

Research Article

Complex Multisnapshot Sparse Bayesian Learning for Offgrid DOA Estimation

Qinghua Liu ¹, Yuanxin He,¹ Kai Ding ², and Quanmin Xie^{3,4}

¹State and Local Joint Engineering Research Centre for Satellite Navigation and Location Service, Guilin University of Electronic Technology, Guilin 541004, China

²Science and Technology on Near-Surface Detection Laboratory, Wuxi 214035, China

³Ordnance NCO Academy, Army Engineering University, Wuhan 430075, China

⁴Hubei (Wuhan) Explosions and Blasting Technology Institute, Jiangnan University, Wuhan 430056, China

Correspondence should be addressed to Kai Ding; winfast113@sina.com

Received 19 August 2021; Revised 4 December 2021; Accepted 25 January 2022; Published 28 February 2022

Academic Editor: Ping Li

Copyright © 2022 Qinghua Liu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Direction of arrival (DOA) estimation has recently been developed based on sparse signal reconstruction (SSR). Sparse Bayesian learning (SBL) is a typical method of SSR. In SBL, the two-layer hierarchical model in Gaussian scale mixtures (GSMs) has been used to model sparsity-inducing priors. However, this model is mainly applied to real-valued signal models. In order to apply SBL to complex-valued signal models, a general class of sparsity-inducing priors is proposed for complex-valued signal models by complex Gaussian scale mixtures (CGSMs), and the special cases correspond to complex versions of several classical priors are provided, which is helpful to analyze the connections with different modeling methods. In addition, the expression of the SBL form of the real- and complex-valued model is unified by parameter values, which makes it possible to generalize and improve the properties of the SBL methods. Finally, the SBL complex-valued form is applied to the offgrid DOA estimation complex-valued model, and the performance between different sparsity-inducing priors is compared. Theoretical analysis and simulation results show that the proposed algorithm can effectively process complex-valued signal models and has lower algorithm complexity.

1. Introduction

Direction of arrival (DOA) estimation of spatial signals is an important content of array signal processing research. In recent years, compressive sensing (CS) [1] and sparse signal reconstruction (SSR) [2] have been introduced into the field of DOA estimation owing to their ultra-high resolution, good robustness to noise, and low dependence on the number of snapshots. One representative method is the L1-SVD algorithm proposed in [3], which uses L1-norm to construct an SSR model, reduces the complexity of the algorithm by singular value decomposition (SVD) under multisnapshots, and has a high resolution for the correlated sources as well. Sparse Bayesian learning (SBL) [4–7] or Bayesian compressed sensing (BCS) [8, 9] is another representative method for the SSR in CS. In SBL, the signal recovery problem is formulated from a Bayesian perspective,

while the sparsity information is exploited by assuming a sparse prior for the signal of interest. One merit of SBL is its flexibility in modeling sparse signals, which can improve the sparsity of its solution [10]. Therefore, research work on DOA estimation based on SBL has been gaining momentum in recent years [11, 12]. While these methods have shown advantages over the conventional ones, however, there exists still difficulties in practical applications.

For these SSR methods mentioned above, which are called ongrid methods, true DOAs are assumed to lie exactly on a set of fixed sampling grids, and the existing sparse representation techniques can be directly applied. When the assumption that the true DOAs are located on a sampling grid fails, the performance of on-grid methods deteriorates due to the problem of mismatch. Although the offgrid (the distance from the true DOA to the nearest grid) error can be reduced by a dense sampling grid, the computational

complexity will increase significantly. Furthermore, a dense grid may lead to the correlation between two steering vectors becoming high, which means the performance cannot be improved by the theory of CS.

Accordingly, an offgrid DOA estimation model was first studied in [13], where the true DOAs are no longer constrained in the sampling grid, and they proposed a sparsity cognizant total least-squares (S-TLS) method for the perturbed compressive sensing under sparsity constraints. It has been shown in [13] that the S-TLS can yield a MAP (maximum a posteriori) optimal estimate if the matrix perturbation caused by the basis mismatch is Gaussian. But the Gaussian condition cannot be satisfied in all DOA estimations. In [14], a new algorithm named offgrid sparse Bayesian inference (OGSBI) is proposed for the offgrid case. Firstly, the true DOA steering vector is approximated by a first-order Taylor expansion to handle the off-grid problem, and it is assumed to have a uniform distribution for off-grid parameters. Based on the OGSBI method, a linear interpolation between two adjacent grids is adopted in [15] to approximate the true DOA steering vector, and they proposed a perturbed sparse Bayesian learning (PSBL) algorithm to solve the offgrid DOA estimation problem.

Both ongrid and offgrid methods are grid-based methods. Another method to deal with the offgrid problem is the gridless method [16–18], which directly operates in the continuous domain so that can completely avoid the grid mismatch problem. However, this kind of method needs strong theoretical guarantees. In addition, it can only be applied to uniform or sparse linear arrays [19].

The SBL algorithm mainly deals with real-valued signals. In practical applications, such as DOA estimation, the experimental signals are often in complex-valued form. The results obtained in the real-valued model cannot be directly applied to the complex-valued model. In order to apply SBL to the complex-valued model, in [20–22], the real and imaginary parts of the complex-valued signal were separated to construct a new real-valued model, and then the SBL real-valued model was applied for processing. But in [20, 21], the real and imaginary parts of the complex-valued signal were treated as two independent variables, and the different hyperparameters were specified by ignoring the correlation between the real and imaginary parts of the complex-valued signal, which has obvious theoretical defects, and the reconstruction performance of the algorithm is low. It was improved in [22], where the performance of the algorithm was improved by assigning the same hyperparameters to the real and imaginary parts of the complex-valued signal by taking advantage of the fact that they have the same sparse structures. However, by decomposing the complex-valued signal into real and imaginary parts, the dimensionality of the signal is doubled, and the sensing matrix is tripled, so the complexity of the algorithm will increase significantly.

In this paper, the extension of SBL from the real-valued model to the complex-valued model in conjunction with the DOA estimation complex-valued model is considered, and the complex multisnapshot SBL (CSBL) offgrid DOA estimation algorithm is proposed. The expression of the SBL form is unified under the two models by parameter values,

which enables us to generalize and improve the properties of the SBL methods. Theoretical analysis and simulation results show that the proposed algorithm can effectively process complex-valued signals with lower complexity.

A comment on notation: we use boldface lowercase letters for vectors and boldface uppercase letters for matrices. $(\cdot)^T$, $(\cdot)^{-1}$, $(\cdot)^*$, and $(\cdot)^H$ denote the transpose, inverse, conjugation, and conjugation-transpose operations, respectively. $\|\cdot\|_F$ and $\|\cdot\|_2$ are the F-norm and 2-norm of a matrix, respectively. $\text{tr}(\cdot)$ denotes the trace operation of a matrix. \odot describes Hadamard product operator. $p(\cdot)$ represents the probability density function of random variables.

2. Offgrid DOA Estimation Model

Consider a uniform linear array (ULA) with N isotropic sensors and interelement spacing Δd as shown in Figure 1. Suppose that K far-field narrowband signals of frequency f_0 illuminate on this ULA from directions $\theta_k, k = 1, 2, \dots, K$. By defining the first sensor of the array as the reference, the received signals can be expressed as

$$Y = A(\theta)S + E, \quad (1)$$

where $Y = [y(1), y(2), \dots, y(M)]$, $S = [s(1), s(2), \dots, s(M)]$, $E = [e(1), e(2), \dots, e(M)]$, and M is the number of snapshots. $y(m) = [y_1(m), y_2(m), \dots, y_N(m)]^T \in \mathbb{C}^{N \times 1}$, $A(\theta) = [a(\theta_1), a(\theta_2), \dots, a(\theta_K)] \in \mathbb{C}^{N \times K}$, $a(\theta_k) = [1, \exp\{j2\pi f_0 \Delta d \sin \theta_k / c\}, \dots, \exp\{j(N-1)2\pi f_0 \Delta d \sin \theta_k / c\}]^T \in \mathbb{C}^{N \times K}$, $s(m) = [s_1(m), s_2(m), \dots, s_K(m)]^T \in \mathbb{C}^{K \times 1}$, $e(m) = [e_1(m), e_2(m), \dots, e_N(m)]^T \in \mathbb{C}^{N \times 1}$ is the additive complex white Gaussian noise vector with zero mean and variance $\sigma^2 I$. When $M > 1$, equation (3) is the case of multiple measurement vector (MMV).

Let $\tilde{\theta} = \{\tilde{\theta}_1, \tilde{\theta}_2, \dots, \tilde{\theta}_L\}$ be a fixed sampling grid in the DOA range $[-90^\circ, 90^\circ]$, where L denotes the grid number and typically satisfies $L \gg N > K$. Without loss of generality, let $\tilde{\theta}$ be a uniform grid with a grid interval $r = \tilde{\theta}_2 - \tilde{\theta}_1 \propto L^{-1}$. In practical situations, the target locations are randomly distributed in space, which leads to the offgrid problem.

Suppose $\theta_k \notin \tilde{\theta} (k = 1, 2, \dots, K)$ and that $\tilde{\theta}_l (\tilde{l} \in 1, 2, \dots, L)$ is the nearest grid to θ_k , which satisfies $|\theta_k - \tilde{\theta}_l| \leq r/2$. So, the steering vector $a(\theta_k)$ can be approximated by the first-order Taylor [23],

$$a(\theta_k) \approx a(\tilde{\theta}_l) + b(\tilde{\theta}_l)(\theta_k - \tilde{\theta}_l), \quad (2)$$

where $b(\tilde{\theta}_l) = a'(\tilde{\theta}_l)$ is the first-order derivative of $a(\tilde{\theta}_l)$. Let $A(\tilde{\theta}) = [a(\tilde{\theta}_1), a(\tilde{\theta}_2), \dots, a(\tilde{\theta}_L)]$, $B(\tilde{\theta}) = [b(\tilde{\theta}_1), b(\tilde{\theta}_2), \dots, b(\tilde{\theta}_L)]$, and grid parameter $\beta_l = \theta_k - \tilde{\theta}_l, l = \tilde{l}, k = 1, 2, \dots, K$. By considering the approximation error into the noise, the offgrid DOA estimation sparse model can be expressed as

$$Y = \Phi(\beta)X + E, \quad (3)$$

where $\Phi(\beta) = A(\tilde{\theta}) + B(\tilde{\theta})\text{diag}(\beta)$ is abbreviated as Φ in the following paragraphs, $Y \in \mathbb{C}^{N \times M}$, $E \in \mathbb{C}^{N \times M}$, $X = [\mathbf{x}(1), \mathbf{x}(2), \dots, \mathbf{x}(M)] \in \mathbb{C}^{L \times M}$, $A(\tilde{\theta})$, $B(\tilde{\theta})$, and $\Phi(\beta) \in \mathbb{C}^{N \times L}$.

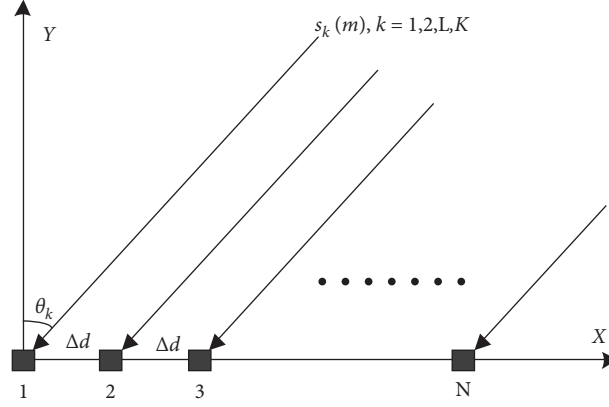


FIGURE 1: The configuration of a ULA.

In SBL, the sparse representation of ΦX is induced by designing a prior $p(\mathbf{x})$ for x . Instead of using the prior $p(\mathbf{x})$ directly, the SBL typically uses a two-layer hierarchical prior model, which involves a conditional prior $p(x|\alpha)$ and a hyperprior $p(\alpha)$ with the hyperparameters $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, \dots, \alpha_L]$, which controls the estimation accuracy and sparsity of x .

3. CSBL Offgrid DOA Estimation Algorithm

3.1. Two-Layer Hierarchical Prior Model. The extension of SBL from the real-valued model is considered to the complex-valued model, and the SBL form is unified, which can be applied to both real-valued and complex-valued models with different values of parameters. Assumed that the signals are independent at different snapshots, i.e., $\mathbf{x}(m)$, $m = 1, 2, \dots, M$ are independent of each other, there exist the following conditional prior equations [24],

$$p(X|\boldsymbol{\alpha}) = \prod_{m=1}^M p(\mathbf{x}(m)|\boldsymbol{\alpha}) = \prod_{m=1}^M \prod_{i=1}^L p(x_i(m)|\alpha_i). \quad (4)$$

or

$$p(X|\boldsymbol{\gamma}) = \prod_{m=1}^M p(\mathbf{x}(m)|\boldsymbol{\gamma}) = \prod_{m=1}^M \prod_{i=1}^L p(x_i(m)|\gamma_i), \quad (5)$$

where $p(x_i|\alpha_i) = (p/\pi\alpha_i)^p \exp\{-p|x_i|^2/\alpha_i\}$, and $p(x_i|\gamma_i) = (p\gamma_i/\pi)^p \exp\{-p\gamma_i|x_i|^2\}$. α_i is the variance of x_i , and $\gamma_i = 1/\alpha_i$ is the precision of x_i . The parameter $p = 1$ when x_i is complex, and $p = 1/2$ when x_i is real. Next, the hyperprior $p(\alpha_i)$ for variance α_i and hyperprior $p(\gamma_i)$ for precision γ_i are modeled, respectively.

The generalized inverse Gaussian (GIG) distribution is chosen [25, 26] as the hyperprior $p(\alpha_i)$ for variance α_i with the following expression with three parameters λ , a , and b .

$$p(\alpha|\lambda, b, a) = \prod_i GIG(\alpha_i|\lambda, b, a) = \prod_i \frac{(a/b)^{\lambda/2}}{2K_\lambda(\sqrt{ab})} \alpha_i^{\lambda-1} \exp\left\{-\frac{1}{2}\left(a\alpha_i + \frac{b}{\alpha_i}\right)\right\}, \quad (6)$$

where $K_\lambda(\cdot)$ denotes the modified Bessel function of the second kind of order λ , and the value range of parameters λ , a , and b are described later. The GIG distribution includes three distributions as special cases: the Gamma distribution, the inverse Gamma distribution, and the IG distribution

according to the different values of the three parameters shown in Figure 2.

For the hyperprior $p(\gamma_i)$, from the inverse relationship between precision and variance $\gamma_i = 1/\alpha_i$, the probability density function transformation formula [26] is used, and the hyperprior $p(\gamma_i)$ is given by

$$p(\gamma|\lambda, a, b) = \prod_i GIG(\gamma_i|-\lambda, a, b) = \prod_i \frac{(b/a)^{-\lambda/2}}{2K_{-\lambda}(\sqrt{ab})} \gamma_i^{-\lambda-1} \exp\left\{-\frac{1}{2}\left(b\gamma_i + \frac{a}{\gamma_i}\right)\right\}. \quad (7)$$

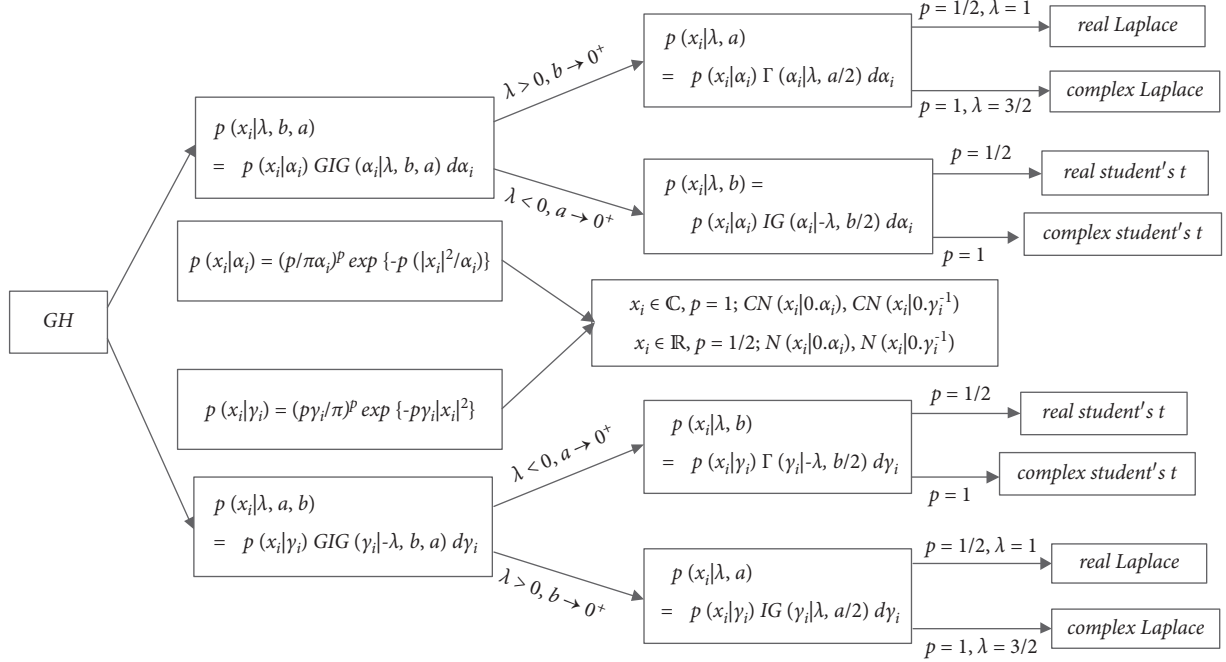


FIGURE 2: The relationship between different two-layer hierarchical priors.

In equations (6) and (7), the parameters satisfy the relationship as follows:

$$\begin{aligned}
 b \geq 0, a > 0 & \quad \text{if } \lambda > 0, \\
 b > 0, a > 0 & \quad \text{if } \lambda = 0, \\
 b > 0, a \geq 0 & \quad \text{if } \lambda < 0.
 \end{aligned} \tag{8}$$

The marginal distribution of \mathbf{x} can be obtained by the two-layer hierarchical prior model, $p(\mathbf{x}|\lambda, b, a) = \prod_i p(x_i|\lambda, b, a)$, where

$$\begin{aligned}
 p(x_i|\lambda, b, a) &= \int p(x_i|\alpha_i) p(\alpha_i|\lambda, b, a) d\alpha_i \\
 &= \left(\frac{p}{\pi}\right)^p \frac{(a/b)^{\lambda/2}}{(a)^{\lambda/2-p/2} K_{\lambda}(\sqrt{ab})} \left[b + 2p|x_i|^2 \right]^{\lambda/2-p/2} K_{\lambda-p}\left(\sqrt{a(b+2p|x_i|^2)}\right).
 \end{aligned} \tag{9}$$

or

$$\begin{aligned}
 p(x_i|\lambda, a, b) &= \int p(x_i|\gamma_i) p(\gamma_i|\lambda, a, b) d\gamma_i \\
 &= \left(\frac{p}{\pi}\right)^p \frac{(b/a)^{-\lambda/2}}{(a)^{\lambda-p/2} K_{-\lambda}(\sqrt{ab})} \left[b + 2p|x_i|^2 \right]^{\lambda/2-p/2} K_{-(\lambda-p)}\left(\sqrt{a(b+2p|x_i|^2)}\right).
 \end{aligned} \tag{10}$$

The abovementioned setting of independent hyper-parameters for each parameter to be estimated is the most significant feature of the SBL model, and it is also the fundamental reason for the sparsity of the model [27, 28]. When $x_i \in \mathbb{R}$ takes the value $p = 1/2$, the

equation (9) corresponds to the generalized hyperbolic (GH) distribution of the real-valued model, where equation (10) is equivalent to (9). In this paper, $p = 1$ corresponds to the GH distribution of the complex-valued model.

As mentioned above, the GIG distribution is selected as the hyperprior $p(\alpha_i)$ and $p(\gamma_i)$, respectively, as it includes a fairly broad class of distributions commonly used as hyperpriors, and the resulting marginal distribution, the GH distribution, again covers a large number of distributions as special cases. Due to this generalization, the connections between different modeling strategies can be analyzed, and several special cases that correspond to standard priors commonly used in sparse modeling are summarized in Figure 2.

3.2. Complex Sparse Bayesian Learning. As mentioned above, the model of the signal matrix \mathbf{X} has been defined by different two-layer hierarchical priors, and the SBL model characterization will be completed in the following by modeling the observation matrix Y in equation (3). Firstly, according to the different values in the model, the noise \mathbf{E} is modeled as Gaussian (real-valued model) or complex Gaussian (complex-valued model) with independent and identical distribution as follows:

$$p(\mathbf{E}|\tau) = \prod_{m=1}^M \prod_{n=1}^N p(e_n(m)|\tau) = \left(\frac{p}{\pi\tau^{-1}}\right)^{pNM} \exp\{-p\tau\|\mathbf{E}\|_F^2\}, \quad (11)$$

where $\tau = 1/\sigma^2$ is the noise precision with variance σ^2 . The parameter $p = 1$ when $e_n(m)$ is complex, and $p = 1/2$ when $e_n(m)$ is real. Therefore, the following expression is obtained:

$$p(Y|X, \beta, \tau) = \left(\frac{p}{\pi\tau^{-1}}\right)^{pNM} \exp\{-p\tau\|\mathbf{Y} - \Phi\mathbf{X}\|_F^2\}. \quad (12)$$

The noise precision τ is unknown, and a Gamma hyperprior is assumed for τ with $p(\tau|c, d) = \Gamma(\tau|c, d)$, where τ obeys the Gamma distribution of parameters c and d . The offgrid parameter β obeys a uniform distribution with $\beta \sim U([-r/2, r/2]^L)$. Taking the two-layer hierarchical prior $CN(x_i|0, \alpha_i)$, $\Gamma(\gamma_i|\lambda, a/2)$ as an example, the relationship between the parameters of the SBL model is shown in Figure 3.

For equation (3) of the off-grid DOA estimation model, when the signal variance α is used to model the two-layer hierarchical prior, the joint probability distribution of the complete SBL model is represented as

$$p(Y, X, \alpha, \beta, \tau) = p(Y|X, \beta, \tau)p(X|\alpha)p(\alpha)p(\beta)p(\tau). \quad (13)$$

According to Bayesian theory, the posterior distribution of the parameters is estimated as

$$p(X, \alpha, \beta, \tau|Y) = \frac{p(Y|X, \beta, \tau)p(X|\alpha)p(\alpha)p(\beta)p(\tau)}{p(Y)}. \quad (14)$$

However, the exact posterior $p(X, \alpha, \beta, \tau|Y)$ cannot be given in closed form due to the fact that $p(Y)$ is not calculated analytically [25]. Therefore, SBL needs an effective approximation $p(X, \alpha, \beta, \tau|Y)$. The sparsity-inducing property of the resulting estimator \mathbf{X} does not only depend on the two-layer hierarchical prior, but also the approximation method used.

In SBL, two widespread approximate approaches referred to as type I and type II estimation have been used. Type II is considered in this paper. In type II estimation [4–9], the impact of hyperparameter α_i or γ_i is concerned on the model, and $p(X, \alpha, \beta, \tau|Y)$ can be decomposed as

$$p(X, \alpha, \beta, \tau|Y) = p(X|Y, \alpha, \beta, \tau)p(\alpha, \beta, \tau|Y), \quad (15)$$

where $p(\alpha, \beta, \tau|Y) = p(Y|\alpha, \beta, \tau)p(\alpha)p(\beta)p(\tau)/p(Y)$. Therefore, type II estimation needs to effectively approximate the posterior distribution $p(\alpha, \beta, \tau|Y)$, by using MAP to estimate hyperparameter α as follows:

$$\begin{aligned} \hat{\alpha}^{\text{MAP}} &= \arg \max_{\alpha} p(Y|\alpha, \beta, \tau) p(\alpha) \\ &= \arg \max_{\alpha} p(\alpha) \int p(Y|X, \beta, \tau)p(X|\alpha)dX \\ &= \arg \max_{\alpha} \left[\ln(p(\alpha)) + \ln \int p(Y|X, \beta, \tau)p(X|\alpha)dX \right], \end{aligned} \quad (16)$$

and type II estimation is equivalent to maximizing the cost function of α , then

$$\mathcal{L}(\alpha) = -p \times \text{tr}\{Y^H \Sigma_Y^{-1} Y\} - pM \ln|\Sigma_Y| + \ln(p(\alpha)), \quad (17)$$

where Σ_Y is the covariance matrix of the likelihood function $p(Y|\alpha, \beta, \tau)$. In the current widely studied real-valued model, it is usually assumed that a uniform hyperprior $p(\alpha) \propto 1$ for α , which leads to maximizing $\hat{\alpha}^{\text{MAP}}$ equivalently to maximizing the likelihood function $p(Y|\alpha, \beta, \tau)$. Therefore, type II estimation is also called as type II maximum likelihood (ML) estimation.

In [26], the results show that type II estimation has superior reconstruction performance than type I estimation. In subsequent paragraphs, type II estimation combined with the above two-layer hierarchical prior extension from the real-valued model is changed to the complex-valued model for the SBL derivation for offgrid DOA estimation.

In type II estimation, equation (15) is used to decompose the exact posterior $p(X, \alpha, \beta, \tau|Y)$, and the following expression can be acquired:

$$\begin{aligned} p(X|Y, \alpha, \beta, \tau) &= \frac{p(Y|X, \beta, \tau)p(X|\alpha)}{p(Y|\alpha, \beta, \tau)} = \frac{p(Y|X, \beta, \tau)p(X|\alpha)}{\int p(Y|X, \beta, \tau)p(X|\alpha)dX} \\ &= \left(\frac{p}{\pi}\right)^{pLM} |\Sigma_X|^{-pM} \exp\{-p \times \text{tr}((X - \mu_X)^H \Sigma_X^{-1} (X - \mu_X))\}, \end{aligned} \quad (18)$$

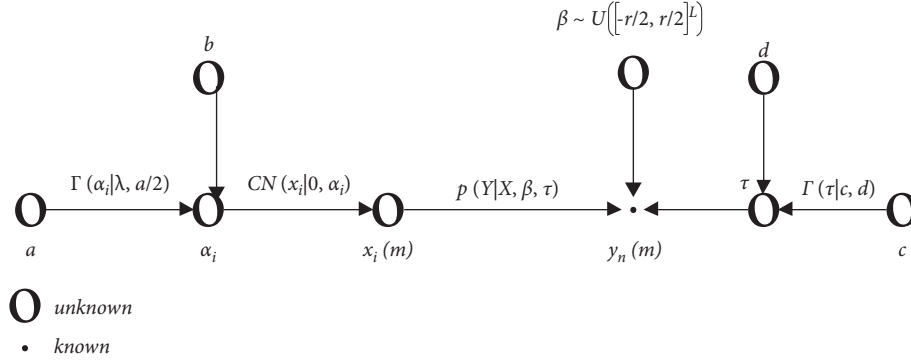


FIGURE 3: SBL off-grid DOA estimation parameter model.

where Σ_X and μ_X are the covariance matrix and the mean of $p(X|Y, \alpha, \beta, \tau)$, respectively,

$$\begin{cases} \Sigma_X = (\tau \Phi^H \Phi + \Lambda_\alpha^{-1})^{-1} \\ \mu_X = \tau \Sigma_X \Phi^H Y \end{cases} \quad (19)$$

The signal variance α and noise precision τ are estimated by using MAP as follows:

$$\begin{aligned} p(\alpha, \beta, \tau|Y) &\propto p(Y|\alpha, \beta, \tau) p(\alpha) p(\beta) p(\tau) \\ &= \int p(Y|X, \beta, \tau) p(X|\alpha) p(\alpha) p(\beta) p(\tau) dX \\ &= \left(\frac{p}{\pi}\right)^{pNM} |\Sigma_Y|^{-pM} \exp\{-p \times \text{tr}(Y^H \Sigma_Y^{-1} Y)\} p(\alpha) p(\beta) p(\tau), \end{aligned} \quad (20)$$

where Σ_Y is the covariance matrix of $p(\alpha, \beta, \tau|Y)$, and $\Sigma_Y = \tau^{-1} \mathbf{I}_N + \Phi \Lambda_\alpha^{-1} \Phi^H$. If α and τ are known, type II estimated value of X is obtained as

$$\hat{X} = \mu_X = \Lambda_\alpha \Phi^H \Sigma_Y^{-1} Y, \quad (21)$$

where $\Lambda_\alpha = \text{diag}(\alpha)$ is a diagonal matrix composed of signal variance, and its diagonal elements control the row-sparsity of \hat{X} . As α_i tends to a particularly small value (theoretically

zero), the corresponding i -th row of \hat{X} becomes $\mathbf{0}^T$. DOA parameters $\hat{\theta}_k$ (the closest grid to the true DOA $\tilde{\theta}_k$, $k = 1, 2, \dots, K$) can be estimated from the position of nonzero rows in \hat{X} .

When signal precision γ is used to model two-layer hierarchical prior, similar expression results are acquired as

$$\begin{cases} \Sigma_X = (\tau \Phi^H \Phi + \Lambda_\gamma)^{-1}, \\ \Sigma_Y = \tau^{-1} \mathbf{I}_N + \Phi \Lambda_\gamma^{-1} \Phi^H, \\ \hat{X} = \mu_X = \tau \Sigma_X \Phi^H Y \Lambda_\gamma^{-1} \Phi^H \Sigma_Y^{-1} Y. \end{cases} \quad (22)$$

where $\Lambda_\gamma = \text{diag}(\gamma)$ is a diagonal matrix composed of signal precision, as γ_i tends to a particularly large value (theoretically infinity), the corresponding i -th row of \hat{X} becomes $\mathbf{0}^T$.

In the above, the complete SBL model is analyzed with signal modeling variance α or precision γ separately. Next, the iterative process of α will be mainly discussed here, the iterative process of γ is similar and will not be repeated.

3.3. Two-Layer Hierarchical Prior $p(x_i|\alpha_i)$, $IG(\alpha_i|\lambda, b/2)$ Type II Estimation. Maximizing the cost function of α is equivalent to maximizing the following function:

$$\mathcal{L} = -pM \ln|\Sigma_Y| - p \times \text{tr}(Y^H \Sigma_Y^{-1} Y) + \sum_{i=1}^L \left((\lambda - 1) \ln(\alpha_i) - \frac{b}{2\alpha_i} \right) + (c - 1) \ln(\tau) - d\tau. \quad (23)$$

Then the partial derivative of \mathcal{L} with respect to α_i is

$$\frac{\partial \mathcal{L}}{\partial \alpha_i} = pM (\Sigma_X)_{i,i} \alpha_i^{-2} - \frac{pM}{\alpha_i} + p \|(\mu_X)_i\|_2^2 \alpha_i^{-2} + \frac{\lambda - 1}{\alpha_i} + \frac{b}{2} \alpha_i^{-2}. \quad (24)$$

Let the partial derivative be zero, it will result in the hyperparameter with the following expression:

$$\alpha_i = \frac{pM(\Sigma_x)_{ii} + p\|(\mu_x)_i\|_2^2 + b/2}{pM - \lambda + 1}. \quad (25)$$

In this paper, the Gamma hyperprior for the noise precision τ is considered, the update of τ can be similarly obtained as follows:

$$\tau = \frac{pM \times N + c - 1}{pM \times \text{tr}(\Phi^H \Phi \Sigma_x) + p\|Y - \Phi \mu_x\|_F^2 + d}. \quad (26)$$

3.4. Estimate of the Offgrid Parameter β . As mentioned above, the updates of signal variance and noise precision by SBL mode are derived. In the following, the update of the important parameter β is explained in the offgrid DOA estimation model. Considering that the only information $\beta \sim U([-r/2, r/2]^L)$ is bounded, the update of β cannot be obtained by MAP, so the maximum expectation algorithm is used to estimate the offgrid parameter β .

The quantity X is treated as a hidden variable and maximize $E_{p(X|Y, \alpha, \beta, \tau)}[\ln(p(Y, X, \alpha, \beta, \tau))]$, where $E_{p(X|Y, \alpha, \beta, \tau)}[\cdot]$ denotes an expectation with respect to the posterior of X as given in equation (18) using the current estimates of the hyperparameters, and is abbreviated as $E[\cdot]$. For the offgrid parameter β , the irrelevant items are ignored, and the value of β is computed by

$$\beta^{\text{new}} = \arg \max_{\beta} E[\ln(p(Y|X, \beta, \tau)p(\beta))]. \quad (27)$$

According to equations (12) and (14), maximizing $E[\ln(p(Y|X, \beta, \tau)p(\beta))]$ is equivalent to minimizing $E[\|Y - \Phi X\|_F^2] = \beta^T P \beta + 2v^T \beta + C$, where C is a constant independent of β , $P = \text{Re}\{(B^H B)^* \odot (M \times \Sigma_x + \mu_x \mu_x^H)\}$ is a positive semidefinite matrix, $v = \text{Re}\{M \times \text{diag}(B^H A \Sigma X)\} - \text{Re}\{\sum_{m=1}^M \text{diag}(\mu_x^*(m))B^H (y(m) - A\mu_x(m))\}$. The optimization problem of β is given by

$$\beta^{\text{new}} = \arg \min_{\beta \in [-r/2, r/2]^L} \{\beta^T P \beta + 2v^T \beta\}. \quad (28)$$

which is a linear least-squares problem with boundary constraints. Let its partial derivative with respect to β be zero as follows:

$$\frac{\partial}{\partial \beta} \beta^T P \beta + 2v^T \beta = 2(P\beta + v) = 0. \quad (29)$$

If P is invertible, equation (29) has a unique solution, that is

$$\tilde{\beta} = -P^{-1}v. \quad (30)$$

Considering the boundedness, when $\tilde{\beta} \in [-r/2, r/2]^L$, then $\beta^{\text{new}} = \tilde{\beta}$. While one of the above two conditions is not satisfied, β is updated element by element. When updating β_l , other elements are fixed and the l -th equation with the partial derivative is zero, and the expression is given by

$$\hat{\beta}_l = \frac{v_p + (P_l)_{-l}^T \beta_{-l}}{P_l}, \quad l = 1, 2, \dots, L, \quad (31)$$

where $(\cdot)_{-l}$ is the vector to remove the l -th element and P_l is the l -th column in P_l . The calculation formula of β is described as

$$\beta_l^{\text{new}} = \begin{cases} \tilde{\beta}_l & \text{if } P^{-1} \text{ exists and } \tilde{\beta} \in [-r/2, r/2]^L \quad 1) \\ \hat{\beta}_l & \text{if 1) doesn't hold and } \hat{\beta}_l \in [-r/2, r/2] \quad 2) \\ -r/2 & \text{if 1) and 2) don't hold and } \hat{\beta}_l < -r/2 \quad 3) \\ r/2 & \text{if 1) and 2) don't hold and } \hat{\beta}_l > r/2 \quad 4) \end{cases}. \quad (32)$$

Finally, the K DOAs will be estimated in the following equation:

$$\hat{\theta}_k = \tilde{\theta}_k + \beta_k^{\wedge}, \quad k = 1, 2, \dots, K, \quad (33)$$

where $\tilde{\theta}_k$ is the closest grid to the true DOA θ_k , and β_k^{\wedge} is the offgrid parameter corresponding to $\tilde{\theta}_k$.

3.5. Proposed Algorithm Flow. Through the above derivation, the algorithm flow is concluded in Table 1 for complex sparse Bayesian learning for offgrid DOA estimation, where the two-layer hierarchical prior $p(x_i|\alpha_i) = \text{CN}(x_i|0, \alpha_i)$, $IG(\alpha_i|\lambda, b/2)$ as an example.

In this paper, for the two-layer hierarchical prior $p(x_i|\gamma_i) = \text{CN}(x_i|0, \gamma_i^{-1})$, $\Gamma(\gamma_i|\lambda, b/2)$ and $p(x_i|\alpha_i) = \text{CN}(x_i|0, \alpha_i)$, $IG(\alpha_i|\lambda, b/2)$, there are $\lambda = -1, b \rightarrow 0^+$. In the models of the two-layer hierarchical prior $p(x_i|\alpha_i) = \text{CN}(x_i|0, \alpha_i)$, $\Gamma(\alpha_i|\lambda, a/2)$ and $p(x_i|\gamma_i) = \text{CN}(x_i|0, \gamma_i^{-1})$, $IG(\gamma_i|\lambda, a/2)$, with the complex-valued DOA estimation model, in order to obtain the L1-norm of x_i , there are $\lambda = 3/2, a = 1$. The $p(\tau) = \Gamma(\tau|c, d)$ for the noise precision, there are $c = 1, d \rightarrow 0^+$.

4. Simulation and Performance Analysis

This section validates the validity of the different two-level hierarchical priors for the application of DOA estimation. Suppose that ULA with $N = 21$ array elements for receiving far-field narrowband sources, the sources carrier frequency is $f_0 = 10$ GHz with SNR = 0 dB and interelement spacing $\Delta d = 3 \times 10^8 / (2f_0) = 1.5$ cm in equation (1).

Parameter initialization is as follows: there are $\alpha_l = 1, l = 1, 2, \dots, L$ (or $\gamma_l = 1$) and $\sigma^2 = 10^{-\text{SNR}/10} \|Y^o\|_F^2 / NM$, where Y^o is the noise-free received signal matrix. The offgrid parameter is $\beta = 0$. A uniform sampling grid is $\tilde{\theta} = \{\tilde{\theta}_1, \tilde{\theta}_2, \dots, \tilde{\theta}_L\}$ with a grid interval $r = 1^\circ$ in the DOA range $[-90^\circ, 90^\circ]$. The number of snapshots is $M = 200$. The proposed algorithm terminates if $\varepsilon \leq 10^{-4}$ or the maximum number of iterations reaches $\text{max_iter} = 2000$.

The DOA estimation results and details of the estimation results of a source located at DOA 0.5° are shown in Figures 4(a) and 4(b), respectively.

TABLE 1: Complex sparse Bayesian learning algorithm for offgrid DOA estimation.

Input quantities: the received signals Y , and the sampling grid in the DOA range $\bar{\theta}$

Output quantities: offgrid θ

Initialization: giving the initial values of parameters $\lambda, b, c, d, \alpha, \tau, \beta \leftarrow \mathbf{0}$, and setting the initial values of the convergence threshold ε and the maximum number of iterations max_iter , computing the array steering matrices $A(\bar{\theta})$ and $B(\bar{\theta})$.

(1) Matrix construction: Computing $\Phi(\beta) = A + B \text{diag}(\beta)$ from $A(\bar{\theta})$, $B(\bar{\theta})$, and β .

(2) Parameters update:

- 1) Updating matrices Σ_X and μ_X based on equation (19) from the current values of α, τ, Φ .
- 2) Acquiring the estimated values $\hat{\theta}_k$ (the closest grid to the true DOA $\bar{\theta}_k, k = 1, 2, \dots, K$) from the position of nonzero rows in $\hat{X} = \mu_X$.
- 3) Updating quantities α, τ from the current values of Σ_X and μ_X with equations (25) and (26).
- 4) Updating the offgrid parameter β from the current values $A(\bar{\theta}), B(\bar{\theta}), \Sigma_X$, and μ_X according to equation (32).

(3) Iteration termination: Calculating the residual value $\|Y - \Phi(\beta)\mu_X\|_F^2$. If the residual value is less than ε or the number of iterations is more than max_iter , terminating the iteration process, otherwise, jumping to (1) to start the iteration, and updating the matrices $A(\bar{\theta}), B(\bar{\theta})$, and $\Phi(\beta)$ based on equation (33).

(4) DOA detection: K DOAs will be estimated in equation (33).

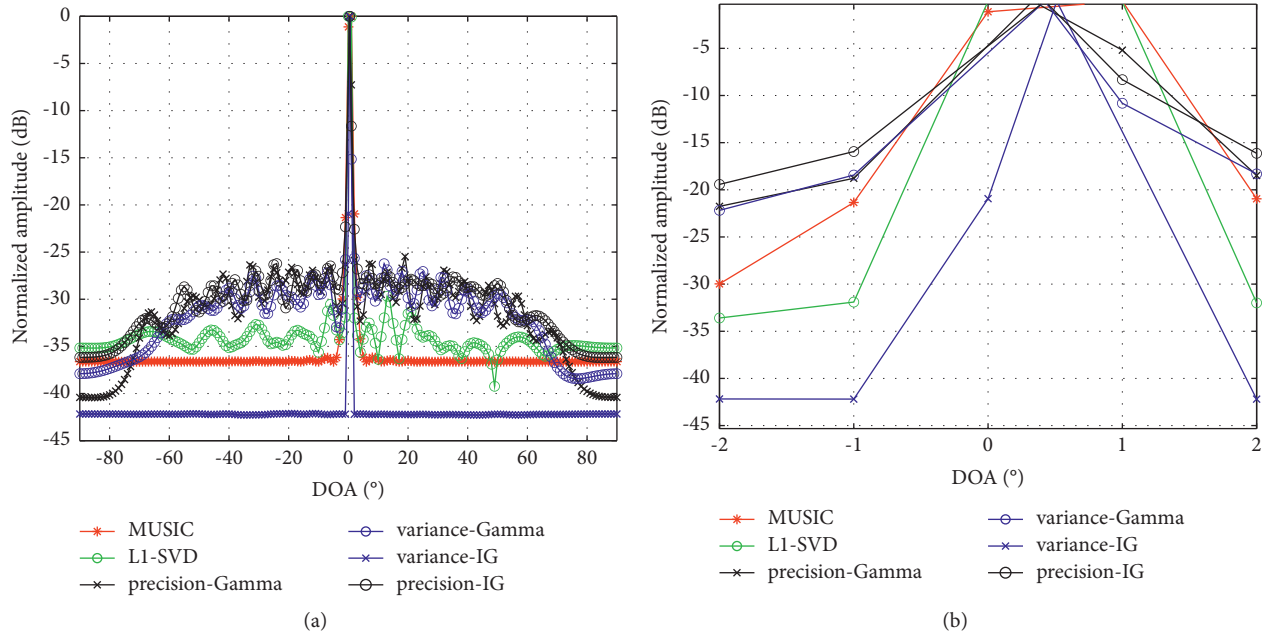


FIGURE 4: Single source of DOA estimation results: (a) single source DOA estimate and (b) single source DOA estimate details.

The DOA estimation results and details of the estimation results of the two sources located at DOAs -2.4° and 3.6° are shown in Figures 5(a) and 5(b), respectively.

Comparing Figures 4 and 5, it can be seen that under the consideration of the offgrid DOA model, different two-layer hierarchical priors used for sparse source x_i modeling can effectively estimate the DOA of the offgrid model, while the traditional MUSIC and L1-SVD algorithms have the DOA estimation results located on the spatial sampling grid nearest to the target point because the sampling grid mismatch is not considered.

In order to analyze the different two-layer hierarchical priors used for sparse source x_i modeling, the performance of the different two-layer hierarchical priors will be compared with the traditional MUSIC algorithm and L1-SVD algorithm, and the performances between the different two-level hierarchical priors are illustrated by simulation results.

The root mean square error (RMSE) of the DOA estimation result is given in the form of dB as follows:

$$\text{RMSE} = 10 \log_{10} \left(\sqrt{\frac{1}{KT} \sum_{k=1}^K \sum_{t=1}^T (\hat{\theta}_{kt} - \theta_k)^2} \right), \quad (34)$$

where K is the total number of sources ($K = 3$), T is the number of experiments ($T = 200$), and $\hat{\theta}_{kt}$ is the estimated DOA of the k -th source in the t -th experiment. Then consider 3 targets with DOAs $\theta_1 = -2.4^\circ$, $\theta_2 = 0.5^\circ$, and $\theta_3 = 3.6^\circ$ respectively.

Figure 6(a) shows the RMSE of DOAs estimation result, with the variation of SNR under the condition that the number of array elements is $N = 21$ and the number of snapshots $M = 200$. Figure 6(b) shows the RMSE of DOAs estimation result, with the variation of array elements under

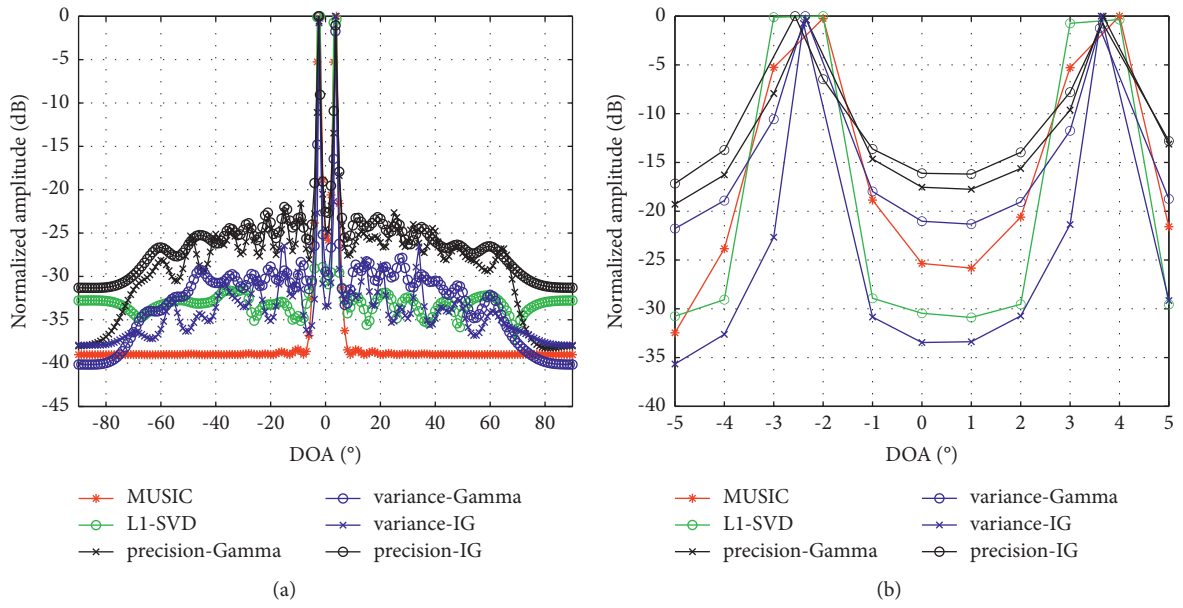


FIGURE 5: Two sources of DOA estimation results: (a) two sources DOA estimates and (b) two sources DOA estimates details.

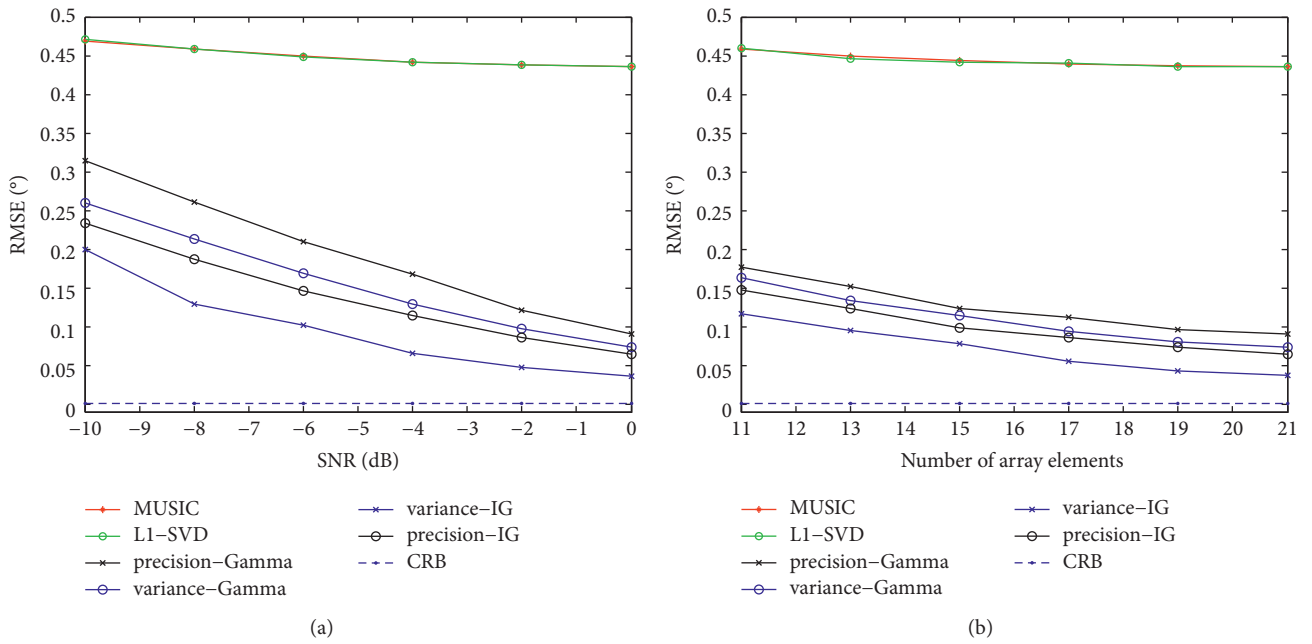


FIGURE 6: Continued.

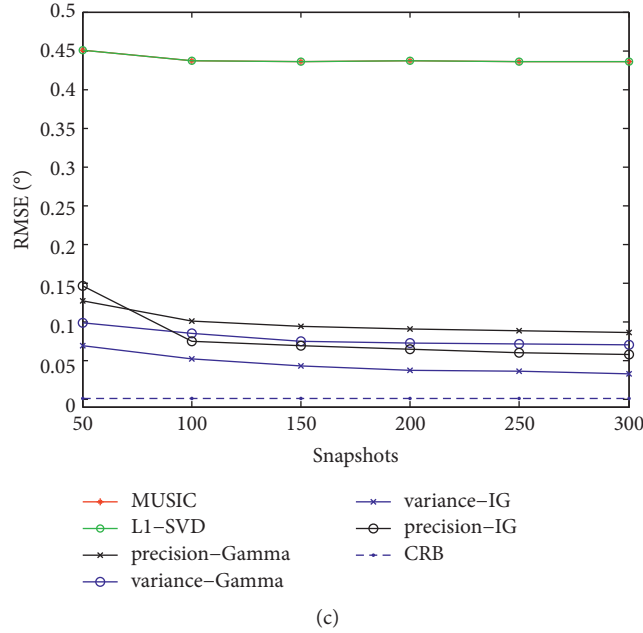


FIGURE 6: RMSE versus parameters: (a) RMSE with different SNR values; (b) RMSE with different array elements; and (c) RMSE with different snapshots.

the condition of SNR = 0 dB and the number of snapshots $M = 200$. Figure 6(c) shows the RMSE of DOAs estimation result, with the variation of snapshots under the condition of SNR = 0dB and the number of array elements $N = 21$. The Cramer-Rao bound (CRB) for the offgrid bias parameters of the offgrid model in (2) and (3) is acquired by $CRB = 1/(2M \text{tr}(\mathbf{B}^H \mathbf{B}) \text{SNR})$ [29–33]. Due to the fact that the value of $\text{tr}(\mathbf{B}^H \mathbf{B})$ tends to 10^4 , the values of CRB in Figure 6 is about $10^{-3} \sim 10^{-4}$, such as $CRB = [0.00107, 0.00085, 0.00067, 0.0005, 0.0004, 0.0003]$ in Figure 6(a), which causes that the CRB waveforms appear relatively flat by compared with RMSE value waveforms.

According to the comparison results of RMSE under different conditions in Figure 6, it can be seen that the RMSE of traditional MUSIC and L1-SVD both show a straight line when the SNR, the number of array elements, and the number of snapshots reach a certain level. Because they cannot estimate the offgrid parameter, and the maximum estimation precision of MUSIC and L1-SVD algorithm is -2° , 0° or 1° , 4° , respectively. In addition, the different two-layer hierarchical priors of the signal x_i are considered, and their performance comparisons are shown in Figure 6. The precision of DOA estimation has been significantly improved due to the fact that the offgrid DOA model takes into account the grid offset.

At the same time, the results of Figure 6 are further analyzed as following, for the Gamma hyperprior of $p(\alpha_i)$ and $p(\gamma_i)$ as an example. The marginal prior from two-layer hierarchical prior $p(x_i|\gamma_i)$ and $\Gamma(\gamma_i|\lambda, b/2)$ is student's t prior, and the marginal prior obtained by setting the values of λ for the two-layer hierarchical prior $p(x_i|\alpha_i)$, $\Gamma(\alpha_i|\lambda, a/2)$

is the Laplace prior. In [12], it is analyzed in detail that the Laplace prior has better performance than the student's t prior under the condition of SMV of the real-valued model. In this paper, this conclusion is extended to the case of MMV of the complex-valued model, the effects of different hyperpriors are considered, and it is found that the two-layer hierarchical prior $p(x_i|\alpha_i)$, $IG(\alpha_i|\lambda, b/2)$ has better performance than $p(x_i|\gamma_i)$, $IG(\gamma_i|\lambda, a/2)$, although the marginal prior obtained by the two-layer hierarchical prior $p(x_i|\gamma_i)$, $\Gamma(\gamma_i|\lambda, b/2)$ and $p(x_i|\alpha_i)$, $IG(\alpha_i|\lambda, b/2)$ are the same as student's t prior, and the marginal prior obtained by the two-layer hierarchical prior $p(x_i|\alpha_i)$, $\Gamma(\alpha_i|\lambda, a/2)$, and $p(x_i|\gamma_i)$, $IG(\gamma_i|\lambda, a/2)$ are the same as Laplace prior.

Through the above analysis, the performance of the two-layer hierarchical prior $p(x_i|\alpha_i)$ and $p(\alpha_i)$ is better than $p(x_i|\alpha_i)$ and $p(\gamma_i)$, and the IG prior model is better than Gamma prior model.

The proposed algorithm is further applied to the angle-distance positioning in the FDA (frequency diverse array) [34]. The complex multisnapshot SBL combined with FDA radar characteristics is used for target angle-distance two-dimensional localization. The log frequency offset is used at the transmitter to remove the distance periodicity of the beam pattern of the full-band frequency receiving mechanism of FDA radar for distance unambiguous localization. Through the transmitting of zero frequency offset and nonzero frequency signals, the angle-distance two-dimensional estimation is converted into two one-dimensional estimates of the azimuth and distance, respectively, thus reducing the amount of calculation. The simulation results are shown in Figure 7 by using the two-layer hierarchical

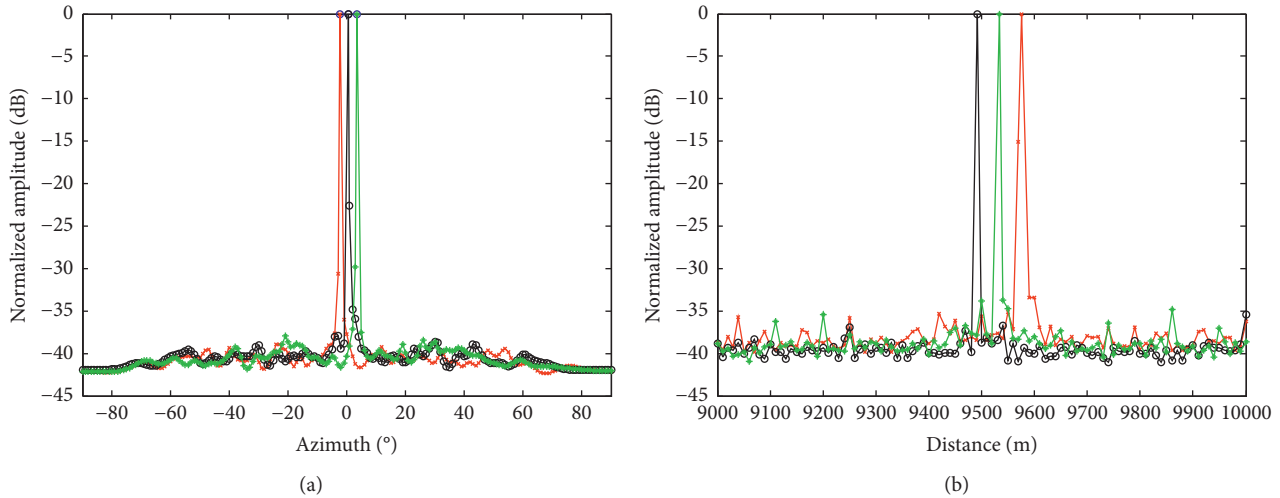


FIGURE 7: Three sources of angle-distance estimation results: (a) three sources azimuth estimates and (b) three sources distance estimates.

prior $p(x_i|\alpha_i)$, $\Gamma(\alpha_i|\lambda, a/2)$. The positions of the three targets are $(-2.4^\circ, 9576 \text{ m})$, $(0.5^\circ, 9492 \text{ m})$, and $(3.6^\circ, 9533 \text{ m})$, respectively. The simulation parameters are consistent with Figure 5. The log frequency offset $\Delta f_n = \Delta f \log(n)$, $n = 1, 2, \dots, N$ with $\Delta f = 2 \text{ MHz}$. The grid interval $r = 10 \text{ m}$ in the distance interval $[9000 \text{ m}, 10000 \text{ m}]$. The simulation results are $(-2.402^\circ, 9576.336 \text{ m})$, $(0.505^\circ, 9492.047 \text{ m})$, and $(3.583^\circ, 9533.193 \text{ m})$.

5. Conclusions

A general complex signal prior construction suitable for complex-valued signal modeling is proposed. By using the complex Gaussian scale mixtures hierarchy, it has been shown that this signal model includes complex versions of a number of signal models commonly used for sparse signal modeling. In addition, the expression of the SBL form of the real-valued and complex-valued models is unified by parameter values. Our motivation is to generalize the current research work of SBL by this unified form and contribute to its generalization to complex-valued signal models. Finally, the SBL complex-valued form is applied to the off-grid DOA estimation complex-valued model and compared with the performance between different sparsity-inducing priors. Theoretical analysis and simulation results show that the proposed algorithm can effectively use SBL to process complex-valued signal models and has lower algorithm complexity.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This research is funded by the National Natural Science Foundation of China (no. 61861011 and 51808554), the Guangxi Natural Science Foundation (no. 2019GXNSFBA245072), the Dean Project of Guangxi Wireless Broadband Communication and Signal Processing Key Laboratory (no. GXKL06160110), and the Science and Technology on Near-Surface Detection Laboratory Foundation (no. TCGZ2017A010).

References

- [1] E. J. Candes and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 21–30, 2008.
- [2] E. J. Candes and T. Tao, "Near-optimal signal recovery from random projections: universal encoding strategies?" *IEEE Transactions on Information Theory*, vol. 52, no. 12, pp. 5406–5425, 2006.
- [3] D. M. Malioutov, *A Sparse Signal Reconstruction Perspective for Source Localization with Sensor Arrays*, Massachusetts Institute of Technology, Cambridge, MA, USA, 2003.
- [4] M. E. Tipping, "Sparse Bayesian learning and the relevance vector machine," *Journal of Machine Learning Research*, vol. 1, no. 3, pp. 211–244, 2001.
- [5] 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.
- [6] D. P. Wipf and B. D. Rao, "An empirical Bayesian strategy for solving the simultaneous Sparse approximation problem," *IEEE Transactions on Signal Processing*, vol. 55, no. 7, pp. 3704–3716, 2007.
- [7] D. Wipf, J. Palmer, and B. Rao, "Perspectives on sparse Bayesian learning," *Advances in Neural Information Processing Systems*, vol. 16, pp. 249–256, 2003.
- [8] S. Ji, Y. Xue, and L. Carin, "Bayesian compressive sensing," *IEEE Transactions on Signal Processing*, vol. 56, no. 6, pp. 2346–2356, 2008.
- [9] 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, 2009.

- [10] Z. Yang, L. Xie, and C. Zhang, "Bayesian Compressed Sensing with New Sparsity-Inducing Prior," pp. 1–11, *Information Theory*, 2012.
- [11] Z.-M. Liu, Z.-T. Huang, and Y.-Y. Zhou, "An efficient maximum likelihood method for Direction-of-Arrival estimation via sparse Bayesian learning," *IEEE Transactions on Wireless Communications*, vol. 11, no. 10, pp. 1–11, 2012.
- [12] M. A. Hannan and P. Rocca, "Directions-of-Arrival estimation in linear sub-arrayed array through compressive sensing," in *Proceedings of the 2019 IEEE International Symposium on Antennas and Propagation and USNC-URSI Radio Science Meeting*, July 2019.
- [13] 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.
- [14] Z. Yang, L. Xie, and C. Zhang, "Off-grid Direction of arrival estimation using sparse Bayesian inference," *IEEE Transactions on Signal Processing*, vol. 61, no. 1, pp. 38–43, 2013.
- [15] X. Wu, W. P. Zhu, and J. Yan, "Direction of arrival estimation for off-grid signals based on sparse Bayesian learning," *IEEE Sensors Journal*, vol. 16, no. 7, pp. 2004–2016, 2015.
- [16] Z. Yang and L. Xie, "On gridless sparse methods for multi-snapshot direction of arrival estimation," *Circuits, Systems, and Signal Processing*, vol. 36, no. 8, pp. 3370–3384, 2017.
- [17] X. Wu, W.-P. Zhu, J. Yan, and Z. Zhang, "A spatial filtering based gridless DOA estimation method for coherent sources," *IEEE Access*, vol. 6, Article ID 56402, 2018.
- [18] M. Wagner, P. Gerstoft, and Y. Park, "Gridless DOA estimation via alternating projections," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 4215–4219, Brighton, UK, May 2019.
- [19] Z. Yang, J. Li, P. Stoica, and L. Xie, "Sparse methods for direction-of-arrival estimation," *Academic Press Library in Signal Processing*, vol. 7, pp. 509–581, 2018.
- [20] G. Oliveri, M. Carlin, and A. Massa, "Complex-weight sparse linear array synthesis by bayesian compressive sampling," *IEEE Transactions on Antennas and Propagation*, vol. 60, no. 5, pp. 2309–2326, 2012.
- [21] M. Carlin, P. Rocca, G. Oliveri, F. Viani, and A. Massa, "Directions-of-Arrival estimation through bayesian compressive sensing strategies," *IEEE Transactions on Antennas and Propagation*, vol. 61, no. 7, pp. 3828–3838, 2013.
- [22] Q. Wu, Y. D. Zhang, M. G. Amin, and B. Himed, "Complex multitask Bayesian compressive sensing," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 3375–3379, Florence, Italy, May 2014.
- [23] X. Wu, W.-P. Zhu, J. Yan, and Z. Zhang, "Two sparse-based methods for off-grid direction-of-arrival estimation," *Signal Processing*, vol. 142, pp. 87–95, 2018.
- [24] A. Das and T. J. Sejnowski, "Narrowband and wideband off-grid direction-of-arrival estimation via sparse Bayesian learning," *IEEE Journal of Oceanic Engineering*, vol. 43, no. 1, pp. 108–118, 2018.
- [25] B. Jorgensen, *Statistical Properties of the Generalized Inverse Gaussian Distribution*, Springer Science & Business Media, Berlin, Germany, 2012.
- [26] N. Abramson and D. J. Braverman, *Pattern Recognition and Machine Learning*, Springer, Berlin, Germany, 2006.
- [27] D. J. C. MacKay, "Bayesian interpolation," *Neural Computation*, vol. 4, no. 3, pp. 415–447, 1992.
- [28] R. Giri and B. Rao, "Type I and type II Bayesian methods for sparse signal recovery using scale mixtures," *IEEE Transactions on Signal Processing*, vol. 64, no. 13, pp. 3418–3428, 2016.
- [29] R. Jagannath and K. V. S. Hari, "Block sparse estimator for grid matching in Single snapshot DOA estimation," *IEEE Signal Processing Letters*, vol. 20, no. 11, pp. 1038–1041, 2013.
- [30] R. Prasad and C. R. Murthy, "Cramer-rao-Type bounds for sparse bayesian learning," *IEEE Transactions on Signal Processing*, vol. 61, no. 3, 2013.
- [31] A. Das, "Theoretical and experimental comparison of off-grid sparse bayesian direction-of-arrival estimation algorithms," *IEEE Access*, vol. 5, Article ID 18075, 2017.
- [32] Y. Qin, Y. Liu, and Z. Yu, "Underdetermined DOA estimation using coprime array via multiple measurement sparse Bayesian learning," *Signal, Image and Video Processing*, vol. 13, no. 7, pp. 1311–1318, 2019.
- [33] X. Wu, W.-P. Zhu, and J. Yan, "Direction of arrival estimation for off-grid signals based on sparse bayesian learning," *IEEE Sensors Journal*, vol. 16, no. 7, pp. 2004–2016, 2016.
- [34] Q. H. Liu, K. Ding, B. S. Wu, and Q. Xie, "Frequency diverse array target localization based on IPSO-BP," *International Journal of Antennas and Propagation*, vol. 2020, Article ID 2501731, 8 pages, 2020.