference matrix. One recently published paper [20] proposed a completely perturbed model that models the effect of model error on the reference matrix using a perturbation matrix; accordingly, the regularization-focal underdetermined system solver (R-FOCUSS) was proposed, which is based on a Bayesian framework and regularization method.

By contrast, RCI can also be modeled as a sparse reconstruction [21–23] or CS problem by exploiting the sparse prior of the target since RCI can be formed as a linear inverse problem [1, 24] where the reference matrix is a random matrix. Thus, RCI with model error can be regarded as a perturbed CS problem where the sensing matrix is completely perturbed. Herman and Strohmer analyzed the effect of perturbation of the sensing matrix on the performance of the basis pursuit algorithm and found the condition of stable recovery [25]. Moreover, based on the total least squares (TLS) model, Zhu et al. devised the sparse-TLS (S-TLS) method via the iterative block coordinate descent method for perturbed CS [26] to cope with sparse and underdetermined "errors-in-variables" models. S-TLS was then applied to solve the direction of arrival (DOA) estimation with model errors. However, as S-TLS is inefficient and time-consuming, two faster and more robust algorithms, namely, TLS-FOCUSS and SD-FOCUSS, were proposed [27]. A unified framework for the DOA estimation with array imperfections was formulated based on the Bayesian approach, and the sparse Bayesian array calibration (SBAC) method was proposed to realize array calibration and DOA estimation, resorting to the expectation-maximization (EM) algorithm for an iterative solution [28].

However, the previous work mostly focused on a certain type of model error, such as grid mismatch, amplitude, or phase error. The actual situation is that multiple model errors often exist simultaneously. Therefore, this paper aims to propose a unified framework for multiple model errors, while the previous imaging algorithms are not effective.

This paper focuses on the sparsity-driven RCI with model errors and proposes a unified framework to address

the RCI in the presence of reference matrix perturbation, which is induced by radar system error and off-grid error simultaneously. To characterize the perturbation matrix, Taylor expansion is utilized to approximate the actual reference matrix perturbed by the off-grid error; in this way, the joint sparse reconstruction model is built. For radar system error, the parametric sparse representation model is established based on the parametric reference matrix. Accordingly, the parametric joint sparse reconstruction model combines the two models to uniformly characterize RCI with model errors. The merit of SBL is its flexibility in modeling sparse signals that can not only promote the sparsity but also exploit the possible structure of the signal to be recovered. Furthermore, SBL still presents good performance, while the wave-front holds bad stochastic property. Using Bayesian hierarchical modeling, the sparse coefficients are assigned to an appropriate structured sparse prior. An sparse Bayesian learning- (SBL-) based method named the structured sparse Bayesian autofocus (SSBA) algorithm within the variational Bayesian expectation maximization (VBEM) framework [29–31] is then proposed to realize sparse imaging and model error calibration jointly. Finally, a scenario in which off-grid error and gain-phase error coexist is taken as an example in order to demonstrate that the SSBA method can be extended to calibrate more than one type of model error and realize sparse imaging simultaneously. Numerical simulations show that the algorithm realizes the imaging robustly and achieves high resolution and outstanding imaging quality in the presence of model errors. This paper proposes a unified framework to solve the imaging problems under various model errors. In conclusion, the proposed method can be applied to various types of model errors, especially when multiple model errors exist simultaneously.

The rest of this paper is organized as follows. Section 2 presents the unified parametric joint sparse reconstruction model for RCI with model errors. In Section 3, the proposed structured sparse Bayesian model and SSBA algorithm are outlined. Section 4 conducts several numerical simulations involving the proposed algorithm. Finally, conclusions are drawn, and future research directions are provided in Section 5.

Notations $(\cdot)^H$ and $(\cdot)^{-1}$ denote the conjugate transpose and inverse operations, respectively. $\|\cdot\|$ is the norm operator, $\langle\cdot\rangle$ denotes the posterior expectation operator, and $\odot$ is the Hadamard product.

2. Problem Formulation

2.1. Imaging Model of RCI. The principle of RCI is depicted in Figure 1. Consider an RCI system configured as a monostatic array, where the transmitters emit independent stochastic waveforms [1, 11]. The RCI system has M transmitters and a single receiver. The random-modulated wavefront shows spatial stochastic diversities to resolve the targets within the beam and presents temporal stochastic diversities to decouple and resolve the targets with different ranges. The temporal-spatial stochastic radiation field is constructed by modulating the wavefront of the detecting
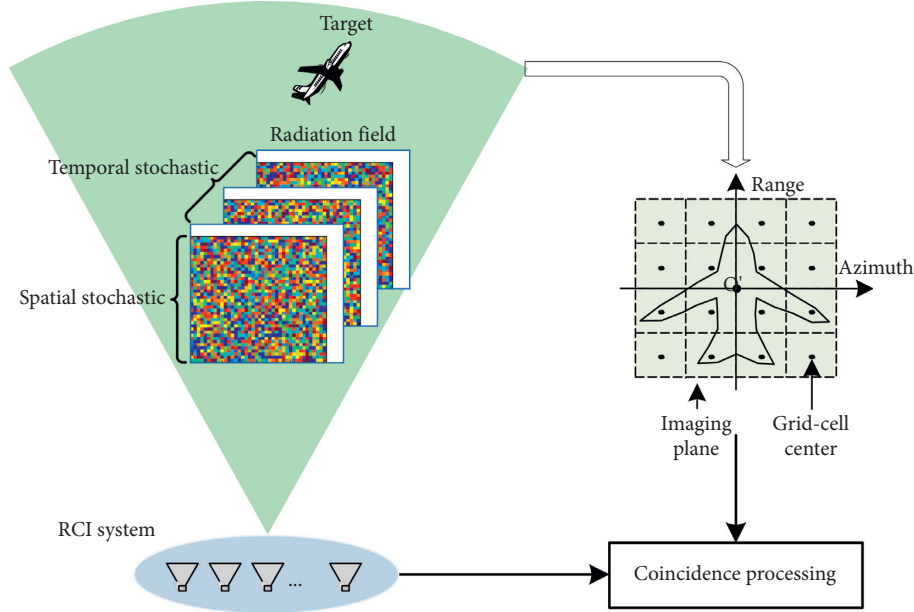
FIGURE 1: Principle of radar coincidence imaging.

signals. The imaging performance depends on the stochastic characteristics of the radiation field. Intuitively speaking, a radiation field with a higher stochastic degree could improve the stability and performance of RCI, exhibiting superior super-resolution potential. The characteristics of the stochastic radiation fields are not only related to the internal physical configurations of the antenna (array type, number, and spacing of array elements) but also to the system-transmitted waveform (waveform type, central frequency, and operation bandwidth). Besides, the detailed coincidence processing has been introduced in [1].

The received scattering echo can be described using the following simplified form:

$$y(t) = \int_A \beta(\mathbf{r}_k) \cdot S(t, r_k) \mathrm{d}k + w(t), \quad (1)$$

where A is the imaging area, $\beta(\mathbf{r}_k)$ is the backscattering coefficient of the target at location $\mathbf{r}_k$, and $w(t)$ denotes the measurement noise. $S(t, \mathbf{r}_k)$ is referred to as the reference signal (or detecting signal) [1], which is defined as

$$S(t, \mathbf{r}_k) = \sum_{m=1}^{M} St_m(t - \tau_m(\mathbf{r}_k)), \quad (2)$$

where M denotes the number of transmitters and $St_m(t)$ is the transmitted waveform for the m th transmitter. $\tau_m(\mathbf{r}_k)$ is the propagation delay corresponding to the m th transmitter and the receiver with respect to the k th scatterer, which is defined as

$$\tau_m(\mathbf{r}_k) = \frac{1}{c}\left(\|\mathbf{R}_m - \mathbf{r}_k\| + \|\mathbf{R}_0 - \mathbf{r}_k\|\right), \quad (3)$$

where $\mathbf{r}_k$, $\mathbf{R}_0$, and $\mathbf{R}_m$ are the position vectors of the k th scatterer, the receiver, and the m th transmitter, respectively, and c is the speed of light in free space.

In practical coincidence processing, the received echo $y(t)$ should be discretized in both space and time domains. Thus, the range-azimuth imaging area is divided into K_1 azimuth cells, K_2 range cells, and associated cell size Δ_1 and Δ_2. Accordingly, the number of grid-cells is $K_1 \times K_2 = K$, and the received signal is sampled to generate N samples, such that the imaging equation can then be given by

$$\mathbf{y} = \mathbf{S} \cdot \boldsymbol{\beta} + \mathbf{w},$$

$$\begin{bmatrix} y(t_1) \\ y(t_2) \\ \vdots \\ y(t_N) \end{bmatrix} = \begin{bmatrix} S(t_1, \mathbf{r}_1) & S(t_1, \mathbf{r}_2) & \cdots & S(t_1, \mathbf{r}_K) \\ S(t_2, \mathbf{r}_1) & S(t_2, \mathbf{r}_2) & \cdots & S(t_2, \mathbf{r}_K) \\ \vdots & \vdots & \cdots & \vdots \\ S(t_N, \mathbf{r}_1) & S(t_N, \mathbf{r}_2) & \cdots & S(t_N, \mathbf{r}_K) \end{bmatrix} \cdot \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_K \end{bmatrix}$$

$$+ \begin{bmatrix} w(t_1) \\ w(t_2) \\ \vdots \\ w(t_N) \end{bmatrix}, \quad (4)$$

where $\mathbf{S}$ is the reference matrix [1], while $\mathbf{y}$, $\mathbf{w}$, and $\boldsymbol{\beta}$ are the echo, noise, and unknown scattering coefficient vectors, respectively. Obviously, $\mathbf{S}$ is a random matrix that is used commonly in CS theory as the sensing matrix due to its improved statistical features and low implementation complexity. Thus, the imaging model reduces to a familiar linear model problem that estimates $\boldsymbol{\beta}$ from reference matrix $\mathbf{S}$ and the echo vector $\mathbf{y}$. The echo vector $\mathbf{y}$ can be obtained using A/D sampling, while the reference matrix $\mathbf{S}$ is deduced through the radiation field calculation.

2.2. Unified Model of RCI with Model Errors. In the derivation of the reference matrix, all corresponding parameters,

including the transmitted waveform $St_m(t)$, the array position ($\mathbf{R}_m$ and $\mathbf{R}_0$), and the grid-cell centers $\mathbf{r}_k$, should be known accurately. Moreover, the scatterers should be located exactly at the grid-cell centers. As noted above, model errors (such as gain-phase error [13], off-grid error [14], and array position error [15]) are known to exist in practical RCI applications. Thus, the reference matrix is inevitably perturbed by the unknown model error; accordingly, the actual reference matrix can be expressed as

$$\widehat{S} = \mathbf{S} + \mathbf{E}, \tag{5}$$

where $\mathbf{E} \in \mathbb{C}^{N \times K}$ denotes the perturbation matrix induced by model error. In the presence of model error, the imaging processing will be affected; in cases of large model error, the target image will be severely defocused.

The fully perturbed reference matrix shown in (5) models all types of model errors using the perturbation matrix, which is simple. However, there is insufficient structured information between the reference matrix and the model errors. Fully exploiting the structured information could lead to improved performance. The reference matrix for RCI with gain-phase error is represented as a parametric matrix [13], while the off-grid RCI is modeled using Taylor expansion [14] to utilize the structured information. Since the structured dependences between $\widehat{S}$ and the different model errors differ, building the structured relationship for all different types of model errors would increase the computational complexity to an unacceptable degree. Hence, we classify the set of model errors Θ into two types:

(a) *Type I*: *Radar System Error*. Radar system error Θ results from uncertainty in the radar system, e.g., gain-phase error and array position error. The reference matrix involving radar system error is generally modeled as a parametric matrix, i.e., $\mathbf{S}(\Theta)$. Accordingly, the sparse RCI is translated to a parametric sparse representation problem [32], which is a special case of dictionary learning [33]; thus, it can dynamically learn the unknown factors during imaging and achieve the optimally sparse representation of radar signals.

(b) *Type II*: *Off-Grid Error*. Dividing the imaging area into several discrete grid cells results in the off-grid error. The imaging model (4) is formulated based on the assumption that the scatterers are located precisely at these prediscretized grid-cell centers [1]. However, the scatterers are in fact distributed in a continuous scene and generally located off the grid-cell centers; therefore, the off-grid error emerges. Moreover, the relative motion between target and radar also results in off-grid error. The off-grid-error-induced reference matrix is generally approximated using first-order Taylor expansion, such that the off-grid imaging model is

$$\mathbf{y} \approx \mathbf{S}\boldsymbol{\beta} + \mathbf{S}_1\boldsymbol{\beta}_1 + \mathbf{S}_2\boldsymbol{\beta}_2 + \mathbf{w}, \tag{6}$$

where $\boldsymbol{\beta}_1 = \boldsymbol{\delta}_1 \odot \boldsymbol{\beta}$ and $\boldsymbol{\beta}_2 = \boldsymbol{\delta}_2 \odot \boldsymbol{\beta}$; $\boldsymbol{\delta}_1$ and $\boldsymbol{\delta}_2$ are the off-grid errors in the azimuth and range directions, respectively.

Clearly, $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$ are also sparse vectors and share the same support with $\boldsymbol{\beta}$. $\mathbf{S}_1$ and $\mathbf{S}_2$ are the matrices obtained from the first-order Taylor expansion and share the same dimension with $\mathbf{S}$.

Obviously, the off-grid model (6) is closely related to the on-grid equation (4). In fact, the off-grid model is the first-order approximation of the true signal model, while the on-grid model is the zeroth-order approximation. Herein, the off-grid model has a much smaller modeling error. Finally, off-grid imaging reduces to a joint sparse reconstruction problem, shown in (6), where both $\boldsymbol{\beta}$ and $\{\boldsymbol{\beta}_1, \boldsymbol{\beta}_2\}$ need to be estimated sparsely using three known dictionaries $\{\mathbf{S}, \mathbf{S}_1, \mathbf{S}_2\}$.

Accordingly, in the presence of both radar system error and off-grid error, the RCI model $\mathbf{y} = \mathbf{S}(\Theta) \cdot \boldsymbol{\beta} + \mathbf{w}$ based on the parametric reference matrix is built as the following parametric joint sparse reconstruction model:

$$y \approx S(\Theta) \cdot \boldsymbol{\beta} + S_1(\Theta) \cdot \boldsymbol{\beta}_1 + \mathbf{S}_2(\Theta) \cdot \boldsymbol{\beta}_2 + \mathbf{w}, \tag{7}$$

where $\mathbf{S}(\Theta)$, $\mathbf{S}_1(\Theta)$, and $\mathbf{S}_2(\Theta)$ correspond to $\mathbf{S}$, $\mathbf{S}_1$, and $\mathbf{S}_2$ in formula (6), respectively. Herein, the specific expressions of these variables are not explicitly given for a general purpose. They can be definitely deduced when $\mathbf{S}$ and Θ are explicitly defined in the real application.

The above parametric joint sparse reconstruction model combines the parametric sparse representation model (for radar system error) and joint sparse reconstruction model (for off-grid error). Thus, the unknown parameters $\{\boldsymbol{\beta}, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2\}$ should be estimated jointly, and the model error Θ also needs to be calibrated. Hence, the proposed imaging model characterizes the model error well and utilizes the structured information of the model error to improve the performance.

3. Structured Sparse Bayesian Autofocus Algorithm

Based on the structured imaging model proposed in (7), the RCI with model errors can be investigated in a sparse reconstruction framework by exploiting the sparsity of the target. In fact, the target is assumed to be composed of a very limited amount of strong scatterers for a sufficiently high carrier frequency, which is widely used in the imaging radar system [34]. Accordingly, the number of scatterers is much smaller than that of the number of grid cells in the image plane, which means that the image is spatially sparse. Fully exploiting the sparse prior could facilitate the achievement of super-resolution, denoising, and feature extraction. This section establishes the structured sparse prior and then conducts the Bayesian inference to complete the target imaging in the VBEM framework.

3.1. Structured Sparse Bayesian Model. The sparse reconstruction problem can be reformulated as an estimation problem to be solved in the Bayesian framework [35,36]. In sparse Bayesian modeling, all unknown variables are treated as stochastic quantities with assigned probability distributions. The unknown sparse signal is assumed to follow a prior distribution to model the sparsity property. The

measurement $\mathbf{y}$ is also a random process with conditional distribution.

The measurement noise is generally assumed to be independent and Gaussian, with zero mean and α_0^{-1} variance:

$$p\left(\mathbf{w} \mid \alpha_0\right) = \mathscr{CN}\left(\mathbf{w} \mid 0, \alpha_0^{-1}\mathbf{I}\right) = \left(\pi\alpha_0^{-1}\right)^{-K} \exp\left\{-\alpha_0\|\mathbf{w}\|_2^2\right\}, \tag{8}$$

where α_0 is termed the *hyperparameter*, and an additional prior distribution, called the *hyperprior*, is assigned to it:

$$p\left(\alpha_0; c, d\right) = \Gamma\left(\alpha_0 \mid c, d\right), \tag{9}$$

where $\Gamma\left(\alpha_0 \mid c, d\right)$ is the Gamma distribution with parameters c and d.

To determine the sparse prior of $\boldsymbol{\beta}$, Laplace distribution is a popular choice [36,37]. However, the Laplace prior does not allow for a tractable Bayesian analysis and only provides point estimation without any higher-order statistical information. In order to fully utilize the uncertainty information during estimation, the signal is hierarchically modeled to impose a Gaussian-inverse Gamma prior that promotes sparsity [37].

In the first stage of the hierarchical model, each entry of $\boldsymbol{\beta}$ is independent and assigned a complex Gaussian distribution with prior variance α_k^{-1}:

$$p(\boldsymbol{\beta} \mid \boldsymbol{\alpha}) = \prod_{k=1}^{K} \mathscr{CN}\left(\beta_k \mid 0, \alpha_k^{-1}\right), \tag{10}$$

where $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, \ldots, \alpha_K]^T$. In the second stage, the precision (inverse-variance) of $\boldsymbol{\beta}$ is assumed to obey an independent Gamma distribution in the interests of convenient inference since it is the conjugate prior of Gaussian distribution:

$$p(\boldsymbol{\alpha}; a, b) = \prod_{k=1}^{K} \Gamma\left(\alpha_k \mid a, b\right). \tag{11}$$

The above model incorporates four free parameters, i.e., $\{a, b, c, d\}$. Herein, this paper sets $a = b = c = d = 10^{-6}$, which strongly promotes a sparse estimation and makes the prior noninformative [36]. By combining (10) and (11), it can be shown that the marginalized distribution of $\boldsymbol{\beta}$ is a complex Student-t distribution, which is suitable sparsity [36].

The hyperparameter $\boldsymbol{\alpha}$ is inversely proportional to the width of the probability density function (PDF). Naturally, a large value of α_k will drive the corresponding weight β_k to zero, which promotes the sparse solution accordingly. To investigate the structured information among $\{\boldsymbol{\beta}, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2\}$, the priors of $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$ should be properly assigned. As $\delta_{1,k} \in (-\Delta_1/2, \Delta_1/2]$, $\delta_{2,k} \in (-\Delta_2/2, \Delta_2/2]$, $\beta_{1,k} = \beta_k\delta_{1,k}$, $\beta_{2,k} = \beta_k\delta_{2,k}$, and the prior variance of β_k is α_k^{-1}, then the prior variances of $\beta_{1,k}$ and $\beta_{2,k}$ can be set as $\Delta_1^2\alpha_k^{-1}/4$ and $\Delta_2^2\alpha_k^{-1}/4$; this is to accommodate the worst case when the true scatterer is located in the middle of the two consecutive grids. Moreover, this setting also satisfies the constraint of common support among $\{\boldsymbol{\beta}, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2\}$. Consequently, the priors of $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$ can be expressed as

$$p\left(\boldsymbol{\beta}_1 \mid \boldsymbol{\alpha}\right) = \prod_{k=1}^{K} \mathscr{CN}\left(\beta_{1,k} \mid 0, \frac{\Delta_1^2\alpha_k^{-1}}{4}\right), \tag{12}$$

$$p\left(\boldsymbol{\beta}_2 \mid \boldsymbol{\alpha}\right) = \prod_{k=1}^{K} \mathscr{CN}\left(\beta_{2,k} \mid 0, \frac{\Delta_2^2\alpha_k^{-1}}{4}\right). \tag{13}$$

The structured sparse Bayesian modeling can be interpreted in a statistical way to impose the common support shared by the three unknown coefficients. Meanwhile, the proposed modeling should be considered as a flexible way to allow learning from data. Next, a full Bayesian inference procedure is carried out to make full use of the uncertainty information, as well as structured sparsity based on the hierarchical sparse prior model. The obtained solution has a smaller possibility of converging to a local minimum since the posterior estimation can smooth away the shallow local minimum [35].

3.2. Bayesian Inference. By combining the stages of the hierarchical Bayesian model, the joint distribution of all unknown and observed quantities can finally be acquired in order to derive the sparse Bayesian model of signal reconstruction. Bayesian inference based on the posterior distribution is then implemented to update the parameters and unknown signals. During the inference, $\boldsymbol{\Omega} = \{\boldsymbol{\beta}, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \boldsymbol{\alpha}, \alpha_0\}$ is modeled as the set of hidden and random variables to be estimated, while Θ is modeled as the deterministic parameter. Considering the RCI with model errors, the target imaging can be interpreted as an expectation maximization (EM) procedure that calibrates the model errors Θ and iteratively reconstructs the target. According to (8)–(13), a joint probability distribution can be obtained by

$$\begin{aligned} p(\mathbf{y}, \boldsymbol{\Omega}; \Theta) = {} & p(\mathbf{y} \mid \boldsymbol{\beta}, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \alpha_0; \Theta) \cdot p(\alpha_0; c, d) \\ & \cdot p(\boldsymbol{\beta} \mid \boldsymbol{\alpha}) \cdot p\left(\boldsymbol{\beta}_1 \mid \boldsymbol{\alpha}\right) \cdot p\left(\boldsymbol{\beta}_2 \mid \boldsymbol{\alpha}\right) \cdot p(\boldsymbol{\alpha}; a, b). \end{aligned} \tag{14}$$

The EM algorithm requires the knowledge of the posterior distribution [29] $p(\boldsymbol{\Omega} \mid \mathbf{y}; \Theta) = p(\boldsymbol{\Omega}, \mathbf{y}; \Theta)/p(\mathbf{y})$. However, the EM algorithm cannot be directly applied since $p(\mathbf{y})$ has no closed-form solution. Accordingly, variational Bayesian expectation maximization (VBEM) is used to approximate the intractable posterior. In a VBEM framework (14), the intractable posterior of hidden variables $\boldsymbol{\Omega}$ can be approximated using variational Bayesian inference (VBI) in the variational Bayesian expectation (VBE) step, while the unknown parameter Θ is estimated in the variational Bayesian maximization (VBM) step. The approximating PDF $q(\boldsymbol{\Omega})$ approximates the true posterior $p(\boldsymbol{\Omega} \mid \mathbf{y}; \Theta)$ by minimizing the Kullback–Leibler divergence (KLD) between $q(\boldsymbol{\Omega})$ and $p(\boldsymbol{\Omega} \mid \mathbf{y}; \Theta)$ [31]. Moreover, a structured mean field approximation can be applied to compute $q(\boldsymbol{\Omega})$: $q(\boldsymbol{\Omega}) = q(\boldsymbol{\beta})q(\boldsymbol{\beta}_1)q(\boldsymbol{\beta}_2)q(\boldsymbol{\alpha})q(\alpha_0)$. Consequently, the optimal posterior distribution for the i – th entry of $\boldsymbol{\Omega}$, i.e., $q(\boldsymbol{\Omega}_i)$, can be expressed as

$$q(\boldsymbol{\Omega}_i) \propto \exp\left\{\langle \ln p(\mathbf{y}, \boldsymbol{\Omega}; \Theta)\rangle_{q(\boldsymbol{\Omega}\backslash\boldsymbol{\Omega}_i)}\right\}, \tag{15}$$

where $\langle \cdot \rangle_{q(\boldsymbol{\Omega}\backslash\boldsymbol{\Omega}_i)}$ denotes the expectation with respect to $q(\boldsymbol{\Omega}\backslash\boldsymbol{\Omega}_i)$ and $\boldsymbol{\Omega}\backslash\boldsymbol{\Omega}_i$ represents the set $\boldsymbol{\Omega}$ without $\boldsymbol{\Omega}_i$. By applying the VBEM algorithm and the aforementioned Bayesian model, the best posterior densities for the entries of $\boldsymbol{\Omega}$ can be obtained from the following procedures:

(1) *VBE Step.* Update of $q(\boldsymbol{\Omega})$

In this step, the optimal posterior distribution $q(\boldsymbol{\Omega})$ is updated iteratively for the given matrices $\widehat{S} \triangleq \mathbf{S}(\Theta)$, $\widehat{S}_1 \triangleq \mathbf{S}_1(\Theta)$, and $\widehat{S}_2 = \mathbf{S}_2(\Theta)$, where Θ is the model error estimated in the last iteration.

The posterior distribution of scattering coefficients, i.e., $q(\boldsymbol{\beta})$, can be obtained as follows:

$$
\begin{aligned}
q(\boldsymbol{\beta}) &\propto \exp\left\{\langle \ln p(\mathbf{y} \mid \boldsymbol{\beta}, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \alpha_0; \Theta) p(\boldsymbol{\beta} \mid \boldsymbol{\alpha})\rangle_{q(\boldsymbol{\beta}_1) q(\boldsymbol{\beta}_2) q(\boldsymbol{\alpha}) q(\alpha_0)}\right\} \\
&= \mathscr{CN}(\boldsymbol{\beta} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}),
\end{aligned}
\tag{16}
$$

$$\boldsymbol{\Sigma} = \left(\langle \alpha_0 \rangle \widehat{S}^H \widehat{S} + \langle \boldsymbol{\Lambda} \rangle\right)^{-1}, \tag{17}$$

$$\boldsymbol{\mu} = \langle \alpha_0 \rangle \boldsymbol{\Sigma} \widehat{S}^H \left(\mathbf{y} - \widehat{S}_1 \cdot \boldsymbol{\mu}_1 - \widehat{S}_2 \cdot \boldsymbol{\mu}_2\right), \tag{18}$$

where $\langle \boldsymbol{\Lambda} \rangle = \mathrm{diag}(\langle \boldsymbol{\alpha} \rangle)$. Similarly, the posterior distributions of $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$ are $q(\boldsymbol{\beta}_1) = \mathscr{CN}(\boldsymbol{\beta}_1 \mid \boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1)$ and $q(\boldsymbol{\beta}_2) = \mathscr{CN}(\boldsymbol{\beta}_2 \mid \boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2)$. The parameters of the approximating factors can be computed as follows:

$$\boldsymbol{\Sigma}_1 = \left(\langle \alpha_0 \rangle \widehat{S}_1^H \widehat{S}_1 + 4\Delta_1^{-2} \langle \boldsymbol{\Lambda} \rangle\right)^{-1}, \tag{19}$$

$$\boldsymbol{\mu}_1 = \langle \alpha_0 \rangle \boldsymbol{\Sigma}_1 \widehat{S}_1^H \left(\mathbf{y} - \widehat{S}^H \boldsymbol{\mu} - \widehat{S}_2^H \boldsymbol{\mu}_2\right), \tag{20}$$

$$\boldsymbol{\Sigma}_2 = \left(\langle \alpha_0 \rangle \widehat{S}_2^H \widehat{S}_2 + 4\Delta_2^{-2} \langle \boldsymbol{\Lambda} \rangle\right)^{-1}, \tag{21}$$

$$\boldsymbol{\mu}_2 = \langle \alpha_0 \rangle \boldsymbol{\Sigma}_2 \widehat{S}_2^H \left(\mathbf{y} - \widehat{S}^H \boldsymbol{\mu} - \widehat{S}_1^H \boldsymbol{\mu}_1\right). \tag{22}$$

The hyperparameters $\{\alpha_k, \forall k\}$ are independent; thus, the probability density function (PDF) $q(\alpha_k)$ can be derived as $q(\alpha_k) = \mathrm{Gamma}(\alpha_k \mid \widehat{a}, \widehat{b})$. The parameters $\widehat{a}$ and $\widehat{b}$ of the Gamma distribution are given as

$$\widehat{a} = a + 3, \tag{23}$$

$$
\begin{aligned}
\widehat{b}_k = b &+ \left\{[\boldsymbol{\Sigma}]_{kk} + \left|[\boldsymbol{\mu}]_k\right|^2\right\} + 4\Delta_1^{-2}\left\{[\boldsymbol{\Sigma}_1]_{kk} + \left|[\boldsymbol{\mu}_1]_k\right|^2\right\} \\
&+ 4\Delta_2^{-2}\left\{[\boldsymbol{\Sigma}_2]_{kk} + \left|[\boldsymbol{\mu}_2]_k\right|^2\right\},
\end{aligned}
\tag{24}
$$

where $[\cdot]_{kk}$ represents the kth main diagonal element of a matrix and $[\cdot]_k$ is the kth element of a vector. Using the property of the Gamma distribution, the expected value of α_k can be computed as

$$\langle \alpha_k \rangle = \frac{\widehat{a}}{\widehat{b}}. \tag{25}$$

The optimal posterior distribution of the noise precision α_0 is obtained as $q(\alpha_0) = \mathrm{Gamma}(\alpha_0 \mid \widehat{c}, \widehat{d})$, where the parameters $\widehat{c}$ and $\widehat{d}$ are given as

$$\widehat{c} = c + N, \tag{26}$$

$$\widehat{d} = d + \left\|\mathbf{y} - \widehat{S}\boldsymbol{\mu} - \widehat{S}_1\boldsymbol{\mu}_1 - \widehat{S}_2\boldsymbol{\mu}_2\right\|_2^2 + \langle \alpha_0 \rangle^{-1} \sum_{k=1}^{K} \lambda, \tag{27}$$

where $\lambda_k = 3 - \langle \alpha_0 \rangle\left\{[\boldsymbol{\Sigma}]_{kk} + 4\Delta_1^{-2}[\boldsymbol{\Sigma}_1]_{kk} + 4\Delta_2^{-2}[\boldsymbol{\Sigma}_2]_{kk}\right\}$. The expectation of α_0 is

$$\langle \alpha_0 \rangle = \frac{\widehat{c}}{\widehat{d}}. \tag{28}$$

The optimal distribution $q(\boldsymbol{\Omega})$ is obtained by iteratively calculating the above steps until convergence. In order to estimate the model error Θ, each hidden variable should be updated before proceeding to the next step.

(2) *VBM Step.* Update of Θ

According to the VBEM, Θ is estimated when the estimate of $\boldsymbol{\Omega}$ is given:

$$
\begin{aligned}
\widehat{\Theta} &= \arg\max_{\Theta} \langle \ln p(\mathbf{y}, \boldsymbol{\Omega}; \Theta)\rangle_{q(\boldsymbol{\Omega})} \\
&= \arg\min_{\Theta} \left\{\langle \left\|\mathbf{y} - \mathbf{S}(\Theta)\cdot\boldsymbol{\beta} - \mathbf{S}_1(\Theta)\cdot\boldsymbol{\beta}_1 - \mathbf{S}_2(\Theta)\cdot\boldsymbol{\beta}_2\right\|_2^2\rangle\right\}.
\end{aligned}
\tag{29}
$$

Hence, the estimation of Θ is a nonlinear least-square problem, but it is difficult to obtain the closed-form expression; thus, the Quasi-Newton method is adopted to update the parameter, as follows:

$$\widehat{\Theta} = \Theta - \left[\nabla^2 f(\Theta)\right]^{-1} [\nabla f(\Theta)], \tag{30}$$

where $f(\Theta) = \langle \left\|\mathbf{y} - \mathbf{S}(\Theta)\cdot\boldsymbol{\beta} - \mathbf{S}_1(\Theta)\cdot\boldsymbol{\beta}_1 - \mathbf{S}_2(\Theta)\boldsymbol{\beta}_2\right\|_2^2\rangle$ is the objective function and $\nabla f(\Theta)$ and $\nabla^2 f(\Theta)$ represent the gradient and Hessian with respect to the model error, respectively. The derivations of $\nabla f(\Theta)$ and $\nabla^2 f(\Theta)$ are presented in [13]. After estimating the model error Θ, the matrices $\widehat{S} \triangleq \mathbf{S}(\Theta)$, $\widehat{S}_1 \triangleq \mathbf{S}_1(\Theta)$, and $\widehat{S}_2 = \mathbf{S}_2(\Theta)$ should be updated before stepping into the next iteration.

The proposed method above utilizes the hierarchical modeling procedure to encode signal sparsity which could obtain the approximate posterior distribution $q(\boldsymbol{\beta})$ and is regarded as a full Bayesian method. The statistical

information is used to enhance the estimation performance and avoid converging to a shallow local minimum during the learning procedure due to the utilization of higher-order statistical information.

3.3. Algorithm Summary and Discussion

Remark 1. Convergence. The VBEM procedure can be summarized as a process of iteratively updating the hidden and deterministic parameters. The SSBA algorithm is based on the VBEM framework. The update of $\mathbf{\Omega}$ decreases the KLD between $q(\mathbf{\Omega})$ and $p(\mathbf{\Omega} \mid \mathbf{y}; \Theta)$, i.e., $q(\mathbf{\Omega}) = \mathrm{KLD}_{q(\mathbf{\Omega})}(q(\mathbf{\Omega}) \parallel p(\mathbf{\Omega} \mid \mathbf{y}; \Theta))$, while the update of Θ, respectively, decreases the negative expected log-likelihood function until convergence. Thus, the (marginal) likelihood monotonically increases throughout the iterations, and the convergence is guaranteed [31]. Numerical results presented in Section 4 also validate that the algorithm will converge within tens of iterations.

Remark 2. Computational Complexity. The SBL-based SSBA algorithm operates in an iterative manner; both the number of iterations and the computational cost of each iteration determine the total computational cost. Compared with the conventional Fourier-based approach, the SBL-based algorithm is known to converge rather slowly, and its computational complexity is much higher [6]. Generally, the above procedure involves the matrix inversion in (17) and matrix-vector multiplication, the computational costs of which are $o(K^3)$ and $o(K^2)$, respectively. This procedure is computationally expensive, particularly, for a large K, which is generally the case in practical applications. Moreover, the convergence slows down when model errors exist, meaning that more iterations are necessary to reach convergence. Thus, the algorithm is impractical for use with large-scale applications. Therefore, a faster algorithm should be developed to decrease the computational complexity or to obtain fast convergence.

Fortunately, a grid pruning operation could be applied to decrease the computational burden since most of the prior precisions (i.e., α_k) tend to quite large values upon convergence benefiting from the sparsity of the target. The pruning of the current grids set Ψ^j can be achieved via

$$\Psi^{j+1} = \left\{ k \mid \langle \alpha_k \rangle < \alpha_{\mathrm{th}}, k \in \Psi^j \right\}, \tag{31}$$

where Ψ^{j+1} denotes the grids set after pruning and α_{th} is a large threshold. Finally, the computational burden is reduced, and the convergence rate is improved due to the grid pruning; moreover, the sparsity of the reconstruction results is enhanced, while the common support shared by $\{\mathbf{\beta}, \mathbf{\beta}_1, \mathbf{\beta}_2\}$ is guaranteed.

In summary, the procedure of the proposed SSBA algorithm based on VBEM is outlined in Algorithm 1.

4. Simulations and Discussion

In this section, simulations are carried out to validate the performance of the proposed SSBA algorithm. SSBA

provides a unified framework for RCI in the presence of various model errors, while a scenario with only two types of model errors is considered in the simulations, namely, gain-phase error (type I error) and off-grid error (type II error). The RCI system consists of one receiver and $M = 16$ transmitters configured as uniform linear array (ULA) and interelement spacing $d = 0.5$ m. The transmitters emit random-modulated frequency-hopping (FH) waveforms [39] with carrier frequency $f_c = 10$ GHz and signal bandwidth $B = 0.5$ GHz. The azimuth-range imaging is discretized to 40×40 grid cells with grid-cell size $0.001 \, \mathrm{rad} \times 0.1$ m. First, we calculate the matrices $\widehat{S} = \mathbf{S}(\Theta)$, $\widehat{S}_1 = \mathbf{S}_1(\Theta)$, and $\widehat{S}_2 = \mathbf{S}_2(\Theta)$ by setting $\Theta = \mathbf{0}$ to initialize the input parameters of the SSBA algorithm. Furthermore, we initialize $\alpha_0 = 10^2/\mathrm{var}(\mathbf{y})$, $\mathbf{\alpha} = N/|\widehat{S}^H \mathbf{y}|$, $\alpha_{\mathrm{th}} = 10^6$, and $a = b = c = d = 10^{-6}$. To terminate the algorithm, this paper sets the tolerance $\gamma = 10^{-3}$ and the maximum number of iterations to $J_{\max} = 200$. In the simulations, the gain and phase errors randomly vary at $[0.7, 1.3]$ and $[-\pi/4, \pi/4]$, respectively. The scatterers of grid mismatch are randomly located in between the grid cells. Therefore, the maximum size of grid mismatch is half the grid size in theory.

In addition to the proposed SSBA algorithm, SBL [36], S-TLS [26], and R-FOCUSS [24] are also simulated to demonstrate the superior performance of SSBA. SBL reconstructs the target without the model errors. S-TLS and R-FOCUSS consider sparse recovery with dictionary perturbation: $\mathbf{y} = (\mathbf{S} + \mathbf{E})\mathbf{\beta} + \mathbf{w}$. S-TLS solves the nonconvex optimization problem, while R-FOCUSS is proposed based on the SD-FOCUSS algorithm, which aims to find the maximum *a posteriori* (MAP) estimation:

$$\min_{\mathbf{\beta}, \mathbf{E}} \quad \left\{ \|\mathbf{y} - (\mathbf{S} + \mathbf{E})\mathbf{\beta}\|_2^2 + \lambda_1 \|\mathbf{E}\|_F^2 + \lambda_2 \|\mathbf{\beta}\|_p^p \right\}, \tag{32}$$

in practical applications, and the regularization parameters λ_1 and λ_2 in R-FOCUSS are properly chosen rather than known *a priori* as in SD-FOCUSS. The original R-FOCUSS uses the output of the Tikhonov regularization method [40] as initialization, while the result of SBL is employed to initialize R-FOCUSS in this paper. In addition, the sparsity-driven autofocus (SDA) method [41] is also extended to solve the imaging problem (7) as

$$\min_{\mathbf{\beta}, \Theta} \left\{ \left\| \mathbf{y} - \mathbf{S}(\Theta) \cdot \mathbf{\beta} - \mathbf{S}_1(\Theta) \cdot \mathbf{\beta}_1 - \mathbf{S}_2(\Theta) \cdot \mathbf{\beta}_2 \right\|_2^2 \right.$$
$$\left. + \lambda \left(\|\mathbf{\beta}\|_1 + 4\Delta_1^{-2} \|\mathbf{\beta}_1\|_1 + 4\Delta_2^{-2} \|\mathbf{\beta}_2\|_1 \right) \right\}, \tag{33}$$

where λ is the regularization parameter, which specifies the strength of the contribution of the regularization term into the solution.

4.1. Qualitative Results and Comparison with Other Algorithms.

To demonstrate the validity of the SSBA algorithm, the paper conducts a numerical simulation where the conventional SBL algorithm is also implemented as a comparison. Gaussian noise is added to the echo vector $\mathbf{y}$ with a signal-to-noise ratio (SNR) of 20 dB. Figure 2 presents the imaging results. The actual target image is

Initialization $j = 0$, $\Theta = \mathbf{0}$, $\widehat{S} = \mathbf{S}\,(\Theta = \mathbf{0})$, $\widehat{S}_1 = \mathbf{S}_1\,(\Theta = \mathbf{0})$, $\widehat{S}_2 = \mathbf{S}_2\,(\Theta = \mathbf{0})$, and $\Psi^0 = \{k \mid k = 1, \cdots, K\}$
Input $\mathbf{y}$, $\widehat{S}$, $\widehat{S}_1$, $\widehat{S}_2$, $\{a, b, c, d\}$, α_{th}, γ, and $J_{\max}$
while not converged do
   (1) VBE step: update $\mathbf{\Sigma}$, $\mathbf{\mu}$, $\mathbf{\Sigma}_1$, $\mathbf{\mu}_1$, $\mathbf{\Sigma}_2$, $\mathbf{\mu}_2$, $\langle \alpha_k \rangle$, and $\langle \alpha_0 \rangle$ from (17)–(22), (25), and (28), where $k \in \Psi^j$
   (2) VBM step: estimate Θ from (30) and update $\widehat{S} = \mathbf{S}\,(\Theta)$, $\widehat{S}_1 = \mathbf{S}_1\,(\Theta)$, and $\widehat{S}_2 = \mathbf{S}_2\,(\Theta)$
   (3) Prune the grids: $\Psi^{j+1} = \{k \mid \langle \alpha_k \rangle < \alpha_{\mathrm{th}}, k \in \Psi^j\}$
   (4) Check for convergence: $\|\mathbf{\mu}^j - \mathbf{\mu}^{j-1}\|_2^2 / \|\mathbf{\mu}^j\|_2^2 < \gamma$ or $j = J_{\max}$ and then $j = j + 1$
end while
  Output reconstructed scattering coefficient vector $\widehat{\beta} = \mathbf{\mu}$

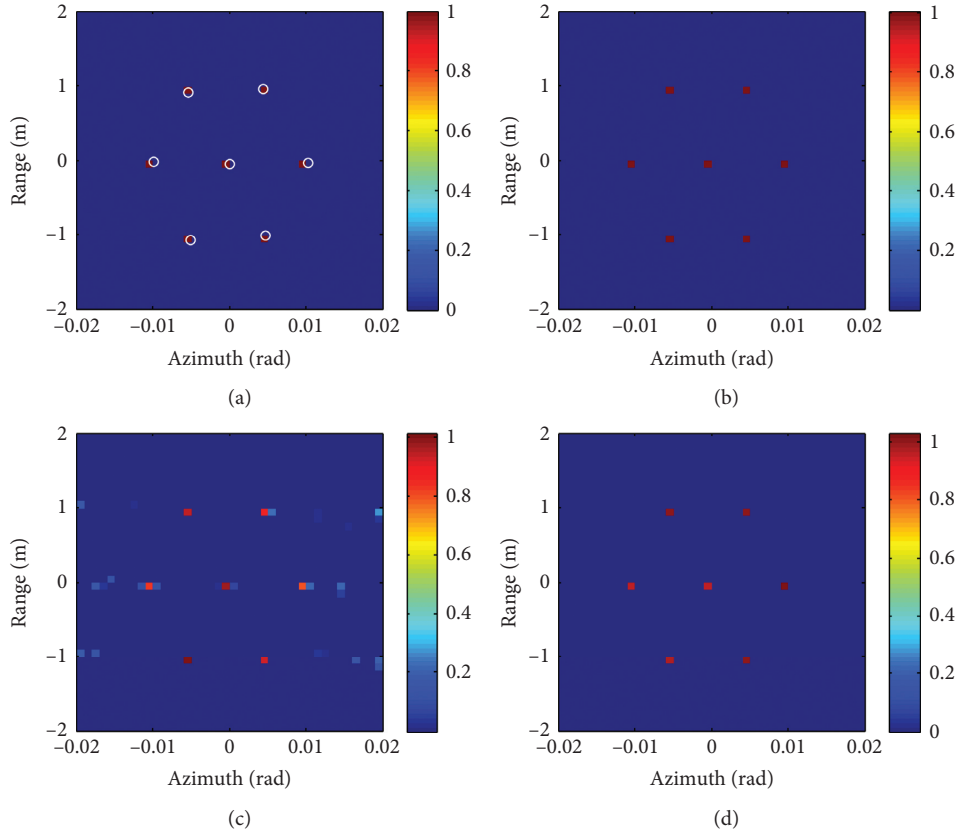ALGORITHM 1: The procedure of the proposed SSBA algorithm based on VBEM.

FIGURE 2: RCI results: (a) ground truth; (b) imaging result of conventional SBL when model errors are calibrated accurately; (c) imaging result of conventional SBL without considering model errors; (d) imaging result of SSBA.

shown in Figure 2(a), where the scatterers are randomly distributed in between the grid cells. The white circles in Figure 2(a) represent the actual locations of scatterers, while the red points represent the closest grid cells to the scatterers. The target image is deemed to have been reconstructed successfully if the off-grid scatterers are captured by the closest grid-cell centers.

Figure 2(b) plots the image reconstructed by SBL when model errors are completely calibrated. It can be seen that SBL achieves perfect performance while the scatterers are reconstructed accurately. However, it is too difficult to calibrate model errors accurately in practical applications; when complete calibration is impossible, the performance of SBL degrades, as shown in Figure 2(c). The image in

Figure 2(c) is defocused, and many spurious scatterers exist. By contrast, compared with Figure 2(c), the image shown in Figure 2(d) is significantly improved. The scatterers are reconstructed much better without any obvious spurious scatterers, mainly because the prior knowledge of sparsity is fully utilized and the model errors are calibrated well. Accordingly, the proposed SSBA algorithm can be applied to RCI with multiple model errors.

To demonstrate the superior performance of SSBA, four sparsity-based algorithms are simulated, i.e., SBL, S-TLS, SDA, and R-FOCUSS. The imaging results are depicted in Figure 3. As an outstanding sparse recovery method, SBL has shown superior performance in radar imaging but still cannot focus the target image well as the model errors exist.
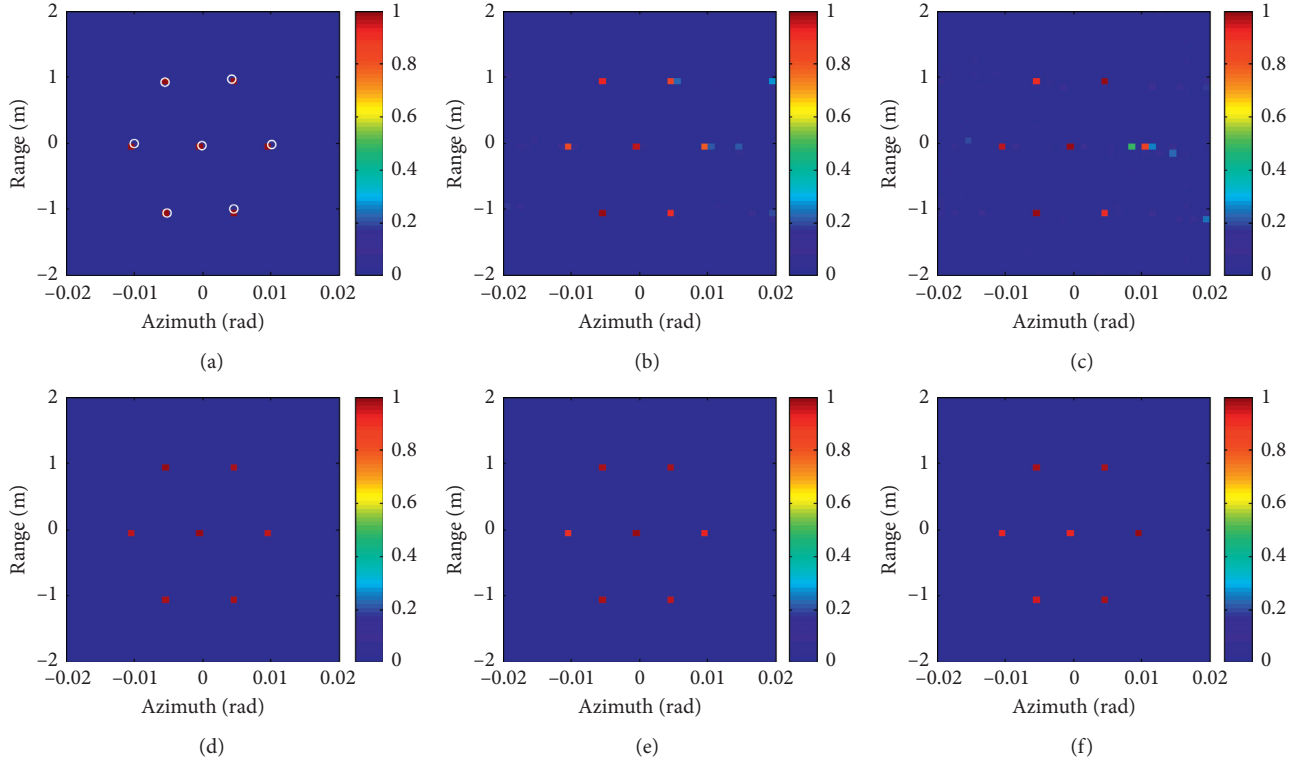
FIGURE 3: RCI results: (a) ground truth; (b) SBL; (c) S-TLS; (d) SDA; (e) R-FOCUSS; (f) SSBA.

From the Bayesian perspective, S-TLS yields an MAP solution when the perturbation matrix **E** is Gaussian, which is hard to be satisfied. S-TLS does not utilize the structured information of model errors and tends to generate nonsparse and defocused images. Comparably, the three algorithms (SDA, R-FOCUSS, and SSBA) generate focused images with quite a few spurious scatterers, as shown in Figures 3(d)–3(f). The three algorithms fully exploit both the sparsity of target and the structured information of model errors (or statistical information of perturbation matrix for R-FOCUSS) and thus perform well for fully perturbed RCI.

To evaluate the numerical complexity of SSBA, the time needs of the five algorithms are recorded in Table 1. The simulations are performed on a computer with Intel Xeon CPU E5-2670 at 2.5 GHz and 32 GB of memory. S-TLS uses the convex optimization method to reconstruct the target; thus, it is time-consuming and suffers from heavy computational complexity. SBL and SDA are also time-consuming because the matrix inversion and matrix-vector multiplication are calculated in each iteration. Comparably, R-FOCUSS has the lowest computation complexity. Due to the grid pruning, SSBA converges rapidly and is also computationally efficient and could thus be applied to practical RCI. Furthermore, SSBA has advantages over R-FOCUSS considering both RIE and image entropy, which is compared in detail in Figure 4.

4.2. Quantitative Results for Different SNRs. In this part, the performance of SSBA is evaluated quantitatively under different SNRs by means of Monte Carlo simulations and is

TABLE 1: Time needs of the five algorithms.

Algorithm	SBL	S-TLS	SDA	R-FOCUSS	SSBA
Runtime (s)	107.1	4514.4	454.1	6.6	47.2

also compared with that of the other algorithms (i.e., SBL, S-TLS, SDA, and R-FOCUSS). In the simulations, the SNR varies from −10 dB to 20 dB. Two criteria are introduced to qualify the performance, i.e., relative imaging error (RIE) and image entropy. RIE is defined as $20 \log_{10} (\|\beta - \hat{\beta}\|_2 / \|\beta\|_2)$, where β and $\hat{\beta}$ are the true and estimated scattering coefficient vectors, respectively. Image entropy is defined as

$$E = - \sum_{k=1}^{K} \left[\left(\left| \hat{\beta}_k \right|^2 / P \right) \ln \left(\left| \hat{\beta}_k \right|^2 / P \right) \right], \quad (34)$$

where $P = \sum_{k=1}^{K} |\hat{\beta}_k|^2$ is the energy of the reconstructed image. Generally, lower image entropy indicates a better-focused image, which means that the image is sparser.

The results are presented in Figure 4. Figures 4(a) and 4(b) show the RIE and image entropy versus SNR, respectively. It can be seen from Figure 4(a) that the reconstructed RIEs decrease quickly, and that the proposed SSBA tends to reconstruct the most accurate image, as the SNR increases. In terms of the focusing performance shown in Figure 4(b), the three SBL-type algorithms (i.e., SBL, SDA, and SSBA) achieve superior performance in high-SNR regimes, while SSBA generates a much sparser and more focused image with smaller RIE. The image reconstructed by S-TLS is defocused and blurry. Apart from the strong
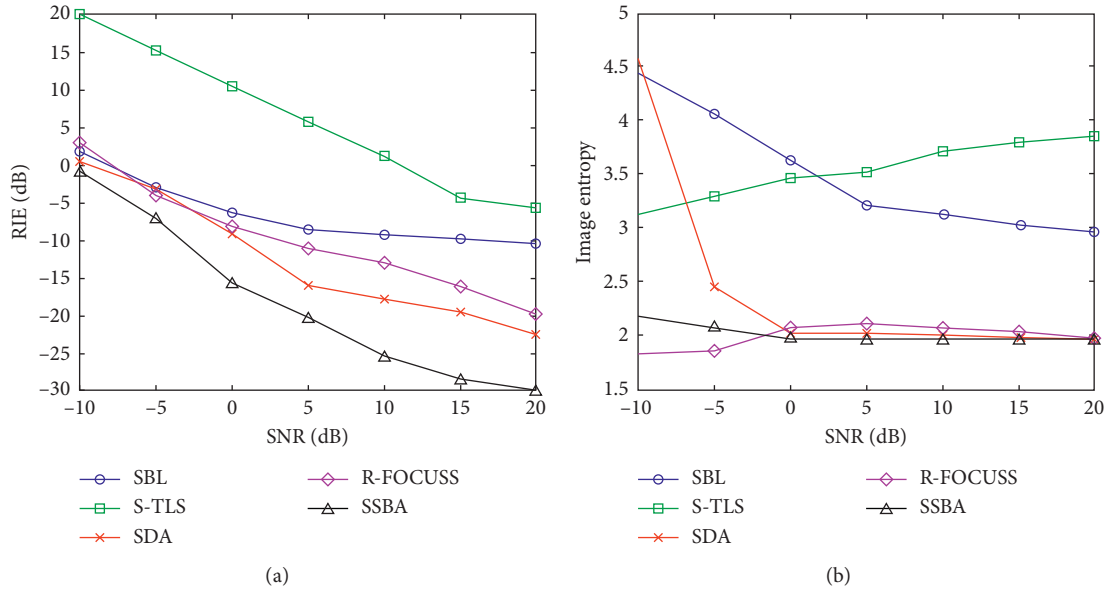
FIGURE 4: RIE and image entropy for different SNRs: (a) RIE vs. SNR; (b) image entropy vs. SNR.

scatterers, many spurious scatterers exist. In low-SNR regimes (SNR < 0 dB), S-TLS and R-FOCUSS seem to reconstruct the target with a lower image entropy; however, the RIEs are much higher, meaning that the target is not imaged successfully. S-TLS and R-FOCUSS employ the perturbation matrix **E** to model the effect of model errors without utilizing the structured information fully so that the model errors are not calibrated well. Moreover, when the noise is strong, overfitting emerges, which makes S-TLS and R-FOCUSS converge to the local optima easily. Generally, while all five algorithms are sensitive to noise, SSBA achieves superior results relative to the other imaging algorithms and significantly improves the practical imaging performance.

Benefiting from the fact that the sparsity prior and model errors are fully exploited, the proposed SDA and SSBA exhibit superior performance. Although they assign different sparse priors of $\boldsymbol{\beta}$, both of them are SBL-based algorithms that share the same imaging models. SDA is a l_1-based regularization method, in which a sparsity-inducing Laplace prior is directly imposed on the signal, after which the sparse solution is exploited from the MAP estimation corresponding to the point estimation of the sparse coefficients. Comparably, SSBA utilizes hierarchical modeling to encode signal sparsity and achieve better sparse solutions, which can obtain the approximate posterior distribution, and is regarded as a full Bayesian method. Herein, higher-order statistical information is used to enhance the imaging performance and avoid converging to a shallow local minimum. Moreover, the alternate optimization between the sparse coefficients and model errors inevitably results in error propagation [6,42], meaning that the estimation error of the sparse coefficients degrades the estimation accuracy of model errors during iterations [43,44]. As for SSBA, the utilization of a higher-order statistical information alleviates the effect of error propagation and improves the imaging performance considerably.

5. Conclusion

In summary, this paper focuses on the sparsity-based RCI with model errors. Firstly, the parametric joint sparse reconstruction model was built to utilize the structured information of model errors; then, an appropriately structured sparse Bayesian prior was assigned to the sparse coefficients and the SSBA algorithm was proposed under the VBEM framework. SSBA fully exploited the sparse prior and structured information of model errors and jointly realized model error calibration and sparse imaging. Numerical experiments demonstrated the superior performance of SSBA, which could yield a well-focused target image with high reconstruction accuracy in the presence of model errors. Thus, SSBA can be applied in practical radar systems.

Although SSBA performs well for RCI with model errors, its performance is sensitive to the parameters of the RCI model (e.g., grid-cell size and array aperture). Besides, the incoherence of the reference matrix determines the imaging performance. In the future, the effects of model parameters should be investigated. Moreover, SSBA still suffers from high computational complexity, especially when the grids are denser or a number of model errors coexist, despite the fact that the unified framework and grid pruning are applied. Thus, the development of fast SSBA algorithms is also of great interest.

Data Availability

In this paper, the proposed SSBA algorithm is dedicated to theoretical research. So, all the data are generated by computer simulation. The specific method of generating the simulated data has been described in detail in Section 2 from equations (1) to (7). Moreover, the values of the simulated parameters are also given in Section 4. Accordingly, all the readers can easily reproduce the needed data to verify our

conclusions. It means that the simulated data are the "internal" data, and no "external" data are needed to validate or reproduce the results. Therefore, we claim that "the simulated data used to support the findings of this study are included within the article."

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] Y. Cheng, X. Zhou, X. Xu, Y. Qin, and H. Wang, "Radar coincidence imaging with stochastic frequency modulated array," *IEEE Journal of Selected Topics in Signal Processing*, vol. 11, no. 2, pp. 414–427, 2017.

[2] Z. Wu, L. Zhang, H. Liu, and N. Kou, "Enhancing microwave metamaterial aperture radar imaging performance with rotation synthesis," *IEEE Sensors Journal*, vol. 16, no. 22, pp. 8035–8043, 2016.

[3] D. A. Ausherman, A. Kozma, J. L. Walker, H. M. Jones, and E. C. Poggio, "Developments in radar imaging," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-20, no. 4, pp. 363–400, 1984.

[4] A. Moreira, P. Prats-Iraola, M. Younis, G. Krieger, I. Hajnsek, and K. P. Papathanassiou, "A tutorial on synthetic aperture radar," *IEEE Geoscience and Remote Sensing Magazine*, vol. 1, no. 1, pp. 6–43, 2013.

[5] M. Cetin and W. C. Karl, "Feature-enhanced synthetic aperture radar image formation based on nonquadratic regularization," *IEEE Transactions on Image Processing*, vol. 10, no. 4, pp. 623–631, 2001.

[6] L. Zhao, L. Wang, L. Yang, A. M. Zoubir, and G. Bi, "The race to improve radar imagery: an overview of recent progress in statistical sparsity-based techniques," *IEEE Signal Processing Magazine*, vol. 33, no. 6, pp. 85–102, 2016.

[7] L. Zhao, L. Wang, G. Bi, and L. Yang, "An autofocus technique for high-resolution inverse synthetic aperture radar imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, pp. 6392–6403, 2014.

[8] R. Wang, Y. K. Deng, O. Loffeld et al., "Processing the azimuth-variant bistatic SAR data by using monostatic imaging algorithms based on two-dimensional principle of stationary phase," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 10, pp. 3504–3520, 2011.

[9] R. Wang, W. Wang, Y. Shao et al., "First bistatic demonstration of digital beamforming in elevation with TerraSAR-X as an illuminator," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 2, pp. 842–849, 2016.

[10] D. Li, X. Li, Y. Qin, Y. Cheng, and H. Wang, "Radar coincidence imaging: an instantaneous imaging technique with stochastic signals," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 4, pp. 2261–2277, 2014.

[11] S. Zhu, A. Zhang, Z. Xu, and X. Dong, "Radar coincidence imaging with random microwave source," *IEEE Antennas and Wireless Propagation Letters*, vol. 14, pp. 1239–1242, 2015.

[12] A. C Fannjiang, "Exact localization and superresolution with noisy data and random illumination," *Inverse Problems*, vol. 27, pp. 1–27, 2011.

[13] X. Zhou, H. Wang, Y. Cheng, and Y. Qin, "Sparse autocalibration for radar coincidence imaging with gain-phase errors," *Sensors*, vol. 15, no. 11, pp. 27611–27624, 2015.

[14] X. Zhou, H. Wang, Y. Cheng, Y. Qin, and H. Chen, "Off-grid radar coincidence imaging based on variational sparse bayesian learning," *Mathematical Problems in Engineering*, vol. 201612 pages, 2016.

[15] X. Xu, X. Zhou, Y. Cheng, and Y. Qin, "Radar coincidence imaging with array position error," in *Proceedings of the IEEE International Conference on Signal Processing, Communications and Computing*, Ningbo, China, 2015.

[16] D. Li, X. Li, Y. Cheng, Y. Qin, and H. Wang, "Radar coincidence imaging under grid mismatch," *ISRN Signal Processing*, vol. 2014, Article ID 987803, 8 pages, 2014.

[17] G. Zha, H. Wang, Z. Yang, Y. Cheng, and Y. Qin, "Adaptive sparse reconstruction with joint parametric estimation for high-speed uniformly moving targets in coincidence imaging radar," *Journal of Applied Remote Sensing*, vol. 10, no. 2, pp. 1–17, 2016.

[18] E. J. Candes and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 21–30, 2008.

[19] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[20] K. Cao, X. Zhou, Y. Cheng, B. Fan, and Y. Qin, "Total variation-based method for radar coincidence imaging with model mismatch for extended target," *Journal of Electronic Imaging*, vol. 26, no. 6, pp. 1–17, 2017.

[21] B. Shi, S. Chen, Y. Tian, X. Fan, and Q. Lian, "FASPR: a fast sparse phase retrieval algorithm via the epigraph concept," *Digital Signal Processing*, vol. 80, no. 9, pp. 12–26, 2018.

[22] B. Shi, Q. Lian, X. Huang, and N. An, "Constrained phase retrieval: when alternating projection meets regularization," *Journal of the Optical Society of America B*, vol. 35, no. 6, pp. 1271–1281, 2018.

[23] B. Shi, Q. Lian, S. Chen, Y. Tian, and X. Fan, "Coded diffraction imaging via double sparse regularization model," *Digital Signal Processing*, vol. 79, no. 8, pp. 23–33, 2018.

[24] B. Shi, Q. Lian, and X. Fan, "PPR: plug-and-play regularization model for solving nonlinear imaging inverse problems," *Signal Processing*, vol. 162, no. 9, pp. 83–96, 2019.

[25] M. A. Herman and T. Strohmer, "General deviants: an analysis of perturbations in compressed sensing," *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, no. 2, pp. 342–349, 2010.

[26] H. Zhu, G. Leus, and G. B. Giannakis, "Sparsity-cognizant total least-squares for perturbed compressive sampling," *IEEE Transactions on Signal Processing*, vol. 59, no. 5, pp. 2002–2016, 2011.

[27] X. Han, H. Zhang, and G. Li, "Fast algorithms for sparse recovery with perturbed dictionary," 2011, https://arxiv.org/abs/1111.6237.

[28] Z.-M. Liu and Y.-Y. Zhou, "A unified framework and sparse Bayesian perspective for direction-of-arrival estimation in the presence of array imperfections," *IEEE Transactions on Signal Processing*, vol. 61, no. 15, pp. 3786–3798, 2013.

[29] L. Zhao, G. Bi, L. Wang, and H. Zhang, "An improved autocalibration algorithm based on sparse bayesian learning framework," *IEEE Signal Processing Letters*, vol. 20, pp. 889–892, 2013.

[30] L. Zhao, L. Wang, G. Bi, S. Li, L. Yang, and H. Zhang, "Structured sparsity-driven autofocus algorithm for high-

resolution radar imagery," *Signal Processing*, vol. 125, pp. 376–388, 2016.

[31] D. G. Tzikas, A. C. Likas, and N. P. Galatsanos, "The variational approximation for bayesian inference," *IEEE Signal Processing Magazine*, vol. 25, no. 6, pp. 131–146, 2008.

[32] W. Rao, G. Li, X. Wang, and X.-G. Xia, "Parametric sparse representation method for ISAR imaging of rotating targets," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 50, no. 2, pp. 910–919, 2014.

[33] I. Tosic and P. Frossard, "Dictionary learning," *IEEE Signal Processing Magazine*, vol. 28, no. 2, pp. 27–38, 2011.

[34] L. C. Potter, D. Da-Ming Chiang, R. Carriere, and M. J. Gerry, "A GTD-based parametric model for radar scattering," *IEEE Transactions on Antennas and Propagation*, vol. 43, no. 10, pp. 1058–1067, 1995.

[35] D. P. Wipf and B. D. Rao, "Sparse bayesian learning for basis selection," *IEEE Transactions on Signal Processing*, vol. 52, no. 8, pp. 2153–2164, 2004.

[36] S. Ji, Y. Xue, and L. Carin, "Bayesian compressive sensing," *IEEE Transactions on Signal Processing*, vol. 56, no. 6, pp. 2346–2356, 2008.

[37] S. D. Babacan, R. Molina, and A. K. Katsaggelos, "Bayesian compressive sensing using laplace priors," *IEEE Transactions on Image Processing*, vol. 19, no. 1, pp. 53–63, 2010.

[38] L. Hu, J. Zhou, Z. Shi, and Q. Fu, "A fast and accurate reconstruction algorithm for compressed sensing of complex sinusoids," *IEEE Transactions on Signal Processing*, vol. 61, no. 22, pp. 5744–5754, 2013.

[39] C. Chen and P. P. Vaidyanathan, "MIMO radar ambiguity properties and optimization using frequency-hopping waveforms," *IEEE Transactions on Signal Processing*, vol. 56, pp. 5926–5936, 2008.

[40] A. N. Tikhonov, "Solution of incorrectly formulated problems and the regularization method," *Soviet Mathematics Doklady*, vol. 4, pp. 1035–1038, 1963.

[41] N. Ö. Onhon and M. Cetin, "A sparsity-driven approach for joint SAR imaging and phase error correction," *IEEE Transactions on Image Processing*, vol. 21, no. 4, pp. 2075–2088, 2012.

[42] S. Chen, X. Hua, H. Wang, C. Luo, Y. Cheng, and B. Deng, "Three-dimensional terahertz coded-aperture imaging based on geometric measures," *Sensors*, vol. 18, no. 5, p. 1582, 2018.

[43] B. Yuan, Y. Guo, W. Chen, and D. Wang, "A novel microwave staring correlated radar imaging method based on bi-static radar system," *Sensors*, vol. 19, no. 4, p. 879, 2019.

[44] M. Zhao, S. Zhu, J. Li et al., "Frequency-diverse bunching metamaterial antenna for coincidence imaging," *Materials*, vol. 12, no. 11, p. 1817, 2019.