Clustered Sparsity-Driven SAR Imaging and Autofocus Algorithm in Structured Phase-Noisy Environments

YUE YANG[1], **(Student Member, IEEE), XUEJING ZHANG**[1], **(Student Member, IEEE), GUAN GUI**[2], **(Senior Member, IEEE), AND QUN WAN**[1], **(Member, IEEE)**

[1]School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China
[2]College of Telecommunication and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210003, China

Corresponding author: Qun Wan (wanqun@uestc.edu.cn)

ABSTRACT This paper presents a novel synthetic aperture radar (SAR) imaging and autofocus algorithm by exploiting the clustering property of the sparse scene in structured phase-noisy environments. The SAR imaging and autofocus problem are reformulated in a Bayesian framework where a clustered sparsity-inducing prior is imposed on the target scene and a correlated probabilistic model is assumed on phase error. In particular, a pattern coupled Gaussian distribution and a Markov random fields (MRF) model are assigned to the scattering coefficients and the support of the sparse scatterers, respectively, to enforce the underlying continuity of illuminated target scene. Furthermore, the multivariate Von Mises prior is incorporated to capture the spatial fluctuation structures of the phase errors along with the aperture positions. The posterior inference is then obtained by using the mean-field variational Bayesian expectation-maximization (VBEM) method. By taking advantage of the inherent target correlation characteristics and the uncertain structures introduced on phase errors, the devised method can result in enhanced imaging and autofocus performance. The experimental results are provided to demonstrate the superiority of the devised scheme in obtaining the concentrated target region over the other reported sparsity-driven SAR imaging and autofocus algorithms.

INDEX TERMS Synthetic aperture radar (SAR), phase noise, Markov random fields (MRF), multivariate Von Mises distribution, variational Bayesian expectation-maximization (VBEM).

I. INTRODUCTION

Synthetic aperture radar (SAR) works well in all-weather and dim-light circumstances and has been extensively applied in target imaging. Recently, sparse representation (SR) based SAR imaging has been proposed and attracted much attention for its superiority in achieving high resolution with limited measurements, see e.g. [1]–[5]. The sparsity-driven SAR image formation problem is essentially an inverse problem, in which we are desired to reconstruct an image of the complex reflectivity field from the given noisy SAR measurement. One common strategy to solve this inverse problem is to regularize the solution space in a deterministic or probabilistic manner by incorporating the sparsity-inducing prior.

The associate editor coordinating the review of this manuscript and approving it for publication was Luca Barletta.

For instance, a variety of radar imaging algorithms have been developed by exploiting the $l_p(0 \leq p \leq 1)$ norm based sparse regularization term, see such as the greedy methods in [6] with $p = 0$, the basis pursuit methods in [7] with $p = 1$ and the sparse Bayesian learning methods in [3], [8] with $p \in (0, 1]$. To further enhance the performance of radar reconstruction, the inherent clustered structures of scatterers underlying sparsity patterns are introduced in [9]–[13]. More specifically, the first-order neighborhood scatterers are discussed in [9]–[11], and the extended second-order neighborhood scatterers are investigated in [12]. In addition to the above two structures, a more generalized clustered model with the logistic Gaussian kernel is introduced in [13] to obtain an improved imaging performance.

One drawback of the existing algorithms is that they may fail when the measurements suffer multiplicative phase noise,

which is induced due to atmosphere turbulence or navigational errors [14], [15]. The SAR sensor transmits pulses to the ground at points corresponding to equal angular increments on its flight path, and the error for each pulse is different. Therefore, the most widely encountered phase errors are in the cross-range direction [14]. If these errors are not well compensated, the resulting target image of the sparsity-driven algorithms would be substantially blurred. To overcome this imperfection, the sparsity-driven autofocus algorithms have been developed recently. For example, the alternating regularization approaches [14]–[16] and sparse Bayesian inference [17] have been applied to realize high-resolution target imaging and phase error correction simultaneously. In particular, the autofocus sparse Bayesian learning (AFSBL) method exploits uncertainty information during iterations, and performs better than the deterministic regularization method. Although the sparsity-driven autofocus methods are effective in some cases, they may cause the undesirable over-shrinkage problems. The reason for this is that the sparsity-driven approach only seeks the sparse solution. As a result, the weak scatterers may not be well-preserved and the back ground noise may not be properly shrunk with a simple sparsity constraint. To enhance the performance of sparsity-driven approach, the concept of structured autofocus sparse Bayesian learning (SAFSBL) is proposed in [18] by exploiting the correlation of the scatterers. Note that the above algorithms do not impose any prior information and/or constraint on phase error. As a result, they may lead to the undesirable phase ambiguity as mentioned in [14].

The drawbacks of the existing methods motivate us to develop a novel adaptive SAR imaging and autofocus algorithm in this paper. Our algorithm incorporates the structured prior knowledge about phase noise and the clustered sparse constraints on the scene in a Bayesian framework. More specifically, a pattern coupled Gaussian prior is introduced in our method to characterize the dependencies among neighboring scattering coefficients, thus making the cumbersome pattern selecting procedure (indispensable in [9]–[12]) unnecessary. In the proposed method, the Markov random fields (MRF) prior [19]–[21] is assumed on the support of the target scene to enforce the nonzero or zero scatterers to cluster in a spatial consistent manner. More importantly, we utilize the multivariate Von Mises model [22] to capture the spatial fluctuations along the aperture positions among phase errors. To surmount the difficulty of the calculation of posterior, the mean-field variational Bayesian expectation-maximization (VBEM) method [23]–[25] is utilized to jointly estimate the hidden variables and the model parameters. In addition, the convergence of the proposed method can be readily diagnosed by commonly used stopping criterion. Since the inherent target correlation characteristics are exploited and the uncertain structures on phase errors are introduced, the devised method can obtain enhanced performance on imaging and autofocus, as presented later in the simulation part.

The rest of this paper is organized as follows. In Section II, the phase-corrupted SAR observation model is described. The problem is formulated by a Bayesian model in Section III, and the subsequent Bayesian inference technique is derived in Section IV. Synthetic and practical data experimental results are given in Section V to demonstrate the effectiveness of the proposed algorithm. Finally, Section VI concludes the paper.

The following mathematical notations are used throughout this paper. We denote the vectors by bold lowercase letters, e.g., $\mathbf{b}$ and its mth element b_m, and the matrices by bold uppercase letters, e.g., $\mathbf{A}$ and its mth column $\mathbf{a}_m$ as well as its (m, n)th entry $a_{m,n}$. In particular, we use $\mathbf{I}$ to denote the identity matrix. $(\cdot)^{-1}, (\cdot)^T, (\cdot)^*$ stand for the inverse, transpose and conjugate transpose, respectively. $\mathbf{x} \odot \mathbf{y}$ represents the Hadamard product operator of vectors $\mathbf{x}$ and $\mathbf{y}$. diag$\{\cdot\}$ and blkdiag$\{\cdot\}$ stand for the diagonal matrix and the block diagonal matrix, respectively. We use arg$(\cdot)$ to denote the argument of a complex number, $\Re(\cdot)$ to represent the real part and $j \triangleq \sqrt{-1}$ to stand for the imaginary unit.

II. PROBLEM STATEMENT

The ground-plane geometry for spotlight-mode SAR is shown in Fig. 1. Data are collected by utilizing a radar sensor traversing a flight path and pointing at a fixed ground patch. At points corresponding to equal angular increments, high-bandwidth pulses are transmitted and returns from the ground patch of radius R_0 are then received and processed to form an image of the complex reflectivity field, denoted as $r(x, y)$. Using high-frequency assumption [26], the overall response of a complex scene is well approximated as a superposition of a set of the scene's differential scatterers. Consider that the radar transmits a time-limited linear frequency modulated signal as $\gamma(t) = e^{j(2\pi f_0 t + \alpha t^2)}$, where f_0 is the carrier frequency and α is the chirp rate limited to time $-T/2 < t < T/2$. After pre-processing steps of downconversion and matched filtering, the measurement signal can be described
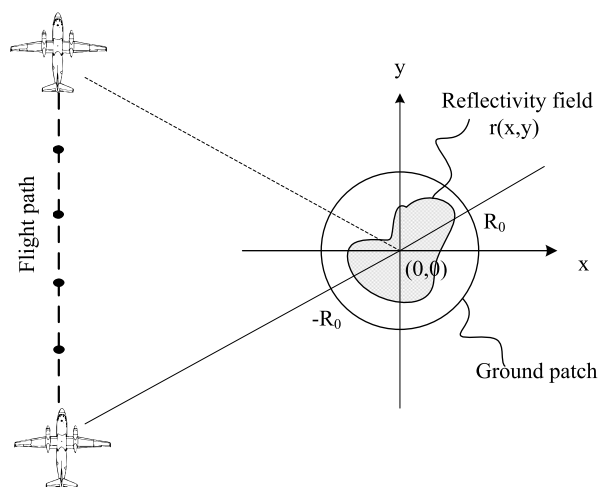

FIGURE 1. The ground-plane geometry for spotlight-mode SAR.

as

$$w(t, \phi) = \int \int_{x^2+y^2 \leq R_0^2} r(x, y)e^{-jk(t)(x\cos\phi + y\sin\phi)}dxdy$$

where ϕ is aspect angle, $k(t) = \frac{2}{c}(2\pi f_0 + 2\alpha(t - \frac{2R_c}{c}))$ is the spatial frequency variable with R_c representing the distance from radar to the center of the ground patch $(0, 0)$ and c denoting the speed of light. We can see that taking samples in effective time interval is tantamount to acquiring measurements over a frequency band. Assuming that the scene under surveillance consists of M point scatterers, the received signal can be written as

$$w(k, \phi) = \sum_{m=1}^{M} r(x_m, y_m)e^{-jk(x_m\cos\phi + y_m\sin\phi)} \quad (1)$$

where $r(x_m, y_m)$ represents the scattering coefficient of the scatterer located at (x_m, y_m). We suppose the measurements are at F discrete frequencies and K discrete aspect angles. Taking the multiplicative phase noise and additive noise into account, (1) can be expressed in matrix formation as

$$\mathbf{w} = \boldsymbol{\Theta}\boldsymbol{\Psi}r + \varepsilon \quad (2)$$

in which $\mathbf{w} \in \mathbb{C}^L$ is the received data with $L = FK$. $\boldsymbol{\Theta} = \text{blkdiag}\{e^{j\theta_1}\mathbf{I}_F, e^{j\theta_2}\mathbf{I}_F, ..., e^{j\theta_K}\mathbf{I}_F\}$ plays the role of the multiplicative phase noise. It is seen that the phase error is different for every aperture position which means that it affects the reconstructed image along the cross-range. $\boldsymbol{\Psi} \in \mathbb{C}^{L \times M}$ is the two-dimensional (2D) Fourier transform operator, $\mathbf{r} \in \mathbb{C}^M$ is the complex SAR imagery to be recovered, and $\boldsymbol{\varepsilon} \in \mathbb{C}^L$ denotes the additive noise.

In the literature, model (2) has been already considered in autofocus for radar imaging problems without any prior on the phase noise [15], [17], [18]. Here we look for a more complex and informative model enforcing uncertain structures on the phase errors. Due to the circular nature of the phase errors, it naturally leads to directional statistics [27]. Among them, the most popular one is probably the multivariate Von Mises distribution, defined as

$$p(\boldsymbol{\theta}) = \frac{1}{B(\boldsymbol{\varrho}, \Delta)} \exp\left(\boldsymbol{\varrho}^T \mathbf{g}(\boldsymbol{\theta}, \boldsymbol{\mu}) - \mathbf{s}(\boldsymbol{\theta}, \boldsymbol{\mu})^T \Delta \mathbf{s}(\boldsymbol{\theta}, \boldsymbol{\mu}) \right.$$
$$\left. - \mathbf{g}(\boldsymbol{\theta}, \boldsymbol{\mu})^T \Delta \mathbf{g}(\boldsymbol{\theta}, \boldsymbol{\mu})\right) \quad (3)$$

where $\boldsymbol{\mu}$ represents the mean direction, $\boldsymbol{\varrho}$ is the concentration parameter and $B(\boldsymbol{\varrho}, \Delta)$ is a normalizing constant. Functions $\mathbf{g}$ and $\mathbf{s}$ are defined by $g_k(\boldsymbol{\theta}, \boldsymbol{\mu}) = \cos(\theta_k - \mu_k)$ and $s_k(\boldsymbol{\theta}, \boldsymbol{\mu}) = \sin(\theta_k - \mu_k)$, respectively. The matrix Δ is real-symmetric with zeros on its diagonal and captures dependencies between phase errors. Without loss of generality, we assume in the sequel that $\boldsymbol{\mu} = [0, 0, ..., 0]^T \triangleq \mathbf{0}_K$. In this paper, we consider the particular case where only the first two subdiagonals of Δ are non-zero. Considering a small variance of the phases $\boldsymbol{\theta}$, straightforward calculus leads to [22]

$$p(\boldsymbol{\theta}) \simeq \frac{1}{B(\boldsymbol{\varrho}, \Delta)} \exp\left(-\sum_k ((\Omega_\theta^{-1})_{k,k}\theta_k^2 \right.$$

$$\left. + 2(\Omega_\theta^{-1})_{k,(k-1)}\theta_k\theta_{k-1} + o(\theta_k^2))\right) \quad (4)$$

where $\forall(k, i) \in \{1, ..., K\}^2$, $\boldsymbol{\Omega}_\theta^{-1}$ is linked to the parameters Δ and $\boldsymbol{\varrho}$ through

$$(\Omega_\theta^{-1})_{k,i} = \begin{cases} \Delta_{k,i}, & \text{if } i \neq k \\ \frac{\varrho_k}{2} - \sum_{j \neq k} \Delta_{k,j}, & \text{if } i = k. \end{cases} \quad (5)$$

Inspired by the idea in [28], expression (4) can be directly identified to a Markov chain such as

$$p(\boldsymbol{\theta}) = \prod_{k=2}^{K} p(\theta_k|\theta_{k-1})p(\theta_1) \quad (6)$$

where $p(\theta_k|\theta_{k-1}) = \mathcal{N}(\beta_0\theta_{k-1}, \beta^{-1})$, $\forall k \in \{2, ..., K\}$, $\theta_1 \sim \mathcal{N}(0, \beta^{-1})$ and $\beta_0 \in \mathbb{R}_+$, provided that

$$\Omega_\theta^{-1} = \frac{\beta}{2}\begin{bmatrix} 1+\beta_0^2 & -\beta_0 & 0 & \cdots & 0 \\ -\beta_0 & 1+\beta_0^2 & -\beta_0 & \ddots & \vdots \\ 0 & -\beta_0 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & 1+\beta_0^2 & -\beta_0 \\ 0 & \cdots & 0 & -\beta_0 & 1 \end{bmatrix}.$$

In this case, according to formula (5), we have $\varrho_k = (\beta_0 - 1)^2\beta$ for $k \in \{2, ..., K - 1\}$, $\varrho_1 = (\beta_0^2 - \beta_0 + 1)\beta$ and $\varrho_K = (1 - \beta_0)\beta$. If $|k - i| = 1$, the connection between phase errors can be denoted as $\Delta_{k,i} = -\beta_0\beta/2$. Otherwise, we have $\Delta_{k,i} = 0$. From a practical point of view, this phase Markov model allows us to describe spatial fluctuations of the phase errors along the aperture positions, and the strength of the fluctuations is related to the value of parameter β.

III. BAYESIAN FORMULATION

We address the problem of reconstructing SAR image $\mathbf{r}$ from noisy phase-corrupted data $\mathbf{w}$. To that end, we first place this problem into a Bayesian framework by defining suitable additional prior distributions on the unknown quantities $\boldsymbol{\varepsilon}$, $\mathbf{r}$ and $\boldsymbol{\theta}$. Suppose that $\boldsymbol{\varepsilon}$ follows complex white Gaussian distribution with zero mean and precision τ, then the likelihood of the measurement vector $\mathbf{w}$ follows:

$$p(\mathbf{w}|\boldsymbol{\Theta}\boldsymbol{\Psi}\mathbf{r}, \tau) = \mathcal{CN}(\mathbf{w}|\boldsymbol{\Theta}\boldsymbol{\Psi}\mathbf{r}, \tau^{-1}\mathbf{I}). \quad (7)$$

To facilitate the inference of the noise precision, a Gamma prior is placed on τ with parameters c and d, conjugate to Gaussian distribution. Then the distribution for τ is expressed as

$$p(\tau) = \mathcal{G}(\tau|c, d). \quad (8)$$

To account for the clustered sparsity of the coefficient vector $\mathbf{r}$, we utilize the fact that the sparsity patterns of neighboring coefficients are statistically dependent. We assume that $\mathbf{r} = \mathbf{h} \odot \mathbf{d}$ to separate the signal support $\mathbf{d}$ from the coefficient vector. The support indicator encodes which dictionary elements are activated for phase history. Similar to [29], a pattern coupled hierarchical prior is adopted for the sparse coefficient vector $\mathbf{h}$ to encourage the spatial continuity of the target

$$\ln q(\boldsymbol{\theta}) \propto \ln p(\boldsymbol{\theta}) + \langle \ln p(\mathbf{w}|\Theta\Psi\mathbf{r}, \tau) \rangle_{q(\mathbf{r})q(\tau)}$$

$$\propto \ln p(\boldsymbol{\theta}) + \left\langle 2\tau \sum_{l=1}^{L} \Re \left\{ w_l \left(\sum_{m=1}^{M} r_m^* \psi_{l,m}^* \right) e^{-j\theta_l} \right\} \right\rangle_{q(\mathbf{r})q(\tau)} \qquad (15)$$

scene. Specifically, the Gaussian prior for each coefficient not only involves its own hyperparameter, but also its immediate neighboring hyperparameters. More precisely, a prior over $\mathbf{h}$ is given by

$$p(\mathbf{h}|\boldsymbol{\sigma}) = \prod_{m=1}^{M} \mathcal{CN}(h_m|0, \eta_m^{-1}) \qquad (9)$$

with

$$\eta_m = \sigma_m + \iota \sum_{n \in \mathcal{O}_m} \sigma_n \qquad (10)$$

where $\mathcal{O}_m$ is the set of the neighbors of m, and $\iota \in (0, 1]$ is a parameter indicating the pattern relevance between h_m and its neighbors h_n. Then we use a Gamma distribution as hyperprior over the hyperparameter $\boldsymbol{\sigma}$, i.e.,

$$p(\boldsymbol{\sigma}) = \prod_{m=1}^{M} \mathcal{G}(\sigma_m|a, b). \qquad (11)$$

On the other hand, it is well known that the MRF based statistical dependency model has the capability to capture the correlations of the investigated elements in the support with its neighbors. Therefore, we adopt a special case of MRF prior, the Ising model [21], to encourage the clustered property of the support of the target scene $\mathbf{d}$. The MRF prior distribution is given by

$$p(\mathbf{d}'; \boldsymbol{\chi}) = \frac{1}{Z(\boldsymbol{\chi})} \exp \left\{ \sum_{m=1}^{M} (\chi_{0m} d_m' + \chi_m \sum_{n \in \mathcal{O}_m} d_m' d_n') \right\}$$

where $Z(\boldsymbol{\chi})$ is the partition function for normalizing the distribution and $\boldsymbol{\chi}$ is the model parameter vector. The contribution of the interaction between two elements $\{d_m', d_n'\}$ is controlled by the parameter $\chi_m > 0$, while the parameter χ_{0m} controls the spatial sparsity of scattered field. $\mathbf{d}' \in \{-1, 1\}^M$ is a binary vector linked to the support indicator $\mathbf{d}$ with $d_m' = -1$ implying $d_m = 0$ and $d_m' = 1$ indicating $d_m = 1$. It is easy to verify that the MRF encourages d_m' and its neighbors sharing the same state since the consistence of adjacent states will result in higher probability. Note that by exploiting the pattern coupled Gaussian distribution and MRF prior, a general model for clustered sparsity of the target scene is established instead of the cumbersome model selection in [9]–[12].

Finally, the probabilistic model describing the behavior of $\boldsymbol{\theta}$ is given by (6). As mentioned in the previous section, this simple model accounts for local variations of the position errors due to atmosphere turbulence or navigational errors

along the aperture positions. Based on the probabilistic models presented, the posterior of the hidden variables given the measurement vector $\mathbf{w}$ follows:

$$p(\boldsymbol{\theta}, \mathbf{r}, \boldsymbol{\sigma}, \tau|\mathbf{w}; \boldsymbol{\chi}, \beta) \propto p(\mathbf{w}|\Theta\Psi\mathbf{r}, \tau) \cdot p(\mathbf{h}|\boldsymbol{\sigma}) \cdot p(\mathbf{d}; \boldsymbol{\chi})$$
$$\cdot p(\boldsymbol{\theta}; \beta) \cdot p(\boldsymbol{\sigma}) \cdot p(\tau). \qquad (12)$$

The intractability of the marginal distribution, $p(\mathbf{w})$, results in the unattainability of the full posterior. As a result, we employ mean-field VBEM method to infer the hidden variables $\mathcal{H} = \{\boldsymbol{\theta}, \mathbf{r}, \boldsymbol{\sigma}, \tau\}$ and the model parameters including the MRF parameter $\boldsymbol{\chi}$ as well as the phase Markov model parameter β.

IV. THE PROPOSED PROCEDURE

In this section, the updating rules for the hidden variables and the model parameters are derived. A two-stage procedure is involved in VBEM. In the VB-E stage, an approximate posterior, denoted by $q(\mathcal{H})$, is calculated to approximate the intractable posterior $p(\boldsymbol{\theta}, \mathbf{r}, \boldsymbol{\sigma}, \tau|\mathbf{w}; \boldsymbol{\chi}, \beta)$. Then the updating rules for all unknown hidden variables in $\mathcal{H}$ are obtained based on $q(\mathcal{H})$. The rules for updating the MRF and phase Markov model parameters are given in the maximization stage.

A. VB-E STAGE

In this stage, it is assumed that $q(\mathcal{H})$ is factorizable with the mean-field approximation, i.e.,

$$q(\boldsymbol{\theta}, \mathbf{r}, \boldsymbol{\sigma}, \tau) = q(\boldsymbol{\theta})q(\tau) \prod_{m=1}^{M} q(r_m)q(\sigma_m). \qquad (13)$$

The optimal distribution of each element in the hidden variables $\mathcal{H}$ follows [23]:

$$\ln q(\mathcal{H}_z) = \langle \ln p(\mathcal{H}, \mathbf{w}) \rangle_{q(\mathcal{H} \backslash \mathcal{H}_z)} \qquad (14)$$

where $\langle \cdot \rangle_{q(\mathcal{H} \backslash \mathcal{H}_z)}$ represents the expectation with respect to $q(\mathcal{H} \backslash \mathcal{H}_z)$ and $\mathcal{H} \backslash \mathcal{H}_z$ denotes the set $\mathcal{H}$ without its zth element $\mathcal{H}_z$. The updated estimation for hidden variable $\mathcal{H}_z$ is represented by $\tilde{\mathcal{H}}_z$.

1) PHASE ERROR ESTIMATION

The estimate of the phase error $\boldsymbol{\theta}$ is obtained by computing the term of $\ln q(\boldsymbol{\theta})$ that depends on $\boldsymbol{\theta}$, as formulated in (15), as shown at the top of this page. Then defining $\xi_l \overset{\Delta}{=} w_l \left(\sum_{m=1}^{M} \tilde{r}_m^* \psi_{l,m}^* \right)$, we have

$$\ln q(\boldsymbol{\theta}) \propto \ln p(\boldsymbol{\theta}) + 2\tilde{\tau} \sum_{l=1}^{L} \Re \left\{ \xi_l e^{-j\theta_l} \right\}$$

$$\ln q(r_m) \propto \langle \ln p(\mathbf{w}|\Theta\Psi r, \tau)p(\mathbf{h}|\sigma)p(\mathbf{d}; \boldsymbol{\chi})\rangle_{q(\boldsymbol{\theta})q(\mathbf{r}\backslash r_m)q(\sigma)q(\tau)}$$

$$\propto -\tilde{\tau}\left[(\boldsymbol{\psi}_m^*\boldsymbol{\psi}_m + \tilde{\eta}_m)r_m^2 - 2\sum_{l=1}^{L}\Re\left\{w_l^*\psi_{l,m}\langle e^{j\theta_l}\rangle r_m\right\} + 2\Re\left\{\sum_{i\neq m}\tilde{r}_i^*\boldsymbol{\psi}_i^*\boldsymbol{\psi}_m r_m\right\}\right] + \langle \ln p(\mathbf{d}; \boldsymbol{\chi})\rangle_{q(\boldsymbol{\theta})q(\mathbf{r}\backslash r_m)q(\sigma)q(\tau)} \quad (21)$$

$$\propto \ln p(\boldsymbol{\theta}) + 2\tilde{\tau}\sum_{l=1}^{L}\left\{|\xi_l|\cos\left[\arg(\xi_l) - \theta_l\right]\right\}. \quad (16)$$

To go further into the clarification of the procedure, we employ the approximation that a Von Mises is equivalent to a Gaussian distribution for small values of $\kappa > 0$ [27], i.e.,

$$\frac{\exp\left\{\frac{1}{\kappa}\cos(e_1 - e_2)\right\}}{2\pi I_0\left(\frac{1}{\kappa}\right)} \simeq \frac{\exp\left\{-\frac{1}{2\kappa}(e_1 - e_2)^2\right\}}{\sqrt{2\pi\kappa}}. \quad (17)$$

Under this condition, (16) can be rewritten as

$$\ln q(\boldsymbol{\theta}) \propto \ln p(\boldsymbol{\theta}) - \frac{1}{2}\left[\boldsymbol{\theta} - \arg(\boldsymbol{\xi})\right]^T \Omega^{-1}\left[\boldsymbol{\theta} - \arg(\boldsymbol{\xi})\right] \quad (18)$$

where $\boldsymbol{\xi} \triangleq [\xi_1, \xi_2, ..., \xi_L]^T$ and $\Omega^{-1} = \text{diag}(2\tilde{\tau}|\boldsymbol{\xi}|)$. Finally, replacing $\ln p(\boldsymbol{\theta})$ by its expression (6), we obtain the estimate of $q(\boldsymbol{\theta})$ as a Gaussian distribution, i.e.,

$$q(\boldsymbol{\theta}) = \mathcal{N}\left(\boldsymbol{\mu}_\theta, \Lambda_\theta\right) \quad (19)$$

where $\Lambda_\theta = (\Omega^{-1} + \Omega_\theta^{-1})^{-1}$ and $\boldsymbol{\mu}_\theta = \Lambda_\theta\left[\Omega^{-1}\arg(\boldsymbol{\xi})\right]$. Consequently, $\boldsymbol{\theta}$ is updated by its mean $\boldsymbol{\mu}_\theta$

$$\tilde{\boldsymbol{\theta}} = \boldsymbol{\mu}_\theta. \quad (20)$$

2) SPARSE SCATTERERING COEFFICIENT ESTIMATION

The posterior of the sparse scatterering coefficient r_m can be expressed in (21), as shown at the top of this page, where

$$\tilde{\eta}_m = \tilde{\sigma}_m + \iota\sum_{n\in\mathcal{O}_m}\tilde{\sigma}_n. \quad (22)$$

More detailed derivation of (21) can be found in Appendix A. According to [28], we have

$$\langle e^{j\theta_l}\rangle \simeq \frac{I_1(1/\Lambda_{\theta_l})}{I_0(1/\Lambda_{\theta_l})}e^{j\mu_{\theta_l}} \quad (23)$$

where $I_0(\cdot)$ and $I_1(\cdot)$ stand for the modified Bessel of the first kind of order 0 and order 1, respectively. Recalling that $q(r_m) = q(h_m, d_m)$, we obtain

$$q(h_m|d_m) = \mathcal{CN}\left(\mu_{h_m}(d_m), \Lambda_{h_m}(d_m)\right) \quad (24)$$

and

$$q(d_m) \propto \sqrt{\Lambda_{h_m}(d_m)}\exp\left\{\frac{[\mu_{h_m}(d_m)]^2}{\Lambda_{h_m}(d_m)}\right\}$$
$$\cdot \exp\left\{\chi_{0m}d_m + \chi_m d_m\sum_{n\in\mathcal{O}_m}q(d_n = 1)\right\} \quad (25)$$

with

$$\Lambda_{h_m}(d_m) = \left(\tilde{\eta}_m + d_m\tilde{\tau}\boldsymbol{\psi}_m^*\boldsymbol{\psi}_m\right)^{-1} \quad (26a)$$
$$\mu_{h_m}(d_m) = \Lambda_{h_m}(d_m) \cdot d_m\tilde{\tau}\boldsymbol{\psi}_m^*\tilde{\boldsymbol{\rho}}_m \quad (26b)$$
$$\tilde{\boldsymbol{\rho}}_m = \bar{\mathbf{w}} - \sum_{i\neq m}q(d_i = 1)\mu_{h_i}(d_i = 1)\boldsymbol{\psi}_i \quad (26c)$$
$$\bar{\mathbf{w}} = \left[w_l e^{-j\mu_{\theta_l}}\frac{I_1\left(1/\Lambda_{\theta_l}\right)}{I_0\left(1/\Lambda_{\theta_l}\right)}\right]_{l=\{1,...,L\}}. \quad (26d)$$

With the above notations, we can obtain the updating rule for r_m as

$$\tilde{r}_m = q(d_m = 1)\mu_{h_m}(d_m = 1). \quad (27)$$

3) COEFFICIENT PRECISION PARAMETER ESTIMATION

Due to the conjugacy of Gamma and Gaussian distribution, the approximated posterior of the coefficient precision parameter σ_m can be obtained by

$$\ln q(\sigma_m) \propto \langle \ln p(\sigma) + \ln p(h_m|\sigma_m)\rangle_{q(\sigma\backslash\sigma_m)q(h_m)}$$
$$\propto (a - 1)\ln\sigma_m - b\sigma_m$$
$$+ \left(\ln\tilde{\eta}_m + \sum_{n\in\mathcal{O}_m}\ln\tilde{\eta}_n\right) - \upsilon_m\sigma_m$$
$$\propto (a - 1)\ln\sigma_m - (b - \delta_m + \upsilon_m)\sigma_m \quad (28)$$

where

$$\upsilon_m = \langle h_m^2\rangle + \iota\sum_{n\in\mathcal{O}_m}\langle h_n^2\rangle \quad (29a)$$
$$\langle h_m^2\rangle = \mu_{h_m}^2(d_m = 1) + \Lambda_{h_m}(d_m = 1). \quad (29b)$$

Moreover, denoting by δ_m the first-order item of the Taylor expansion of the term $(\ln\tilde{\eta}_m + \sum_{n\in\mathcal{O}_m}\ln\tilde{\eta}_n)$, one can express δ_m as

$$\delta_m = \left(\iota\sum_{n\in\mathcal{O}_m}\tilde{\sigma}_n\right)^{-1} + \sum_{n\in\mathcal{O}_m}\iota\left(\tilde{\sigma}_n + \iota\sum_{i\in\mathcal{O}_n\backslash m}\tilde{\sigma}_i\right)^{-1}.$$

On this basis, $q(\sigma_m)$ can be approximately treated as a Gamma distribution $\mathcal{G}(\sigma_m|\hat{a}_m, \hat{b}_m)$ with $\hat{a}_m = a$ and $\hat{b}_m = b - \delta_m + \upsilon_m$. As a result, the updating rule for σ_m can be given by

$$\tilde{\sigma}_m = \hat{a}_m/\hat{b}_m. \quad (30)$$

4) NOISE PRECISION ESTIMATION

The variational distribution of the noise precision τ is similarly computed as

$$\ln q(\tau) \propto \langle \ln p(\mathbf{w}|\Theta\Psi r, \tau)p(\tau)\rangle_{q(\boldsymbol{\theta})q(\mathbf{r})} = \ln\mathcal{G}(\tau|\hat{c}, \hat{d})$$

$$\hat{d} = d + \mathbf{w}^*\mathbf{w} - 2\sum_{m=1}^{M}\Re\left\{\langle d_m h_m\rangle\bar{\mathbf{w}}^*\boldsymbol{\psi}_m\right\} + \sum_{m=1}^{M}\sum_{i=1}^{M}(\langle d_m d_i h_m^* h_i\rangle\boldsymbol{\psi}_m^*\boldsymbol{\psi}_i) \tag{31}$$

$$\langle d_m d_i h_m^* h_i\rangle = \begin{cases} q(d_m=1)\left\{[\mu_{h_m}(d_m=1)]^2 + \Lambda_{h_m}(d_m=1)\right\}, & \text{if } m=i \\ q(d_m=1)q(d_i=1)[\mu_{h_m}(d_m=1)]^*\mu_{h_i}(d_i=1), & \text{if } m \neq i \end{cases} \tag{33}$$

$$\chi = \arg\max_{\chi}\langle\ln p(\mathbf{d}';\chi)\rangle_{q(\mathbf{d}')} \tag{35a}$$

$$= \arg\max_{\chi} -\ln Z(\chi) + \left\langle\sum_{m=1}^{M}\left(\chi_{0m}d_m' + \chi_m\sum_{n\in\mathcal{O}_m}d_m'd_n'\right)\right\rangle_{q(d_m')} \tag{35b}$$

$$= \arg\max_{\chi}\sum_{m=1}^{M}\left\langle -\ln\sum_{d_m'\in\{-1,1\}}e^{\chi_{0m}d_m' + \chi_m\sum_{n\in\mathcal{O}_m}d_m'd_n'} + \chi_{0m}d_m' + \chi_m\sum_{n\in\mathcal{O}_m}d_m'd_n'\right\rangle_{q(d_m')}. \tag{35c}$$

$$\beta = \arg\max_{\beta}\langle\ln p(\boldsymbol{\theta};\beta)\rangle_{q(\boldsymbol{\theta})} \tag{38a}$$

$$= \arg\max_{\beta}\frac{K}{2}\ln\beta - \left\langle\sum_{k}\left((\hat{\Omega}_\theta^{-1})_{k,k}\theta_k^2 + 2(\hat{\Omega}_\theta^{-1})_{k,(k-1)}\theta_k\theta_{k-1}\right)\right\rangle_{q(\theta_k)}\cdot\frac{\beta}{2} \tag{38b}$$

where $\hat{c} = c + L$, and the expression of $\hat{d}$ can be found in (31), as shown at the top of this page. The term $\langle d_m h_m\rangle$ in (31) is expressed as

$$\langle d_m h_m\rangle = q(d_m=1)\mu_{h_m}(d_m=1) \tag{32}$$

and the expression of $\langle d_m d_i h_m^* h_i\rangle$ is defined in (33), as shown at the top of this page. Finally, τ is updated by

$$\tilde{\tau} = \hat{c}_m/\hat{d}_m. \tag{34}$$

B. VB-M STAGE
The parameters χ and β are estimated in the maximization stage of the VBEM.

1) MRF MODEL PARAMETER ESTIMATION
The maximization is intractable due to the difficulty in calculating $Z(\chi)$. We follow the idea of [12] by using a mean field assumption of $p(\mathbf{d}')$ to alleviate this problem. The parameter χ can be estimated by (35) on the top of this page. Then, χ_{0m}, χ_m can be approximately expressed as

$$[\chi_{0m}, \chi_m] = \arg\max_{\chi_{0m},\chi_m} -\ln\sum_{d_m'\in\{-1,1\}}e^{\chi_{0m}d_m' + \chi_m\sum_{n\in\mathcal{O}_m}d_m'\tilde{d}_n'}$$
$$+ \chi_{0m}\tilde{d}_m' + \chi_m\sum_{n\in\mathcal{O}_m}\tilde{d}_m'\tilde{d}_n'. \tag{36}$$

For simplicity, we assume that $\chi_{0m} = \delta$ is a constant for $m \in \{1, 2, ..., M\}$. Thus, by calculating the derivative of (36) with respect to χ_m and setting it to zero, we obtain the estimation

of $\tilde{\chi}_m$ as

$$\tilde{\chi}_m = \frac{\ln[(1+\tilde{d}_m')/(1-\tilde{d}_m')] - 2\delta}{2\sum_{n\in\mathcal{O}_m}\tilde{d}_n'}. \tag{37}$$

2) PHASE MARKOV MODEL PARAMETER ESTIMATION
The solution to β can be found by maximizing the expected log-likelihood function, as formulated in (38) on the top of this page. In (38), $\hat{\Omega}_\theta^{-1}$ is given by

$$\hat{\Omega}_\theta^{-1} = \begin{bmatrix} 1+\beta_0^2 & -\beta_0 & 0 & \cdots & 0 \\ -\beta_0 & 1+\beta_0^2 & -\beta_0 & \ddots & \vdots \\ 0 & -\beta_0 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & 1+\beta_0^2 & -\beta_0 \\ 0 & \cdots & 0 & -\beta_0 & 1 \end{bmatrix} \tag{39}$$

with $\beta_0 \in \mathbb{R}_+$ being a constant. Then compute the derivative of (38) with respect to the parameter β and set it to zero, we obtain the following formula to update the parameter β:

$$\tilde{\beta} = \frac{K}{\boldsymbol{\mu}_\theta^T\hat{\Omega}_\theta^{-1}\boldsymbol{\mu}_\theta}. \tag{40}$$

The iterative procedure is repeated until the convergence criterion $\|\mathbf{h}^{(j)} - \mathbf{h}^{(j-1)}\|_2/\|\mathbf{h}^{(j)}\|_2 < \rho_0$ is achieved, where j denotes the iterative index and $\rho_0 > 0$ is a user-selected parameter (e.g. 10^{-2}), or the number of iterations exceeds a predefined maximum iteration number J_{iter}. Since the VBEM is guaranteed to converge, our algorithm always converges to a global or a local optimum [12]. To make the above

descriptions clear, we summarize the proposed method in Algorithm 1.

Algorithm 1 The Proposed Bayesian Clustered Sparsity-Driven SAR Imaging and Autofocus Algorithm

---

Input: The measurement data $\mathbf{w}$ and corresponding dictionary $\mathbf{\Psi}$.

1: Initialize the latent variables $\mathcal{H} = (\boldsymbol{\theta}, \mathbf{h}, \mathbf{d}, \boldsymbol{\sigma}, \tau)$, MRF and phase Markov model parameters $\boldsymbol{\chi}$ and β, hyper-parameters and variables $a, b, c, d, \rho_0, J_{iter}$; set iteration counter $j = 0$;
2: **while** $\|\mathbf{h}^{(j)} - \mathbf{h}^{(j-1)}\|_2 / \|\mathbf{h}^{(j)}\|_2 < \rho_0$ or $j < J_{iter}$ **do**
3:       Hidden variables updating:
4:           Update $\boldsymbol{\theta}$ by (20);
5:           Update r_m by (27);
6:           Update σ_m by (30);
7:           Update τ by (34);
8:       Model parameters updating:
9:           Update χ_m by (37);
10:         Update β by (40);
11: **end while**

Output: The final focused image $\mathbf{r}$ and estimated phase error $\boldsymbol{\theta}$.

V. EXPERIMENTAL ANALYSIS

Two datasets are used to show the merits of our proposed algorithm in this section. The first dataset is synthetic scene and the second one is the simplified tank dataset generated by an electromagnetic scattering calculation software. We compare the proposed algorithm with three benchmark methods: the polar format algorithm (PFA) [30], the regularization based sparsity-driven autofocus (SDA) algorithm [14] and the structured autofocus sparse Bayesian learning (SAFSBL) algorithm [18].

TABLE 1. System parameters for simulated-data.

Parameter	Value
Carrier frequency f_0	9 GHz
Bandwidth $\triangle f$	0.8 GHz
Angular range $\triangle \phi$	5°
Range resolution Rr	0.188 m
Azimuth resolution Ar	0.191 m

TABLE 2. Performance evaluation against VPN.

VPN	Metrics	SDA	SAFSBL	Proposed
	PMSE	0.2593	0.1220	0.0591
0.6	Corr	0.4056	0.5071	0.7716
	Entropy	3.8944	2.9597	2.0109
	PMSE	0.1994	0.1182	0.0380
0.3	Corr	0.5524	0.6247	0.7933
	Entropy	3.4367	2.8273	1.6945
	PMSE	0.1141	0.0699	0.0189
0.1	Corr	0.7886	0.8081	0.8926
	Entropy	2.7990	1.8751	1.0633

A. PERFORMANCE METRICS

In this section, the performance metrics used on the experiments are outlined to verify the imaging and autofocus abilities of the proposed approach. The mean square error (MSE) of phase error estimation is given by

$$\text{PMSE} = \|\arg(\hat{\mathbf{\Theta}}) - \arg(\mathbf{\Theta})\|_F^2 / K \qquad (41)$$

where $\hat{\mathbf{\Theta}}$ is the estimated phase error and $\mathbf{\Theta}$ is true phase error. To quantitatively evaluate the imaging and autofocus performance, we use the measures of correlation and entropy. Supposing that the obtained image by different methods is denoted as $\hat{\mathbf{r}}$, the correlation is then defined as

$$\text{Corr} = \frac{|\hat{\mathbf{r}}^T \mathbf{r}|}{\|\hat{\mathbf{r}}\|_2 \cdot \|\mathbf{r}\|_2} \qquad (42)$$

which measures the similarity of the recovered image with the true one. The higher the correlation value that a method can provide, the better the method preserves the information of the target. The entropy of the image is given by

$$\text{Entropy} = -\sum \mathbf{p} \cdot \log_2 \mathbf{p} \qquad (43)$$

where $\mathbf{p}$ is the histogram of the recovered gray level image. The entropy evaluates the concentration quality of the obtained SAR image. The more the image is focused, the smaller its entropy is.

B. SIMULATED DATA EXPERIMENTS

The radar system parameters of the synthetic dataset are given in Table 1. The simulated scene is shown in Fig. 2(a) consisting of isolated scatterers and continuous areas, where their amplitudes are set to be 5 and phases are random. The

TABLE 3. Performance evaluation against SNR.

SNR	Metrics	SDA	SAFSBL	Proposed
	PMSE	0.0731	0.0601	0.0472
0 dB	Corr	0.4679	0.5781	0.6372
	Entropy	3.3061	2.8955	2.3394
	PMSE	0.0629	0.0457	0.0283
5 dB	Corr	0.6533	0.7407	0.8518
	Entropy	3.2116	2.6992	1.7483
	PMSE	0.0445	0.0318	0.0281
10 dB	Corr	0.8976	0.9095	0.9514
	Entropy	2.0409	1.0518	0.8226
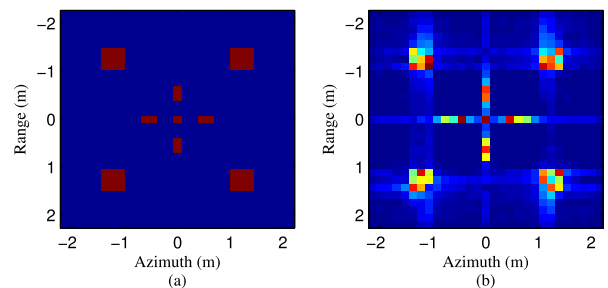
FIGURE 2. Simulated data. (a) Original scene; (b) SAR image obtained by the PFA algorithm.

image obtained by PFA, free of additive and multiplicative noises, is given in Fig. 2(b). In the subsequent experiments, the received signal is corrupted by both the additive Gaussian noise and the phase noise obeying the phase Markov model. We represent the variance of phase noise (VPN) by reciprocal of the phase Markov model parameter β (i.e., β^{-1}). To evaluate the performance of the proposed algorithm, it is necessary to perform experiments by using different VPN under different conditions of signal-to-noise ratio (SNR).

Fig. 3 shows the imaging results obtained by using the four methods in terms of various VPNs when SNR is 15 dB. We set the phase Markov model parameter β_0 as 0.8, β^{-1} as 0.6 ,0.3 and 0.1, respectively. It is seen that the PFA reconstructions contain substantial blurring in cross-range dimension. The three sparse imaging results which have the capability of denoising, are better than those of PFA. However, the target is not well focused using SDA method because it makes use of merely sparsity. Due to utilization of clustered sparsity, both the SAFSBL and the devised algorithm can preserve the weak scatterers in the target region and remove the noise outside target region. It can be validated that our proposed method maintains the most focused target imaging results because of the multivariate Von Mises prior introduced to enforce the connections exist between the phase errors, instead of the uninformative prior on phase errors in SAFSBL. In Table 2, quantitative results, including PMSE, correlation and entropy,

are given to evaluate the imaging and autofocus performance. The results are averaged over 50 independent realizations at different VPN regimes. As indicated in Table 2, the PFA cases exhibit the worst performance since it doesn't have the ability of phase noise mitigation. The results obtained by the deterministic SDA method are worse than the Bayesian inference techniques. This is because in the iterative procedure, the estimation error propagation phenomenon inevitably exists in regularization problems which have been solved by MAP estimation, while the Bayesian formulation of the clustered sparsity and the phase structural correlation in the devised scheme can conveniently preserve the estimation uncertainty information during iterations, which can alleviate effect of error propagation. More concretely, the VBEM technique can help to refine the reconstruction accuracy, since the fully posterior distributions are provided instead of the point estimates.

The image estimates obtained by four methods with different SNRs are presented and compared in Fig. 4. The VPN is set to be 0.05. In general, all of these compared algorithms except PFA can achieve reasonable results when SNR is 10 dB, where the obtained target image is well concentrated. However, when SNR decreases, degraded imagery results appear, particularly those obtained by the regularization based method. This may attributes to the fact that the regularized parameters are selected by experience in SDA, while all the necessary parameters are directly learned from
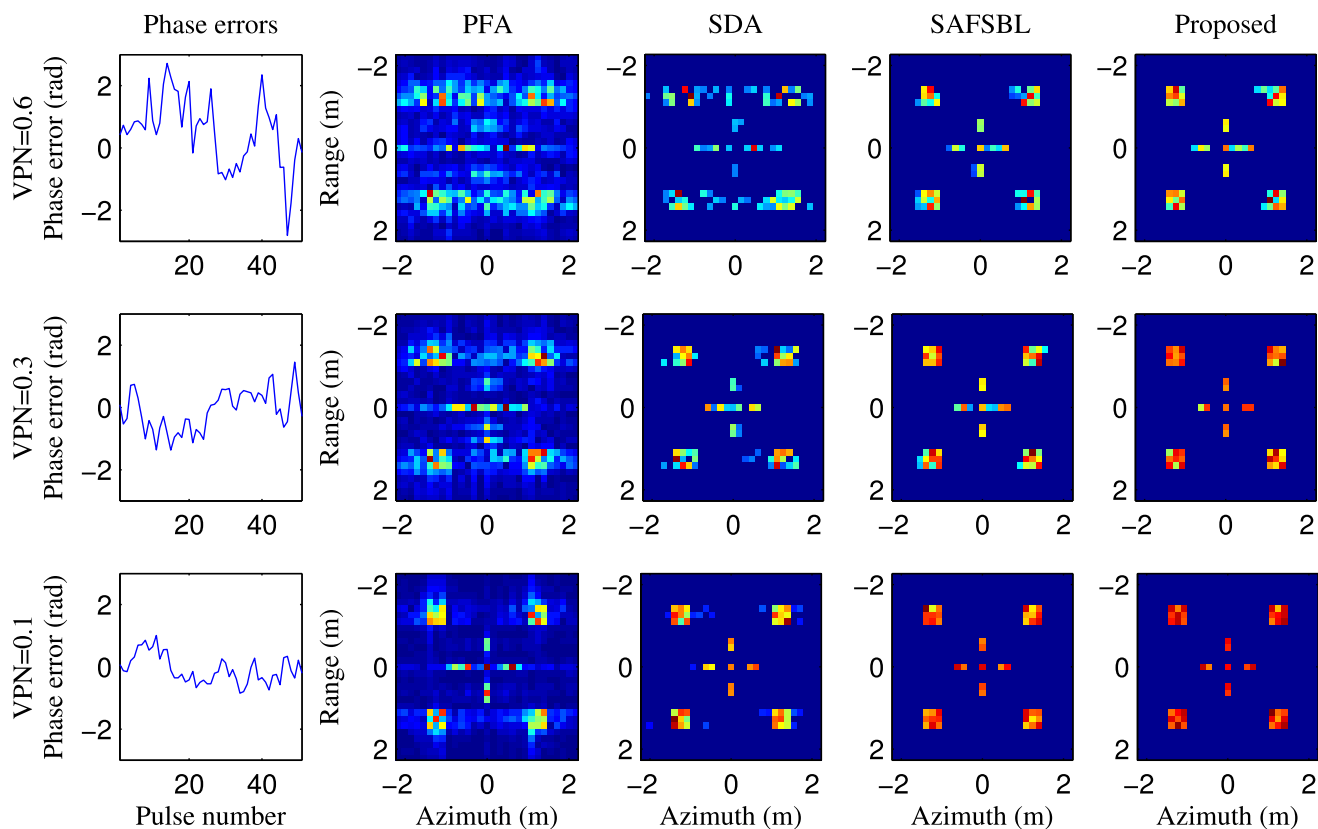

FIGURE 3. Comparison of SAR images of synthetic scene against different VPNs and four algorithms.
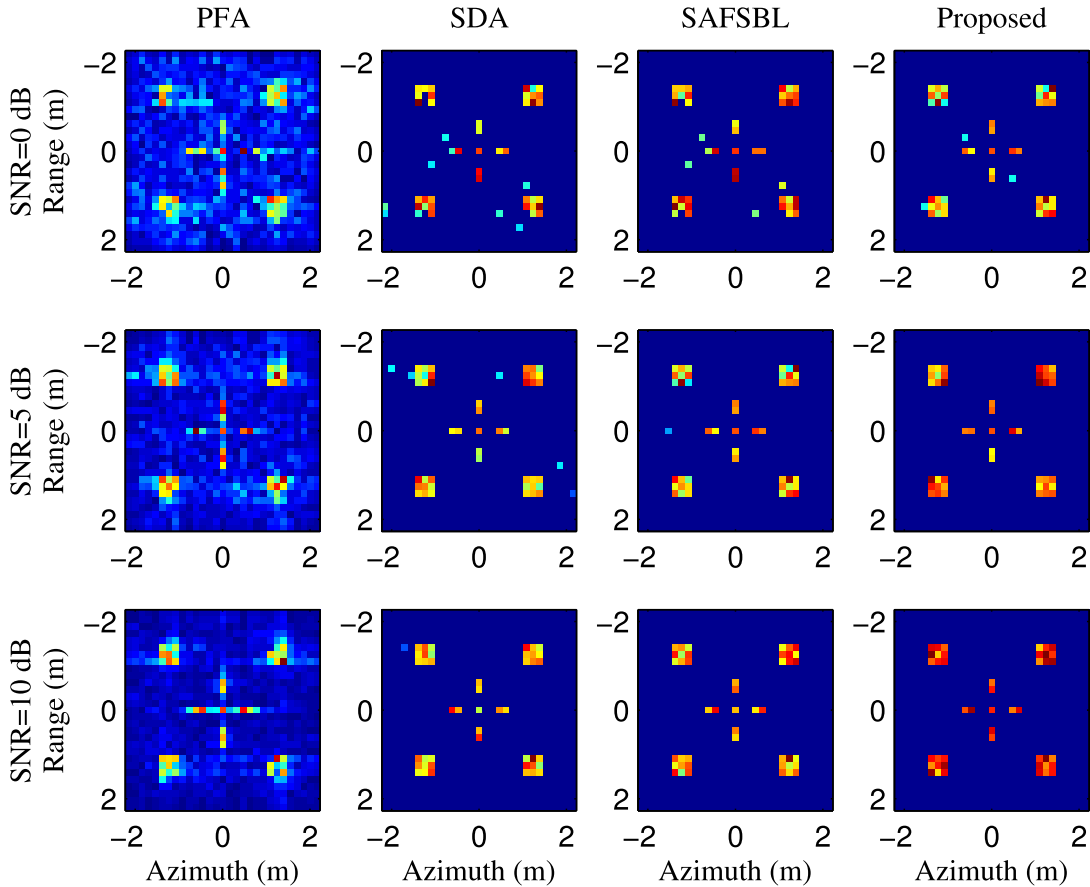
FIGURE 4. Comparison of SAR images of synthetic scene against different SNRs and four algorithms.
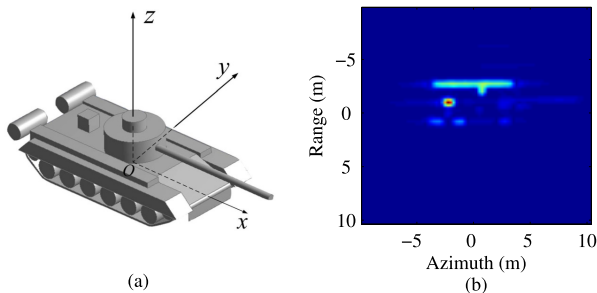

FIGURE 5. Measured data. (a) Photograph of the simplified tank; (b) SAR image obtained by the PFA algorithm.

data under Bayesian inference framework. Table 3 shows the influence of SNR on the performance of the algorithms. A Monte Carlo experiment with 50 times under each SNR is adopted. It's quite obvious that higher SNR generally leads to improved estimation. For all the tested SNRs, it is seen that the performance of the proposed algorithm highly distinguishes from that of the benchmark ones under the same conditions, indicating better performance of suppressing noise and preserving target information.

C. SIMPLIFIED TANK DATA EXPERIMENTS

In the second example, we present reconstruction results based on the simplified tank dataset. The photograph of the

simplified tank model is shown in Fig. 5(a) and the image obtained by PFA without additional noise is given in Fig. 5(b). The simplified tank model is sized by $9.51\text{m} \times 3.56\text{m} \times 2.57\text{m}$ (length $\times$ width $\times$ height). The phase history data is collected at frequencies $8.5 - 9.5\text{GHz}$ in 0.01GHz intervals and over the azimuth range $[86.6°, 93.4°]$ in $0.05°$. The incident plane wave is $30°$ in zenith angle and is horizontally polarized.

In Fig. 6, the performances are evaluated in terms of different VPNs under SNR of 10 dB. The phase error is added into the data according to the phase Markov model by setting β^{-1} as 0.5 and 0.1, respectively. As observed in Fig. 6, the PFA approach almost fails to obtain a meaningful image, and the SDA method can hardly obtain a simplified tank profile with a limited number of true scatterers. The SAFSBL method shows a relatively reasonable profile of the simplified tank, but the image blurring effects still exist in these images, and some of the true scattering points (e.g., the barrel of the simplified tank) are not recovered. In contrast, the proposed algorithm can effectively estimate the phase errors and achieve better concentration results by recovering more true scattering points.

Fig. 7 presents the imaging results obtained by different methods under SNRs of -5 dB and 5 dB. The phase error is added using $\beta^{-1} = 0.1$. When the SNR is as low as -5 dB, the images obtained by PFA are too noisy and the
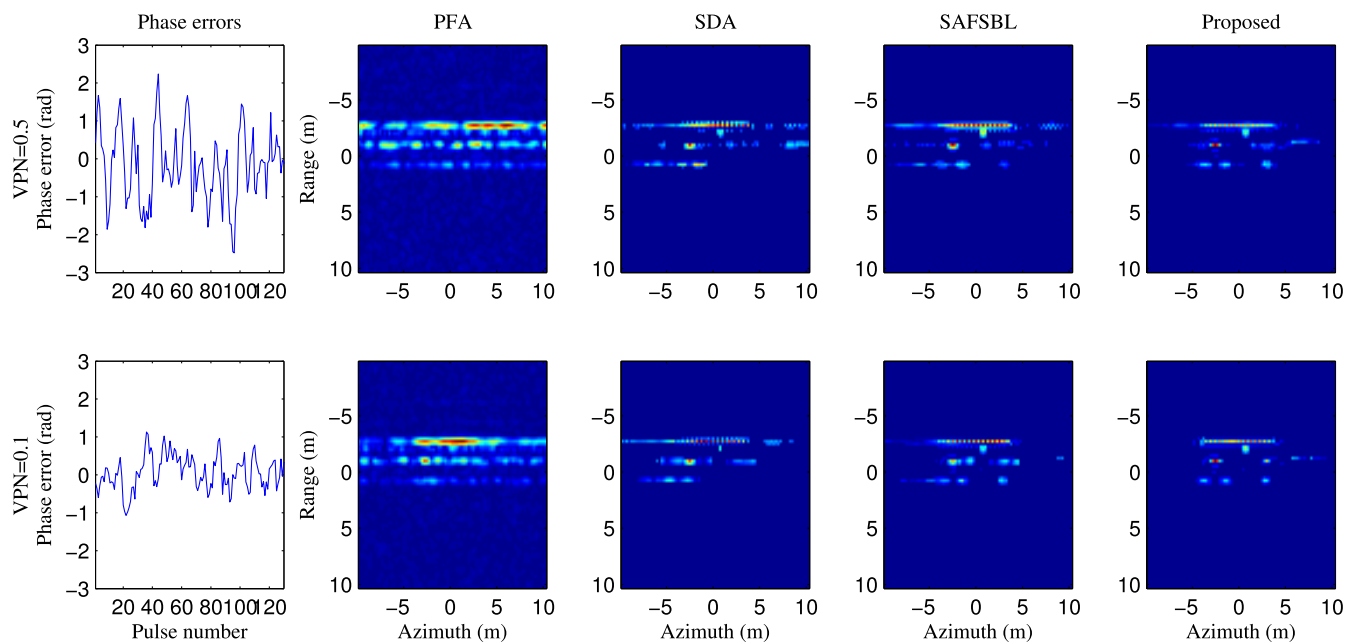
FIGURE 6. Comparison of SAR images of simplified tank against different VPNs and four algorithms.
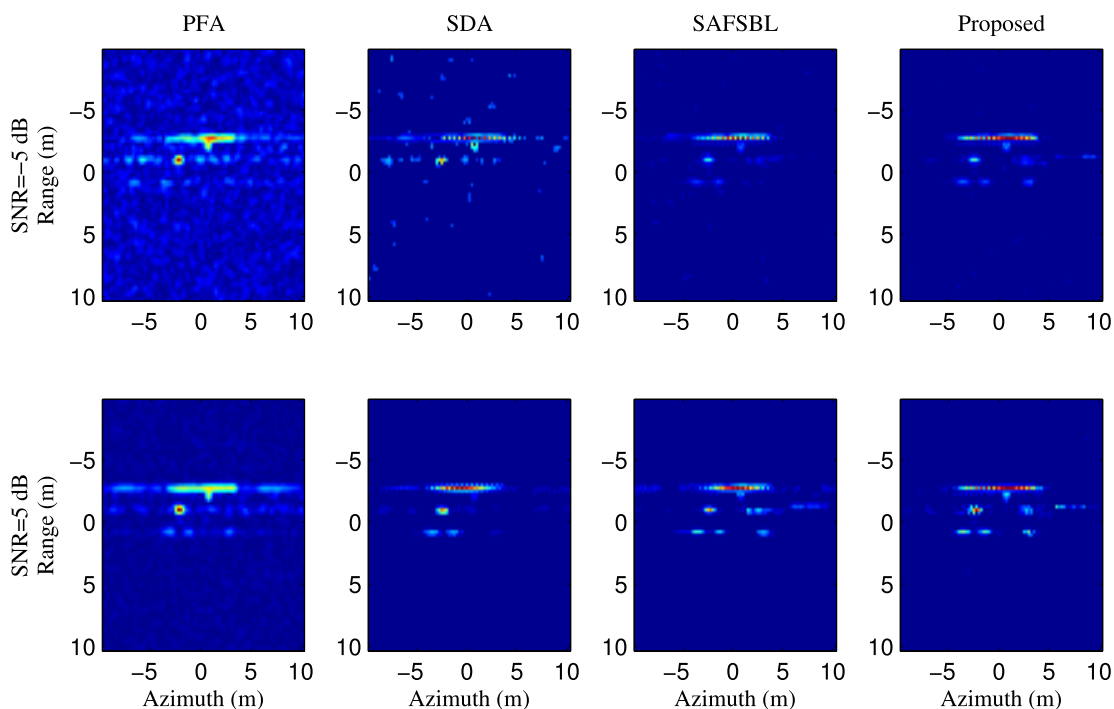

FIGURE 7. Comparison of SAR images of simplified tank against different SNRs and four algorithms.

target can hardly be noticed. The results of SDA appear to have too many artificial points outside the target region and suffers from loss of scattering coefficients. SAFSBL removes most of the undesirable additive noise but is still not well focused. The proposed method demonstrates promising imagery results with more concentrated target region and less artifacts around the region. When the SNR increases, it is

observed that the results of all methods are improved. More importantly, it is shown that our proposed one can preserve weak scatterers and give much more concentrated images by making good use of the cluster-promoting feature and phase structural dependence. With these comparisons, it is reasonable to conclude that the devised method is able to obtain superior imagery results from different SNR scenarios.

$$
\begin{aligned}
\ln q(r_m) &\propto \langle \ln p(\mathbf{w}|\Theta\Psi\mathbf{r}, \tau)p(\mathbf{h}|\boldsymbol{\sigma})p(\mathbf{d}; \boldsymbol{\chi})\rangle_{q(\boldsymbol{\theta})q(\mathbf{r}\backslash r_m)q(\boldsymbol{\sigma})q(\tau)} \\
&\propto \left\langle -\tau||\mathbf{w} - \Theta\Psi_{\neg m}\mathbf{r}_{\neg m} - \Theta\boldsymbol{\psi}_m r_m||^2 + \ln p(\mathbf{h}|\boldsymbol{\sigma})p(\mathbf{d}; \boldsymbol{\chi})\right\rangle_{q(\boldsymbol{\theta})q(\mathbf{r}\backslash r_m)q(\boldsymbol{\sigma})q(\tau)} \\
&\propto \left\langle -\tau\left[(\boldsymbol{\psi}_m^*\boldsymbol{\psi}_m + \eta_m)r_m^2 - 2\Re\left\{(\mathbf{w} - \Theta\Psi_{\neg m}\mathbf{r}_{\neg m})^*\Theta\boldsymbol{\psi}_m r_m\right\}\right] + \ln p(\mathbf{d}; \boldsymbol{\chi})\right\rangle_{q(\boldsymbol{\theta})q(\mathbf{r}\backslash r_m)q(\boldsymbol{\sigma})q(\tau)} \\
&\propto -\tilde{\tau}\left[(\boldsymbol{\psi}_m^*\boldsymbol{\psi}_m + \tilde{\eta}_m)r_m^2 - 2\sum_{l=1}^{L}\Re\left\{w_l^*\psi_{l,m}\langle e^{j\theta_l}\rangle r_m\right\} + 2\Re\left\{\sum_{i\neq m}\tilde{r}_i^*\boldsymbol{\psi}_i^*\boldsymbol{\psi}_m r_m\right\}\right] + \langle \ln p(\mathbf{d}; \boldsymbol{\chi})\rangle_{q(\boldsymbol{\theta})q(\mathbf{r}\backslash r_m)q(\boldsymbol{\sigma})q(\tau)} \quad (44)
\end{aligned}
$$

VI. CONCLUSIONS

In this paper, an adaptive sparsity-driven technique for joint SAR imaging and phase error correction has been proposed by considering the spatial continuity of the scene in structured phase-noisy environments. The SAR imaging and autofocus problem is reformulated in a Bayesian framework exploiting two types of priors. One is the MRF model as well as the neighboring correlated precision priors on the spatial continuity structure, and the other is the particular case of Von Mises model on phase noise. By incorporating the informative priors, the devised approach is able to provide more concentrated target image and preserve the weak scatterers in the target region. In addition, all the necessary parameters are directly learned from data via VBEM method, avoiding a tedious parameter-tuning procedure. Representative simulations have been presented to verify the effectiveness and superiority of the proposed algorithm under various scenarios. Future work will focus on the improvement of the computational burden. By employing program optimization and parallel computing technology, the program can be further accelerated for real-time applications.

APPENDIX A
DERIVATION OF (21)

The variational distribution $\ln q(r_m)$ satisfies (44), as shown at the top of this page, where $\Psi_{\neg m}$ results from the dictionary Ψ after removing its mth column $\boldsymbol{\psi}_m$, $\mathbf{r}_{\neg m}$ results from $\mathbf{r}$ after the exclusion of its mth element. This completes the derivation of the result in (21).

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their careful review and constructive comments.

REFERENCES

[1] L. Yang, L. Zhao, Z. Song, and G. Bi, "Sparsity-driven sar imaging for highly maneuvering ground target by the combination of time-frequency analysis and parametric Bayesian learning," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 4, pp. 1443–1455, Apr. 2017.

[2] G. Xu, Y. Liu, and M. Xing, "Multi-channel synthetic aperture radar imaging of ground moving targets using compressive sensing," *IEEE Access*, vol. 6, pp. 66134–66142, 2018.

[3] J. Xu, Y. Pi, and Z. Cao, "Bayesian compressive sensing in synthetic aperture radar imaging," *IET Radar, Sonar Navigat.*, vol. 6, no. 1, pp. 2–8, Jan. 2012.

[4] Y. Li, S. Fan, J. Yang, J. Xiong, X. Cheng, G. Gui, and H. Sari, "Musai-l$_{(1/2)}$: Multiple sub-wavelet-dictionaries based adaptive iteratively weighted l$_{(1/2)}$ regularization algorithm," *IEEE Access*, vol. 6, pp. 16795–16805, 2018.

[5] Y. Li, L. Yun, X. Cheng, Z. Xiao, and G. Guan, "Nonconvex penalized regularization for robust sparse recovery in the presence of *sαs* noise," *IEEE Access*, vol. 6, pp. 25474–25485, 2018.

[6] G. Li, H. Zhang, X. Wang, and X.-G. Xia, "ISAR 2-D imaging of uniformly rotating targets via matching pursuit," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 48, no. 2, pp. 1838–1846, Apr. 2012.

[7] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM Rev.*, vol. 43, no. 1, pp. 129–159, 2001.

[8] G. Xu, M. Xing, L. Zhang, Y. Liu, and Y. Li, "Bayesian inverse synthetic aperture radar imaging," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 6, pp. 1150–1154, Nov. 2011.

[9] L. Wang, L. Zhao, G. Bi, C. Wan, and L. Yang, "Enhanced ISAR imaging by exploiting the continuity of the target scene," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 9, pp. 5736–5750, Sep. 2014.

[10] Q. Wu, Y. D. Zhang, M. G. Amin, and B. Himed, "Multi-task Bayesian compressive sensing exploiting intra-task dependency," *IEEE Signal Process. Lett.*, vol. 22, no. 4, pp. 430–434, Apr. 2015.

[11] Q. Wu, Y. D. Zhang, F. Ahmad, and M. G. Amin, "Compressive-sensing-based high-resolution polarimetric through-the-wall radar imaging exploiting target characteristics," *IEEE Antennas Wireless Propag. Lett.*, vol. 14, pp. 1043–1047, 2015.

[12] L. Wang, L. Zhao, G. Bi, and C. Wan, "Sparse representation-based ISAR imaging using Markov random fields," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 8, pp. 3941–3953, Aug. 2015.

[13] Q. Wu, Y. D. Zhang, M. G. Amin, and B. Himed, "High-resolution passive SAR imaging exploiting structured Bayesian compressive sensing," *IEEE J. Sel. Topics Signal Process.*, vol. 9, no. 8, pp. 1484–1497, Dec. 2015.

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

[15] N. Ö. Önhon and M. Çetin, "Joint sparsity-driven inversion and model error correction for radar imaging," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, Mar. 2010, pp. 1206–1209.

[16] S.-J. Wei, X.-L. Zhang, and J. Shi, "An autofocus approach for model error correction in compressed sensing SAR imaging," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2012, pp. 3987–3990.

[17] L. Zhao, L. Wang, G. Bi, and L. Yang, "An autofocus technique for high-resolution inverse synthetic aperture radar imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 10, pp. 6392–6403, Oct. 2014.

[18] L. Zhao, L. Wang, G. Bi, S. Li, L. Yang, and H. Zhang, "Structured sparsity-driven autofocus algorithm for high-resolution radar imagery," *Signal Process.*, vol. 125, pp. 376–388, Aug. 2016.

[19] T. Kasetkasem and P. K. Varshney, "Statistical characterization of clutter scenes based on a Markov random field model," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 39, no. 3, pp. 1035–1050, Jul. 2003.

[20] P. Yahampath, *Joint Source-Decoding Large Scale Sensor Networks Using Markov Random Field Models*. Amsterdam, The Netherlands: Elsevier, 2010.

[21] V. Cevher, M. Duarte, C. Hedge, and R. Baraniuk, "Sparse signal recovery using Markov random fields," in *Proc. Conf. Neural Inf. Process. Syst.*, 2008, pp. 257–264.

[22] A. Drémeau and A. Deleforge, "Phase retrieval with a multivariate von mises prior: From a Bayesian formulation to a lifting solution," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, Mar. 2017, pp. 4596–4600.

[23] D. G. Tzikas, A. C. Likas, and N. P. Galatsanos, "The variational approximation for Bayesian inference," *IEEE Signal Process. Mag.*, vol. 25, no. 6, pp. 131–146, Jan. 2008.

[24] K. E. Themelis, A. A. Rontogiannis, and K. D. Koutroumbas, "A variational Bayes framework for sparse adaptive estimation," *IEEE Trans. Signal Process.*, vol. 62, no. 18, pp. 4723–4736, Sep. 2014.

[25] A. Dremeau, C. Herzet, and L. Daudet, "Boltzmann machine and mean-field approximation for structured sparse decompositions," *IEEE Trans. Signal Process.*, vol. 60, no. 7, pp. 3425–3438, Jul. 2012.

[26] J. A. Jackson and R. L. Moses, "Synthetic aperture radar 3D feature extraction for arbitrary flight paths," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 48, no. 3, pp. 2065–2084, Jul. 2012.

[27] K. V. Mardia, *Directional Statistics*. Hoboken, NJ, USA: Wiley, 2000.

[28] A. Drémeau and C. Herzet, "Doa estimation in structured phase-noisy environments," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, Mar. 2017, pp. 3176–3180.

[29] J. Fang, L. Zhang, and H. Li, "Two-dimensional pattern-coupled sparse Bayesian learning via generalized approximate message passing," *IEEE Trans. Image Process.*, vol. 25, no. 6, pp. 2920–2930, Jun. 2016.

[30] S. Quegan, "Spotlight synthetic aperture radar: Signal processing algorithms," *J. Atmos. Sol.-Terr. Phys.*, vol. 59, no. 5, pp. 597–598, 1995.

YUE YANG (S'15) was born in Sichuan, China. She received the B.Eng. degree from the School of Electronic Engineering, University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2015, where she is currently pursuing the Ph.D. degree with the School of Information and Communication Engineering. Since 2019, she has been a Visiting Student with the Department of Electrical and Computer Engineering (ECE), National University of Singapore, Singapore. Her research interests include synthetic aperture radar imaging, sparse signal reconstruction, and statistical signal processing.
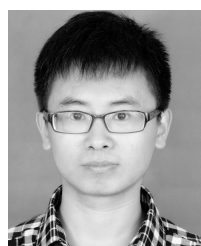
XUEJING ZHANG (S'17) was born in Hebei, China. He received the B.S. degree in electrical engineering from Huaqiao University, Xiamen, China, in 2011, and the M.S. degree in signal and information processing from Xidian University, Xi'an, China, in 2014. He is currently pursuing the Ph.D. degree in signal and information processing with the School of Information and Communication Engineering, University of Electronic Science and Technology of China (UESTC), Chengdu, China. Since 2017, he has been a Visiting Student with the University of Delaware, Newark, DE, USA. His research interests include array signal processing and wireless communications.

GUAN GUI (M'11–SM'17) received the Dr.Eng. degree in information and communication engineering from the University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2012. From 2009 to 2014, he was with the Wireless Signal Processing and Network Laboratory (Prof. Adachi Laboratory), Department of Communications Engineering, Graduate School of Engineering, Tohoku University as a Research Assistant and a Postdoctoral Research Fellow, respectively. From 2014 to 2015, he was an Assistant Professor with the Department of Electronics and Information System, Akita Prefectural University. Since 2015, he has been a Professor with the Nanjing University of Posts and Telecommunications (NJUPT), Nanjing, China. He has published more than 200 international peer-reviewed journal/conference papers. His research interests include deep learning, compressive sensing, and advanced wireless techniques.

He received the Member and Global Activities Contributions Award in the IEEE ComSoc and seven Best Paper Awards, i.e., ICEICT 2019, ADHIP 2018, CSPS 2018, ICNC 2018, ICC 2017, ICC 2014, and VTC 2014-Spring. He was selected as/for Jiangsu Specially-Appointed Professor, in 2016, Jiangsu High-level Innovation and Entrepreneurial Talent, in 2016, Jiangsu Six Top Talent, in 2018, the Nanjing Youth Award, in 2018. He was an Editor of *Security and Communication Networks*, from 2012 to 2016. He has been the Editor of the IEEE Transactions on Vehicular Technology, since 2017, IEEE Access, since 2018, the *KSII Transactions on Internet and Information Systems*, since 2017, and the *Journal of Communications*, since 2019, and the Editor-in-Chief of the *EAI Transactions on Artificial Intelligence*, since 2018.

QUN WAN (M'04) received the B.Sc. degree in electronic engineering from Nanjing University, in 1993, and the M.Sc. and Ph.D. degrees in electronic engineering from the University of Electronic Science and Technology of China (UESTC), in 1996 and 2001, respectively. From 2001 to 2003, he was a Postdoctoral Researcher with the Department of Electronic Engineering, Tsinghua University. Since 2004, he has been a Professor with the Department of Electronic Engineering, UESTC. He is currently the Director of the Joint Research Laboratory of Array Signal Processing and also an Associate Dean of the School of Electronic Engineering. His research interests include direction finding, radio localization, and signal processing based on information criterion.

● ● ●