

Research Article

DOA Estimation for a Mixture of Uncorrelated and Coherent Sources Based on Hierarchical Sparse Bayesian Inference with a Gauss-Exp-Chi² Prior

Pinjiao Zhao ¹, Weijian Si,² Guobing Hu,¹ and Liwei Wang³

¹Department of Electronic and Information Engineering, Jinling Institute of Technology, Nanjing 211169, China

²Department of Information and Communication Engineering, Harbin Engineering University, Harbin 150001, China

³No. 8511 Research of Institute of CASIC, Nanjing 210007, China

Correspondence should be addressed to Pinjiao Zhao; zhaopinjiao@hrbeu.edu.cn

Received 29 November 2017; Accepted 31 May 2018; Published 10 July 2018

Academic Editor: Sotirios K. Goudos

Copyright © 2018 Pinjiao Zhao 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 algorithms based on sparse Bayesian inference (SBI) can effectively estimate coherent sources without recurring to extra decorrelation techniques, and their estimation performance is highly dependent on the selection of sparse prior. Specifically, the specified sparse prior is expected to concentrate its mass on the zero and distribute with heavy tails; otherwise, these algorithms may suffer from performance degradation. In this paper, we introduce a new sparse-encouraging prior, referred to as “Gauss-Exp-Chi²” prior, and develop an efficient DOA estimation algorithm for a mixture of uncorrelated and coherent sources under a hierarchical SBI framework. The Gauss-Exp-Chi² prior distribution exhibits a sharp peak at the origin and heavy tails, and this property makes it an appropriate prior to encourage sparse solutions. A three-layer hierarchical sparse Bayesian model is established. Then, by exploiting variational Bayesian approximation, the model parameters are estimated by alternately updating until Kullback-Leibler (KL) divergence between the true posterior and the variational approximation becomes zero. By constructing the source power spectra with the estimated model parameters, the number and locations of the highest peaks are extracted to obtain source number and DOA estimates. In addition, some implementation details for algorithm optimization are discussed and the Cramér-Rao bound (CRB) of DOA estimation is derived. Simulation results demonstrate the effectiveness and favorable performance of the proposed algorithm as compared with the state-of-the-art sparse Bayesian algorithms.

1. Introduction

Direction of arrival (DOA) estimation has been a crucial issue in various application areas involving radar, wireless communication, and navigation [1–3]. Multiple signal classification (MUSIC) [4] and estimation of signal parameter via rotational invariance technique (ESPRIT) [5], known as the two most classical subspace-based DOA estimation algorithms, have been proposed to resolve uncorrelated sources. However, in multipath propagation environments, sources from an identical target may undergo reflection from various surfaces, and thus, the received sources may be a mixture of uncorrelated and coherent sources. In such environments,

the aforementioned subspace-based algorithms would suffer from serious performance deterioration or even fail [6].

In order to solve the aforementioned problem, several preprocessing techniques are developed for decorrelation. In the related studies, these preprocessing techniques are mainly classified into two categories: spatial smoothing (SS) [7] and matrix reconstruction (MR) [8]. Theoretically, the SS technique is implemented by dividing the whole array into multiple subarrays to combat the rank deficiency, while the MR technique is performed by rearranging the rank-deficient matrix into Hankel matrix, Toeplitz matrix, or other matrices to restore the rank. By combining the SS or MR techniques with subspace-based algorithms, several

DOA estimation algorithms have been developed to handle the scenarios where uncorrelated and coherent sources coexist [9–13]. More specifically, the oblique projection spatial smoothing (OPSS) [9] and moduli property spatial smoothing (MPSS) [10] are the typical SS-based algorithms, and the oblique projection Toeplitz matrix reconstruction (OPT-MR) [11], symmetric non-Toeplitz matrix reconstruction (SNT-MR) [12], and real-valued Hankel matrix reconstruction (RVH-MR) [13] are the popular MR-based algorithms.

Unlike traditional subspace-based DOA estimation algorithms, the emerging sparse source reconstruction (SSR) algorithms [14–19], including matching pursuit (MP) algorithm [14], l_p -norm optimization algorithms [15, 16], and sparse Bayesian inference (SBI) algorithms [17–19], provide a new perspective for DOA estimation. Since SSR-based algorithms realize DOA estimation via sparse source reconstruction, instead of calculating the covariance matrix, they can resolve the coherent sources directly without extra pre-processing techniques. Among these SSR-based algorithms, both the MP and the l_p -norm optimization algorithms are restricted to solve an optimization problem and are relied on the point estimate, so that these algorithms ignore the uncertainty of the underlying source in the process of source reconstruction. By contrast, SBI-based algorithms specially consider the uncertainty of the underlying source and estimate the source via choosing an appropriate prior, which yield favorable reconstruction performance [20]. Moreover, they can provide good estimation performance in the case of low SNR or small number of snapshots [21]. In general, SBI-based algorithms first specify a sparsity-encouraging prior to the unknown source model and then the model parameters are estimated via Bayesian inference. Since the exact Bayesian inference is intractable, two mainstream approximation inference algorithms were presented to estimate the model parameters, in which one is evidence procedure [22] and the other is variational Bayesian approximation [23]. For the evidence procedure, some unknown hyperparameters with respect to the hierarchical prior model are estimated iteratively by maximizing the evidence. For the variational Bayesian approximation, the posterior distribution is approximated as the product of several tractable distributions, and the model parameters keep updating to minimize the Kullback-Leibler (KL) divergence between the true posterior and the variational approximation, which has attractive computational efficiency along with high estimation performance [24]. Note that these approximation inference algorithms operate under the premise that an appropriate sparse prior has been imposed on the source model for the purpose of encouraging sparse solutions. Many sparsity-encouraging prior models have been investigated in the SBI framework [25–27]. In [25], Bayesian compressed sensing (BCS) was proposed with a two-layer hierarchical Gaussian-inverse-gamma prior (or Student's- t prior), where the first layer is a Gaussian probability density function (pdf) and the second layer is a gamma pdf. Babacan et al. [26] proposed an equivalent Laplace prior that is generated by a Gaussian prior and an Exponential prior. In [27], a normal product (NP) prior was developed with two algorithms:

NP-0 (using one-layer source model) and NP-1 (using two-layer source model). However, these priors are concentrated near the origin with relatively light tails, which may cause overshrinkage of the incident sources [28].

In this paper, we develop a new sparse-encouraging prior (called Gauss-Exp-Chi² prior) whose pdf distribution exhibits a sharp peak at the origin and heavy tails. With the proposed prior, the DOA estimation for a mixture of uncorrelated and coherent sources is performed under the hierarchical SBI framework using a uniform linear array (ULA). A three-layer hierarchical Bayesian model is established based on the Gauss-Exp-Chi² prior. Subsequently, according to the variational Bayesian approximation, the model parameters (including the mean and variance of sparse sources and hyperparameters) keep alternately updating until the KL divergence between the true posterior and the variational approximation tends to be zero. By exploiting the estimated model parameters, the source power spectra is constructed, from which the number and locations of the highest peaks are extracted to obtain source number and DOA estimates. Simulation results show that the proposed algorithm has superior estimation performance. Now we briefly summarize the contributions of this work as follows:

- (i) To encourage sparse solutions, we develop a new sparse-encouraging prior, called Gauss-Exp-Chi² prior, whose pdf distribution has a sharp peak at the origin and heavy tails.
- (ii) By constructing the source powers of all the potential directions in the angular space, both source number and DOA estimates are obtained.
- (iii) Variational approximations are adopted for the estimation of the hierarchical sparse Bayesian model parameters.
- (iv) Several implementation details for algorithm optimization including Woodbury matrix identity for dimension-reduction, pruning a basis function and the third kind Bessel function approximation are discussed, and the CRB of DOA estimation is derived.

The remainder of this paper is organized as follows. The DOA estimation model for mixed sources is formulated in Section 2. Section 3 presents the Gauss-Exp-Chi² prior and DOA estimation algorithm for a mixture of uncorrelated and coherent sources within the hierarchical SBI framework. The algorithm optimization and CRB of DOA estimation are discussed in Section 4. Section 5 presents the simulation results of the proposed algorithm. Conclusions are drawn in Section 6.

Notations. Vectors and matrices are denoted by lowercase and uppercase boldface letters, respectively. $(\cdot)^T$, $(\cdot)^H$, $(\cdot)^{-1}$, and $\langle \cdot \rangle$ represent transpose, conjugate transpose, inverse, and the statistical expectation, respectively. \mathbf{I} is an identity matrix. $\|\cdot\|_2$ denotes the Euclidean norm, and diag

(\cdot) denotes a diagonal matrix. Additionally, $\int_a^b v dv$ represents the integral of v from a to b .

2. Problem Formulation

Consider a total of K far-field narrowband sources impinging on the uniform linear array (ULA) consisting of M omnidirectional sensors with the interspacing between adjacent sensors d being a half of the carrier wavelength λ , that is, $d = \lambda/2$. Then, the array output vector at the t th snapshot can be expressed as

$$\mathbf{x}(t) = \sum_{k=1}^K \mathbf{a}(\theta_k) s_k(t) + \mathbf{n}(t), \quad t = 1, 2, \dots, T, \quad (1)$$

where T is the number of snapshots, $\mathbf{a}(\theta_k) = [1, e^{-j\pi \sin \theta_k}, \dots, e^{-j\pi(M-1) \sin \theta_k}]^T$ is the steering vector corresponding to the direction θ_k of the k th impinging source $s_k(t)$, and $\mathbf{n}(t)$ denotes the noise vector. Without loss of generality, the impinging source $\{s_k(t)\}_{k=1}^K$ is a mixture of uncorrelated and coherent sources. Note that there exists power fading among the coherent source, and fading coefficients are used to describe the degree of power fading [5]. Specifically, consider K far-field narrowband sources $\{s_k(t)\}_{k=1}^K$ impinging on the ULA, in which the first K_u sources $s_1(t), s_2(t), \dots, s_{K_u}(t)$ are uncorrelated, and the last K_c sources $s_{K_u+1}(t), s_{K_u+2}(t), \dots, s_K(t)$ are coherent. Then, the uncorrelated source set is denoted as $\{s_k(t)\}_{k=1}^{K_u}$ and the coherent source set is denoted as $\{s_k(t)\}_{k=K_u+1}^K = [1, \rho_1, \rho_2, \dots, \rho_{K_c-1}]^T s_{K_u+1}(t)$ with $\rho_1, \rho_2, \dots, \rho_{K_c-1}$ are the fading coefficients.

Divide the entire angular space into J sampling grids $\widehat{\boldsymbol{\theta}} = [\widehat{\theta}_1, \widehat{\theta}_1, \dots, \widehat{\theta}_J]^T$, where J represents the grid number and generally satisfies $J \gg M > K$. Assume that $\widehat{\theta}_{j_k} \in [\widehat{\theta}_1, \widehat{\theta}_1, \dots, \widehat{\theta}_J]^T$ is the nearest grid to true direction $\theta_j \in [\theta_1, \theta_2, \dots, \theta_K]^T$; thus, we have $h_j(t) = s_{j_k}(t) \delta(j - j_k)$. Thus, $\mathbf{x}(t)$ can be rewritten in a sparse form

$$\mathbf{x}(t) = \widehat{\mathbf{A}} \mathbf{h}(t) + \mathbf{n}(t), \quad t = 1, 2, \dots, T, \quad (2)$$

where $\widehat{\mathbf{A}} = [\boldsymbol{\alpha}(\widehat{\theta}_1), \boldsymbol{\alpha}(\widehat{\theta}_2), \dots, \boldsymbol{\alpha}(\widehat{\theta}_J)]$ and $\mathbf{h}(t) = [h_1(t), h_2(t), \dots, h_J(t)]^T$. Due to the fact that $\mathbf{h}(t)$ has K non-zero elements in J elements, it is a sparse vector. In the case of multiple snapshots, the sparse sources at all the snapshots share the same support, and the array output matrix of T snapshots can be represented by

$$\mathbf{X} = \widehat{\mathbf{A}} \mathbf{H} + \mathbf{N}, \quad (3)$$

where $\mathbf{X} = [\mathbf{x}(1), \mathbf{x}(2), \dots, \mathbf{x}(T)]$, $\mathbf{H} = [\mathbf{h}(1), \mathbf{h}(2), \dots, \mathbf{h}(T)]$, and $\mathbf{N} = [\mathbf{n}(1), \mathbf{n}(2), \dots, \mathbf{n}(T)]$. The goal of this paper is to provide the DOA estimation under the coexistence of uncorrelated and coherent sources from a sparse Bayesian perspective.

3. DOA Estimation

In this section, a DOA estimation algorithm for a mixture of uncorrelated and coherent sources is proposed within the hierarchical SBI framework. A Gauss-Exp-Chi² prior is developed to encourage sparse solutions, and then the parameters of three-layer hierarchical Bayesian model are estimated via variational Bayesian approximations. By constructing source power spectra, the source number and DOA estimation are obtained.

3.1. Bayesian Model. In the Bayesian model, the pdf of a joint distribution $p(\mathbf{X}, \mathbf{H}, \boldsymbol{\alpha}, \eta)$ with respect to all the unknown and observed quantities is required to be known. In this paper, the joint distribution can be expressed as

$$p(\mathbf{X}, \mathbf{H}, \boldsymbol{\alpha}, \eta) = p(\mathbf{X}|\mathbf{H}, \sigma^2) p(\mathbf{H}|\boldsymbol{\alpha}) p(\boldsymbol{\alpha}|\tau, \eta) p(\eta|v), \quad (4)$$

where $\boldsymbol{\alpha}$ and η are referred to as the hyperparameters; $\tau \geq 0$ and $v \geq 0$ are, respectively, referred to as the rate parameter and shape parameter. It is assumed that the components $\mathbf{n}(t), t = 1, 2, \dots, T$ are the independently zero-mean stationary Gaussian noise with known variance σ^2 . Thus, the pdf of the noise matrix \mathbf{N} is given by

$$p(\mathbf{N}|\sigma^2) = \prod_{t=1}^T \mathcal{N}(\mathbf{n}(t)|\mathbf{0}, \sigma^2 \mathbf{I}). \quad (5)$$

Combining (3) and (5), the Gaussian likelihood model is obtained as follows:

$$p(\mathbf{X}|\mathbf{H}, \sigma^2) = \prod_{t=1}^T \mathcal{N}(\mathbf{x}(t)|\widehat{\mathbf{A}} \mathbf{h}(t), \sigma^2 \mathbf{I}). \quad (6)$$

From a Bayesian perspective, the pdf distribution of an assigned prior is appealing to exhibit a sharp peak at the origin and heavy tails, which favors strong shrinkage of noise sources and avoids overshrinkage of the interest sources. This property is generally considered as a desirable property for enforcing sparsity and variable selection [24]. Some typical sparse priors, such as Gaussian-inverse-gamma prior and Laplace prior, are widely used in the relevant research [25, 26], which, however, are concentrated near the origin with relatively light tails [28]. To alleviate this problem, we here develop a three-layer hierarchical prior, referred to as Gauss-Exp-Chi² prior, for \mathbf{H} . In the first layer of prior, we adopt a zero-mean Gaussian prior

$$\begin{aligned} p(\mathbf{H}|\boldsymbol{\alpha}) &= \prod_{t=1}^T \mathcal{N}(\mathbf{h}(t)|\mathbf{0}, \Lambda^{-1}) \\ &= \prod_{t=1}^T \prod_{j=1}^J \mathcal{N}(h_j(t)|0, \alpha_j), \end{aligned} \quad (7)$$

with being $\Lambda^{-1} = \text{diag}(\boldsymbol{\alpha}) = \text{diag}(\alpha_1, \alpha_2, \dots, \alpha_J)$. In the second layer of prior, an exponential hyperprior is imposed on $\boldsymbol{\alpha}$

$$p(\boldsymbol{\alpha}|\tau, \eta) = \prod_{j=1}^J \text{Exp}(\alpha_j|\tau\eta), \quad (8)$$

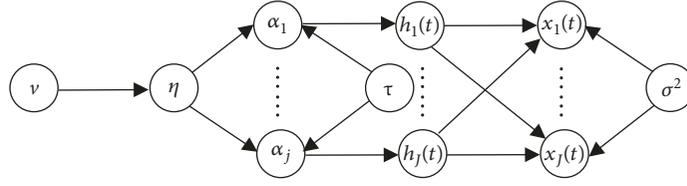


FIGURE 1: Directed graph of the sparse Bayesian model.

where $\text{Exp}(\cdot)$ denotes the exponential distribution. In the third layer of prior, a chi-square (Chi2) hyperprior is considered over η , that is,

$$p(\eta|v) = \chi^2(\eta|v) = \left[\Gamma\left(\frac{v}{2}\right) \right]^{-1} 2^{-v/2} \eta^{v/2-1} \exp\left(-\frac{\eta}{2}\right), \quad (9)$$

where $\Gamma(\cdot)$ denotes the gamma function.

The first two layers (7) and (8) of the proposed prior result in a generative Laplace (Gaussian-Exponential) distribution [26], and the last layer (i.e., a Chi2 distribution) is embedded to obtain the proposed three-layer Gauss-Exp-Chi² prior. Thus, the proposed prior has more free parameters to control the degree of sparseness as compared with the Laplace prior [29].

Based on the above analysis, the directed graph of the sparse Bayesian model is shown in Figure 1, where arrows are used to point to the generative model. Note that the first five blocks from the left (corresponding to the variables v , η , τ , $\alpha_1, \dots, \alpha_j$, and $h_1(t), \dots, h_j(t)$) depict the three-layer hierarchical prior mentioned above.

Since $p(\alpha_j | \tau v) \propto (\tau \alpha_j + 1/2)^{-(v/2+1)}$, $j = 1, 2, \dots, J$, it is obvious that most α_j for $j = 1, 2, \dots, J$ are favored to zero, which leads to the fact that most $h_j(t)$ are favored to zero. Thus, the proposed three-layer hierarchical Gauss-Exp-Chi² prior is a sparsity-encouraging prior, which can help to improve the performance of source reconstruction [30]. By combining the three layers of the proposed prior, the generated prior can be computed via

$$p(\mathbf{H}; \tau, v) = \prod_{t=1}^T \prod_{j=1}^J p(h_j(t); \tau, v), \quad (10)$$

where

$$p(h_j(t); \tau, v) = \int_0^{+\infty} \int_0^{+\infty} p(h_j(t) | \alpha_j) p(\alpha_j | \tau, \eta) p(\eta | v) d\alpha_j d\eta. \quad (11)$$

Note that $p(\alpha_j | \tau v) = \int_0^{+\infty} p(\alpha_j | \tau \eta) p(\eta | v) d\eta$; thus, $p(h_j(t); \tau, v)$ can be computed by marginalizing $p(h_j(t) | \alpha_j) p(\alpha_j | \tau v)$ over α_j (see Appendix for detail).

$$p(\mathbf{H}; \tau, v) = \prod_{t=1}^T \prod_{j=1}^J \sqrt{\frac{\tau}{\pi}} \frac{\Gamma(v/2 + 1) \Gamma(v/2 + 1/2)}{\Gamma(v/2)} \cdot U_{\text{chf}}\left(\frac{v}{2} + \frac{1}{2}, \frac{1}{2}, h_j^2(t)\tau\right), \quad (12)$$

where $U_{\text{chf}}(\cdot, \cdot, \cdot)$ is the confluent hypergeometric function defined by

$$U_{\text{chf}}(a, a - b + 1, c) = \frac{1}{\Gamma(a)} \int_0^{+\infty} t^{a-1} (1+t)^{-b} \exp(-ct) dt. \quad (13)$$

The reason for choosing the Gauss-Exp-Chi² prior is twofold: its pdf distribution has a sharp spike at the origin and heavy tails; it forms in a conjugate manner since the exponential and Chi2 distributions belonging to the exponential distribution family are chosen as the 2nd layer and the 3rd layer of this prior, which significantly simplifies the form of posterior distributions [24]. The tail to spike behavior for Gauss-Exp-Chi² prior, Laplace prior, Student's- t prior, and Gaussian prior is illustrated in Figure 2, wherein the parameter selection of these priors is according to the standard derivation [25, 26]. It is observed that the proposed Gauss-Exp-Chi² prior is a sparsity-encouraging prior which has sharper peak and heavier tails than the existing priors.

3.2. Variational Bayesian Approximations. It is well known that Bayesian inference operates on the basis of the posterior distribution $p(\mathbf{H}, \mathbf{a}, \eta | \mathbf{X}) = p(\mathbf{H}, \mathbf{a}, \eta, \mathbf{X}) / p(\mathbf{X})$. However, this posterior distribution is intractable for the reason that $p(\mathbf{X}) = \int \int p(\mathbf{H}, \mathbf{a}, \eta, \mathbf{X}) d\mathbf{H} d\mathbf{a} d\eta$ cannot be calculated analytically. In order to approximate this posterior distribution, we resort to mean-field variational Bayesian approximation [23], whose task is to seek several analytically approximate distributions that minimize the KL divergence between the posterior $p(\mathbf{H}, \mathbf{a}, \eta | \mathbf{X})$ and its approximation distribution $q(\mathbf{H}, \mathbf{a}, \eta)$. According to [23], $q(\mathbf{H}, \mathbf{a}, \eta)$ can be factorized as the product of three variational distributions $q(\mathbf{H})$, $q(\mathbf{a})$, and $q(\eta)$; thus, we have

$$p(\mathbf{H}, \mathbf{a}, \eta | \mathbf{X}) \approx q(\mathbf{H}, \mathbf{a}, \eta) = q(\mathbf{H})q(\mathbf{a})q(\eta). \quad (14)$$

Denote a set $\zeta = \{\mathbf{H}, \mathbf{a}, \eta\}$, the distribution of each variable ζ_k is expressed as $\ln q(\zeta_k) \propto \langle \ln q(\zeta_k, \mathbf{X}) \rangle_{q(\zeta/\zeta_k)}$, where ζ/ζ_k represents the subset excluding ζ_k . More specifically, the posterior approximation of \mathbf{H} , \mathbf{a} , and η in this paper are, respectively, given by

$$\ln q(\mathbf{H}) \propto \langle \ln p(\mathbf{X}, \mathbf{H}, \mathbf{a}, \eta) \rangle_{q(\mathbf{a})}, \quad (15)$$

$$\ln q(\mathbf{a}) \propto \langle \ln p(\mathbf{X}, \mathbf{H}, \mathbf{a}, \eta) \rangle_{q(\mathbf{H})q(\eta)}, \quad (16)$$

$$\ln q(\eta) \propto \langle \ln p(\mathbf{X}, \mathbf{H}, \mathbf{a}, \eta) \rangle_{q(\mathbf{a})}. \quad (17)$$

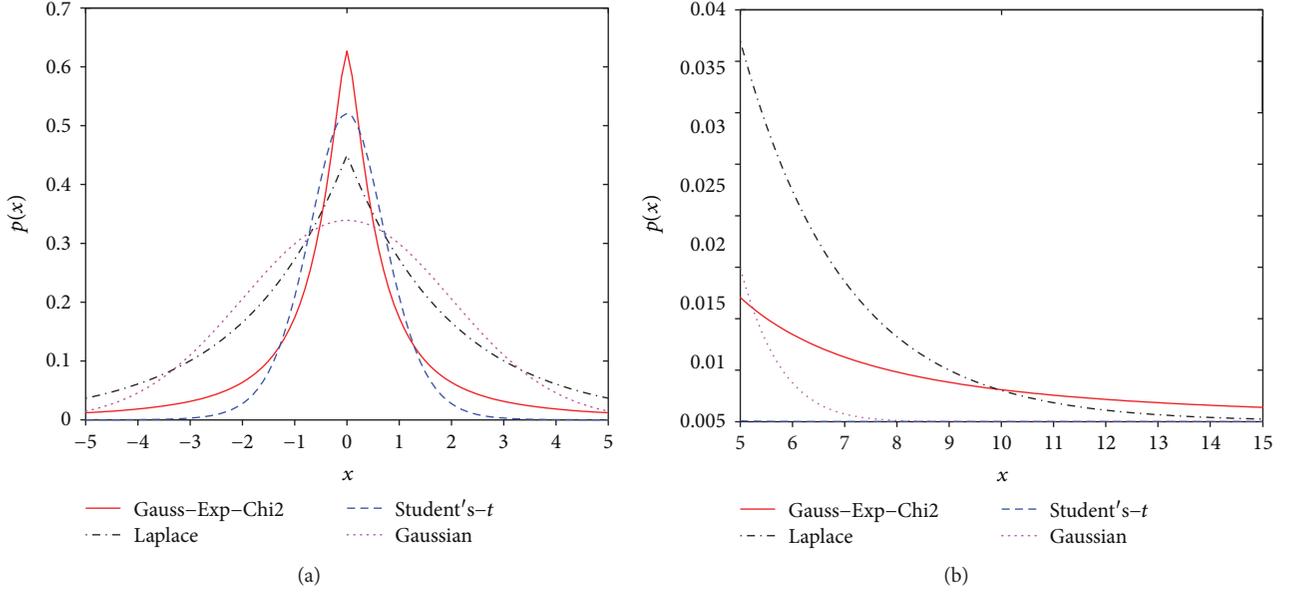


FIGURE 2: Four kinds of pdf curves with the standard derivation: (a) compares four distributions at origin and (b) compares four distributions in tail.

(1) *Update of $q(\mathbf{H})$.* Keeping the terms of $q(\mathbf{H})$ only depend on \mathbf{H} , we have

$$\begin{aligned} \ln q(\mathbf{H}) &\propto \langle \ln p(\mathbf{X}|\mathbf{H})p(\mathbf{H}|\mathbf{a}) \rangle_{q(\mathbf{a})} \\ &\propto \sum_{t=1}^T \sigma^{-2} \mathbf{h}^H(t) \widehat{\mathbf{A}}^H \widehat{\mathbf{A}} \mathbf{x}(t) - \frac{1}{2} \sigma^{-2} \mathbf{h}^H(t) \mathbf{\Lambda}^H \mathbf{h}(t), \end{aligned} \quad (18)$$

with $\mathbf{\Lambda} = \text{diag}(1/\alpha_1, 1/\alpha_2, \dots, 1/\alpha_J)$. By gathering together the similar terms, we have

$$\begin{aligned} \ln q(\mathbf{H}) &\propto \sum_{t=1}^T -\frac{1}{2} \mathbf{h}^H(t) \left(\sigma^{-2} \widehat{\mathbf{A}}^H \widehat{\mathbf{A}} + \langle \mathbf{\Lambda} \rangle \right) \mathbf{h}(t) \\ &\quad + \sigma^{-2} \sum_{t=1}^T \mathbf{h}^H(t) \widehat{\mathbf{A}}^H \mathbf{x}(t) \Rightarrow q(\mathbf{H}) \\ &\propto \sum_{t=1}^T \exp \left(-\frac{1}{2} \mathbf{h}^H(t) \Sigma^{-1} \mathbf{h}(t) + \mathbf{h}^H(t) \Sigma^{-1} \boldsymbol{\mu}(t) \right), \end{aligned} \quad (19)$$

which implies that $q(\mathbf{H})$ is the product of multiple Gaussian distributions $\mathcal{N}(\mathbf{h}(t) | \boldsymbol{\mu}(t), \Sigma)$ with the mean $\boldsymbol{\mu}(t)$ and covariance Σ being

$$\boldsymbol{\mu}(t) = \sigma^{-2} \Sigma^{-1} \widehat{\mathbf{A}}^H \mathbf{x}(t), \quad (20)$$

$$\Sigma = \left(\sigma^{-2} \widehat{\mathbf{A}}^H \widehat{\mathbf{A}} + \langle \mathbf{\Lambda} \rangle \right)^{-1}. \quad (21)$$

(2) *Update of $q(\mathbf{a})$.* Similarly, according to (16), $q(\mathbf{a})$ is derived as

$$\begin{aligned} \ln q(\mathbf{a}) &\propto \langle \ln p(\mathbf{H}|\mathbf{a})p(\mathbf{a}|\tau, \eta) \rangle_{q(\mathbf{H})q(\eta)} \\ &\propto \langle \ln \left[\prod_{t=1}^T \prod_{j=1}^J \frac{1}{\sqrt{2\pi\alpha_j}} \exp \left(-\frac{h_j^2(t)}{2\alpha_j} \right) \right. \\ &\quad \left. \cdot \prod_{j=1}^J \tau \eta \exp(-\tau \eta \alpha_j) \right] \rangle_{q(\mathbf{H})q(\eta)}. \end{aligned} \quad (22)$$

Denoting $h_j^2 = \sum_{t=1}^T h_j^2(t)$ and $\langle h_j^2 \rangle = \mu_j^2 + \Sigma_{jj}$, (22) can be further simplified to

$$\begin{aligned} \ln q(\mathbf{a}) &\propto \sum_{j=1}^J \langle \ln \left[\alpha_j^{-1/2} \exp \left(-\frac{h_j^2}{2\alpha_j} - \tau \eta \alpha_j \right) \right] \rangle_{q(\mathbf{H})q(\eta)} \\ &\Rightarrow q(\mathbf{a}) \propto \prod_{j=1}^J \alpha_j^{-1/2} \exp \left(-\frac{\langle h_j^2 \rangle}{2\alpha_j} - \tau < \eta > \alpha_j \right). \end{aligned} \quad (23)$$

The posterior distribution of $q(\mathbf{a})$ can be approximately equivalent to the product of a series of generalized inverse Gaussian (GIG) distributions, that is, $\mathbf{a} \sim \prod_{j=1}^J \text{GIG}(\alpha_j | 1/2 < h_j^2 > 2\tau < \eta >)$, and thus, we have

$$\langle \alpha_j^n \rangle = \left(\frac{\langle h_j^2 \rangle}{2\tau < \eta >} \right)^{n/2} \frac{\kappa_{n-1/2} \left(\sqrt{2\tau < \eta > \langle h_j^2 \rangle} \right)}{\kappa_{-1/2} \left(\sqrt{2\tau < \eta > \langle h_j^2 \rangle} \right)}, \quad (24)$$

where $\kappa_p(\cdot)$ is the third kind Bessel function with order p . The case of $n = 1$ in (24) gives the calculation of $\langle \alpha_j \rangle$, which can be used in (25). The case of $n = -1$ gives $\langle \alpha_j^{-1} \rangle$ used for the evaluation of Σ in (21).

Input: $\widehat{\mathbf{A}}, \mathbf{X}$
Output: $\boldsymbol{\mu}, \boldsymbol{\Sigma}$

1. **Initialize:** $\langle \alpha_j \rangle, \langle \alpha_j^{-1} \rangle, \langle \eta \rangle, \tau, \nu, \sigma^2, \varepsilon, i_{\max}$
2. **while** $\|\langle \alpha_j \rangle^{(i)} - \langle \alpha_j \rangle^{(i-1)}\| / \|\langle \alpha_j \rangle^{(i-1)}\| > \varepsilon$ and $i < i_{\max}$ **do**
3. update $\boldsymbol{\Sigma}^{(i)}$ using (21)
4. update $\boldsymbol{\mu}^{(i)}$ using (20)
5. update $\langle \alpha_j \rangle^{(i)}$ and $\langle \alpha_j^{-1} \rangle^{(i)}$ using (24)
6. update $\langle \eta \rangle^{(i)}$ using (26)
7. $i = i + 1$
8. **end while**

ALGORITHM 1: The main steps of the model parameter estimation.

(3) *Update of $q(\eta)$.* For $q(\eta)$, we have

$$\begin{aligned} \ln q(\eta) &\propto \langle \ln p(\boldsymbol{\alpha} | \tau, \eta) p(\eta | \nu) \rangle_{q(\boldsymbol{\alpha})} \\ &\propto \langle \ln \left[\prod_{j=1}^J \tau \eta \exp(-\tau \eta \alpha_j) \frac{2^{-\nu/2}}{\Gamma(\nu/2)} \eta^{\nu/2-1} \exp\left(-\frac{\eta}{2}\right) \right] \rangle_{q(\boldsymbol{\alpha})} \\ &\Rightarrow q(\eta) \propto \eta^{J+\nu/2-1} \exp\left(-\tau \sum_{j=1}^J \langle \alpha_j \rangle \eta - \frac{\eta}{2}\right). \end{aligned} \quad (25)$$

Thus, it can be known that $q(\eta)$ is a gamma distribution, that is, $\eta \sim \text{Gamma}(J + \nu/2, \tau \sum_{j=1}^J \langle \alpha_j \rangle + 1/2)$ with the following mean

$$\langle \eta \rangle = \frac{J + \nu/2}{\tau \sum_{j=1}^J \langle \alpha_j \rangle + 1/2}. \quad (26)$$

The estimation of model parameters is implemented by alternately updating the mean $\boldsymbol{\mu}$, variance $\boldsymbol{\Sigma}$, and hyperparameters $\langle \alpha_j \rangle$ and $\langle \eta \rangle$ to minimize the KL divergence between the true posterior and the variational approximation, and the main steps are summarized in Algorithm 1.

3.3. Source Power and DOA Estimation. The DOAs of the impinging sources are estimated via the following two steps: (1) form spatial spectra using the estimated source powers of all the potential directions and (2) extract the number and the corresponding locations of the peaks beyond a power threshold from the spectra. Let $\widehat{\mathbf{U}} = [\boldsymbol{\mu}(1), \boldsymbol{\mu}(2), \dots, \boldsymbol{\mu}(T)]$ and $\widehat{\boldsymbol{\Sigma}}$ be the final estimates of \mathbf{U} and $\boldsymbol{\Sigma}$ (the mean and variance with respect to \mathbf{H}) and consider \mathbf{H} row by row, then we have $\mathbf{h}_j \sim \mathcal{N}(\widehat{\mathbf{U}}_j, \widehat{\boldsymbol{\Sigma}}_{jj} \mathbf{I})$. As outlined in [18], let P_{o_j} be the source power of direction $\widehat{\boldsymbol{\theta}}_j$; then we have

$$\widehat{P}_{o_j} = \frac{1}{T} \left(\|\widehat{\mathbf{U}}_j\|_2^2 + T \widehat{\boldsymbol{\Sigma}}_{jj} \right) = \frac{1}{T} \|\widehat{\mathbf{U}}_j\|_2^2 + \widehat{\boldsymbol{\Sigma}}_{jj}. \quad (27)$$

Therefore, the source powers of all the potential directions in the angular space are estimated by calculating $\widehat{\mathbf{P}}_o = [\widehat{P}_{o_1}, \widehat{P}_{o_2}, \dots, \widehat{P}_{o_K}]^T$. Suppose that there are \mathbf{K} peaks

exceeding the power threshold P_{thres} and the corresponding grid indices are $j_k (j_k = 1, 2, \dots, K)$. Then, the estimated source number and DOAs will be K and $\widehat{\boldsymbol{\theta}}_{j_k} (j_k = 1, 2, \dots, K)$.

4. Discussions

4.1. Algorithm Optimization. For alleviating the computational complexity and speeding up the update efficiency, some implementation details are adopted for algorithm optimization in this subsection.

4.1.1. Woodbury Matrix Identity for Dimension-Reduction. At each iteration, it is inevitable to calculate a $J \times J$ matrix inversion when updating $\boldsymbol{\Sigma}$ according to (21), which requires $\mathcal{O}(J^3)$ computations. To reduce the computational complexity, we adopt the Woodbury matrix identity, and then (21) can be rewritten as $\boldsymbol{\Sigma} = \langle \boldsymbol{\Lambda}^{-1} \rangle - \langle \boldsymbol{\Lambda}^{-1} \rangle \widehat{\mathbf{A}}^H \mathbf{G}^{-1} \widehat{\mathbf{A}} \langle \boldsymbol{\Lambda}^{-1} \rangle$ with an $M \times M$ matrix $\mathbf{G} = \sigma^{-2} \mathbf{I} + \widehat{\mathbf{A}} \langle \boldsymbol{\Lambda}^{-1} \rangle \widehat{\mathbf{A}}^H$. Thus, the computational complexity is reduced to $\mathcal{O}(M^3)$.

4.1.2. Pruning a Basis Function. In order to speed up the update efficiency, $\langle \alpha_j \rangle$ and the corresponding $\mathbf{a}(\widehat{\boldsymbol{\theta}}_j)$ would be pruned from the original sparse Bayesian model when $\langle \alpha_j \rangle$ is smaller than a certain threshold. This procedure is referred to as ‘‘pruning a basis function’’, and similar approaches have been used in [29, 30].

4.1.3. The Third Kind Bessel Function Approximation. It is known that (24) involves the third kind Bessel function, and this integral process is relatively complicated. According to [31], the third kind Bessel function of m can be approximated as $\kappa_p(m) \sim 1/2\Gamma(p)(m/2)^{-p}$, if $p > 0$, and $\kappa_p(m) \sim 1/2\Gamma(-p)(m/2)^p$, if $p < 0$. Then, (24) can be simplified to

$$\langle \alpha_j^n \rangle = \begin{cases} \left(\frac{\langle h_j^2 \rangle}{2\tau \langle \eta \rangle} \right)^{1/2}, & n = 1, \\ \left(\frac{\langle h_j^2 \rangle}{2\tau \langle \eta \rangle} \right)^{-1/2} \frac{2\Gamma(3/2)}{\Gamma(1/2)} \sqrt{\frac{1}{2\tau \langle \eta \rangle \langle h_j^2 \rangle}}, & n = -1. \end{cases} \quad (28)$$

4.2. *Cramér-Rao Bound (CRB)*. To evaluate the DOA estimation performance, the CRB of the DOA estimation for a mixture of uncorrelated and coherent sources is derived in this subsection. Consider the sparse form of the array output vector $\mathbf{x}(t)$ mentioned in (2), the log-likelihood function of $\mathbf{x}(1), \mathbf{x}(2), \dots, \mathbf{x}(T)$ is firstly construct as

$$\begin{aligned} \ln \mathcal{L}(\mathbf{x}(1), \mathbf{x}(2), \dots, \mathbf{x}(T)) \\ = \text{const} - MT \ln \sigma - \sigma^{-1} \sum_{t=1}^T \left(\mathbf{x}^*(t) - \mathbf{h}^*(t) \widehat{\mathbf{A}}^* \right) \\ \cdot \left(\mathbf{x}(t) - \widehat{\mathbf{A}} \mathbf{h}(t) \right). \end{aligned} \quad (29)$$

Let $\mathbf{h}_r(t)$ and $\mathbf{h}_i(t)$ be the real part and imaginary part of $\mathbf{h}(t)$, that is, $\mathbf{h}_r(t) \triangleq \text{Re} \{ \mathbf{h}(t) \}$ and $\mathbf{h}_i(t) \triangleq \text{Im} \{ \mathbf{h}(t) \}$. Then, the partial derivatives of (29) with respect to $\mathbf{h}_r(t)$, $\mathbf{h}_i(t)$, σ , and $\widehat{\boldsymbol{\theta}}$ are, respectively, given as

$$\begin{aligned} \frac{\partial \ln \mathcal{L}}{\partial \mathbf{h}_r(t)} &= 2\sigma^{-1} \text{Re} \left\{ \widehat{\mathbf{A}}^* \mathbf{n}(t) \right\}, \\ \frac{\partial \ln \mathcal{L}}{\partial \mathbf{h}_i(t)} &= 2\sigma^{-1} \text{Im} \left\{ \widehat{\mathbf{A}}^* \mathbf{n}(t) \right\}, \\ \frac{\partial \ln \mathcal{L}}{\partial \sigma} &= MT\sigma^{-1} + \sigma^{-2} \sum_{t=1}^T \mathbf{n}^*(t) \mathbf{n}(t), \\ \frac{\partial \ln \mathcal{L}}{\partial \widehat{\boldsymbol{\theta}}} &= 2\sigma^{-1} \sum_{t=1}^T \text{Re} \{ \mathbf{H}_d^*(t) \mathbf{A}_D^* \mathbf{n}(t) \}, \end{aligned} \quad (30)$$

where $\mathbf{H}_d^*(t) = \text{diag} \{ h_1(t), h_2(t), \dots, h_J(t) \}$ and $\mathbf{A}_D = [d\mathbf{a}(\widehat{\theta}_1)/d\widehat{\theta}_1, d\mathbf{a}(\widehat{\theta}_2)/d\widehat{\theta}_2, \dots, d\mathbf{a}(\widehat{\theta}_J)/d\widehat{\theta}_J]$. Subsequently, define

$$\begin{aligned} \mathfrak{N} &= E \left\{ \frac{\partial \ln \mathcal{L}}{\partial \widehat{\boldsymbol{\theta}}} \left(\frac{\partial \ln \mathcal{L}}{\partial \widehat{\boldsymbol{\theta}}} \right)^T \right\} \\ &= 2\sigma^{-1} \sum_{t=1}^T \text{Re} \{ \mathbf{H}_d^*(t) \mathbf{A}_D^* \mathbf{A}_D \mathbf{H}_d(t) \}, \\ \boldsymbol{\Xi}_r^m &= \text{Re} \{ \boldsymbol{\Xi}^m \} = E \left\{ \frac{\partial \ln \mathcal{L}}{\partial \mathbf{h}_r(t)} \left(\frac{\partial \ln \mathcal{L}}{\partial \widehat{\boldsymbol{\theta}}} \right)^T \right\} \\ &= 2\sigma^{-1} \text{Re} \left\{ \widehat{\mathbf{A}}^* \mathbf{A}_D \mathbf{H}_d(t) \right\}, \\ \boldsymbol{\Xi}_i^m &= \text{Im} \{ \boldsymbol{\Xi}^m \} = E \left\{ \frac{\partial \ln \mathcal{L}}{\partial \mathbf{h}_i(t)} \left(\frac{\partial \ln \mathcal{L}}{\partial \widehat{\boldsymbol{\theta}}} \right)^T \right\} \\ &= 2\sigma^{-1} \text{Im} \left\{ \widehat{\mathbf{A}}^* \mathbf{A}_D \mathbf{H}_d(t) \right\}, \end{aligned}$$

$$\begin{aligned} \mathbf{Q}_r &= E \left\{ \frac{\partial \ln \mathcal{L}}{\partial \mathbf{h}_r(t)} \left(\frac{\partial \ln \mathcal{L}}{\partial \mathbf{h}_r(k)} \right)^T \right\}^{-1} \\ &= E \left\{ \frac{\partial \ln \mathcal{L}}{\partial \mathbf{h}_i(t)} \left(\frac{\partial \ln \mathcal{L}}{\partial \mathbf{h}_i(k)} \right)^T \right\}^{-1} \\ &= \left(2\sigma^{-1} \text{Re} \left\{ \widehat{\mathbf{A}}^* \widehat{\mathbf{A}} \right\} \delta_{kt} \right)^{-1}, \\ \mathbf{Q}_i &= -E \left\{ \frac{\partial \ln \mathcal{L}}{\partial \mathbf{h}_r(t)} \left(\frac{\partial \ln \mathcal{L}}{\partial \mathbf{h}_i(k)} \right)^T \right\}^{-1} \\ &= \left(2\sigma^{-1} \text{Im} \left\{ \widehat{\mathbf{A}}^* \widehat{\mathbf{A}} \right\} \delta_{kt} \right)^{-1}. \end{aligned} \quad (31)$$

Then, the Fisher information matrix (FIM) [32] can be obtained as

$$\begin{aligned} \text{FIM} &= \mathfrak{N} - \left[\boldsymbol{\Xi}_r^1 \boldsymbol{\Xi}_i^1 \boldsymbol{\Xi}_r^2 \boldsymbol{\Xi}_i^2 \dots \boldsymbol{\Xi}_r^M \boldsymbol{\Xi}_i^M \right] \\ &= \begin{bmatrix} \mathbf{Q}_r & -\mathbf{Q}_i & 0 & \dots & 0 \\ \mathbf{Q}_i & \mathbf{Q}_r & -\mathbf{Q}_i & \ddots & \vdots \\ 0 & \mathbf{Q}_i & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \mathbf{Q}_r & -\mathbf{Q}_i \\ 0 & \dots & 0 & \mathbf{Q}_i & \mathbf{Q}_r \end{bmatrix} \begin{bmatrix} \boldsymbol{\Xi}_r^1 \\ \boldsymbol{\Xi}_i^1 \\ \boldsymbol{\Xi}_r^2 \\ \boldsymbol{\Xi}_i^2 \\ \vdots \\ \boldsymbol{\Xi}_r^M \\ \boldsymbol{\Xi}_i^M \end{bmatrix}. \end{aligned} \quad (32)$$

It is noted that the following equality holds

$$\left[\boldsymbol{\Xi}_r^T \boldsymbol{\Xi}_i^T \right] \begin{bmatrix} \mathbf{Q}_r & -\mathbf{Q}_i \\ \mathbf{Q}_i & \mathbf{Q}_r \end{bmatrix} \begin{bmatrix} \boldsymbol{\Xi}_r \\ \boldsymbol{\Xi}_i \end{bmatrix} = \text{Re} \{ \boldsymbol{\Xi}^* \mathbf{Q} \boldsymbol{\Xi} \}, \quad (33)$$

with $\mathbf{Q} = (2\sigma^{-1} \widehat{\mathbf{A}}^* \widehat{\mathbf{A}})^{-1}$ and $\boldsymbol{\Xi} = 2\sigma^{-1} \widehat{\mathbf{A}}^* \mathbf{A}_D \mathbf{H}_d(t)$. Finally, by combing (32) and (33), we have

$$\begin{aligned} \text{CRB}(\widehat{\boldsymbol{\theta}}) &= \text{FIM}^{-1} = 0.5\sigma \\ &\cdot \left(\sum_{t=1}^T \text{Re} \left\{ \mathbf{H}_d(t) \mathbf{A}_D^* \left(\mathbf{I} - \widehat{\mathbf{A}} \left(\widehat{\mathbf{A}}^* \widehat{\mathbf{A}} \right)^{-1} \widehat{\mathbf{A}}^* \right) \mathbf{A}_D \mathbf{H}_d(t) \right\} \right)^{-1}. \end{aligned} \quad (34)$$

5. Simulation Results

In this section, several simulations are presented to illustrate the performance of the proposed algorithm as compared with the BCS [25] and NP-1 [27] algorithms. Consider a 9-sensor ULA with sensor interspacing being a half of the carrier wavelength. In the proposed algorithm, hyperparameters $\boldsymbol{\alpha}$ and η are initialized as $1/T \sum_{t=1}^T \widehat{\mathbf{A}}^H \mathbf{x}(t)$ and 0.1; the values of rate parameter τ , shape parameter ν , and power

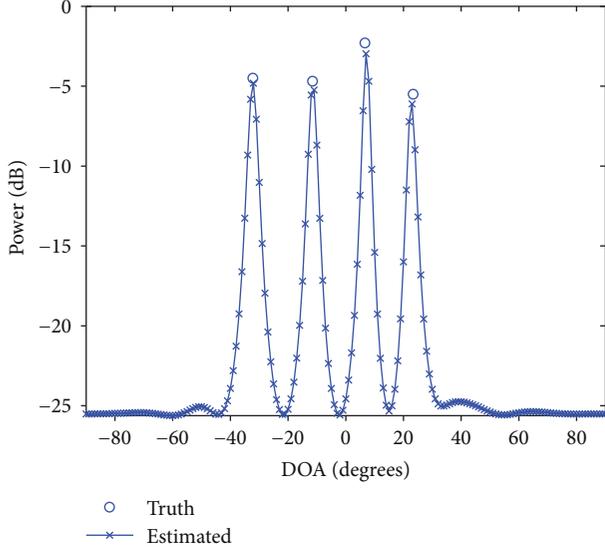


FIGURE 3: Spectra for the proposed algorithm with the fixed SNR 0 dB and number of snapshots 200.

threshold P_{thres} are, respectively, set to be $\tau = 1.5$, $\nu = 1$, and $P_{\text{thres}} = -10$ dB. Two hundred independent Monte Carlo trials are conducted for the following simulations, and the success rate and root mean squared error (RMSE) are two significant performance metrics for evaluating the DOA estimation performance. The success rate is defined as the rate between the number of successful estimates and the total number of Monte Carlo trials, where a successful estimate is declared if the estimation error is within a certain small Euclidean distance of the true directions. The RMSE is defined by

$$\text{RMSE} = \sqrt{\frac{1}{200K} \sum_{i=1}^{200} \sum_{k=1}^K (\hat{\theta}_{k,i} - \theta_k)^2}, \quad (35)$$

where $\hat{\theta}_{k,i}$ is the estimates of θ_k in the i th Monte Carlo trial.

In the first simulation, we evaluate the effectiveness of the proposed algorithm. Assume that two uncorrelated sources from $\{-32.2^\circ, -11.5^\circ\}$ and two coherent sources from $\{6.7^\circ, 23.4^\circ\}$ with fading coefficients $[1, 0.2891 - 0.7567j]$ impinge on this ULA. The grid number, SNR, and number of snapshots are, respectively, set to be 361, 0 dB, and 200. Figure 3 plots the source power spectrum, which is a function of DOA.

As can be seen from Figure 3, the power peaks in the vicinity of the true DOAs can be distinguished from other peaks by their relatively strong powers and the biases between these peaks and the true DOAs are slight, which demonstrate the effectiveness of the proposed algorithm.

The second simulation compares the estimation performance of three algorithms, including the proposed algorithm, BCS, and NP-1 algorithms. Consider one uncorrelated source from 19.2° and two coherent sources from $\{-25.5^\circ, 33.3^\circ\}$ with fading coefficients being $[1, -0.5280 +$

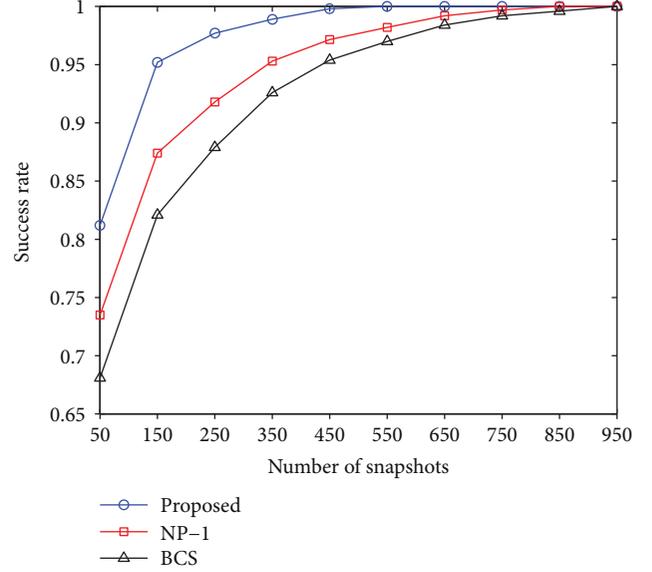


FIGURE 4: Success rates versus number of snapshots with the fixed SNR 0 dB.

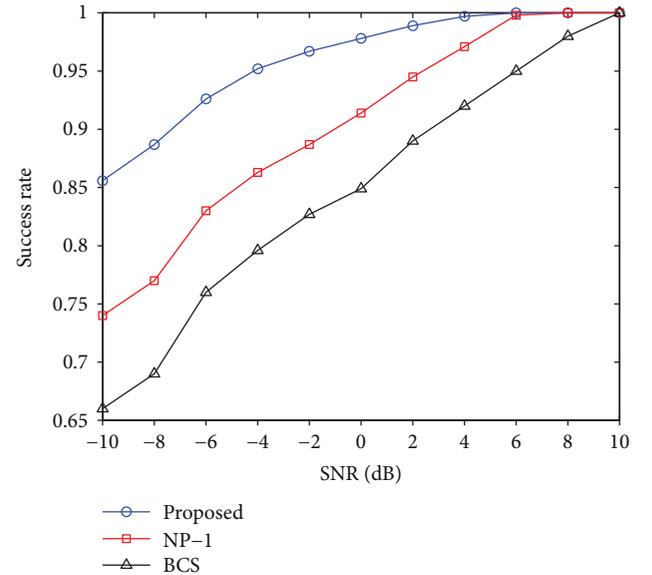


FIGURE 5: Success rates versus SNR with the fixed number of snapshots 200.

$0.6010j]$. The grid number and SNR are, respectively, fixed at 361 and 0 dB. The success rate versus number of snapshots and SNR are depicted in Figures 4 and 5, whereas the RMSE versus number of snapshots and SNR are shown in Figures 6 and 7. The results from Figures 4 and 5 illustrate that all the three algorithms are able to estimate the DOA correctly under the condition that the SNR or number of snapshots are relatively high, but for low SNR or small number of snapshots, the proposed algorithm has a higher success rate than the BCS and NP-1 algorithms. Figures 6 and 7 show that the estimation accuracy of the three algorithms tends to improve with the increase of SNR or the number of snapshots, and the

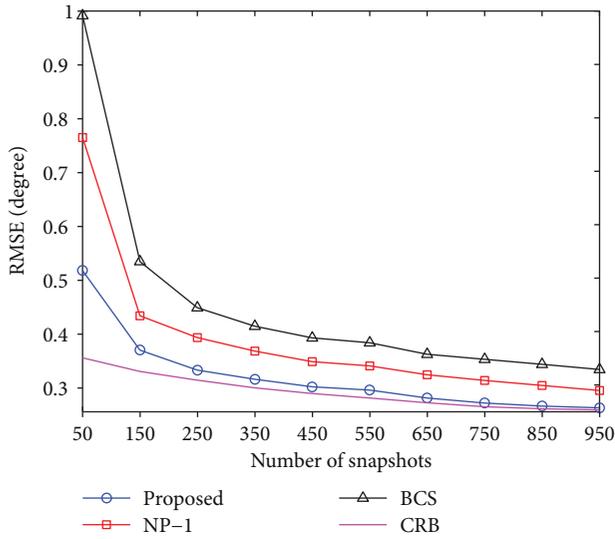


FIGURE 6: RMSE versus number of snapshots with the fixed SNR 0 dB.

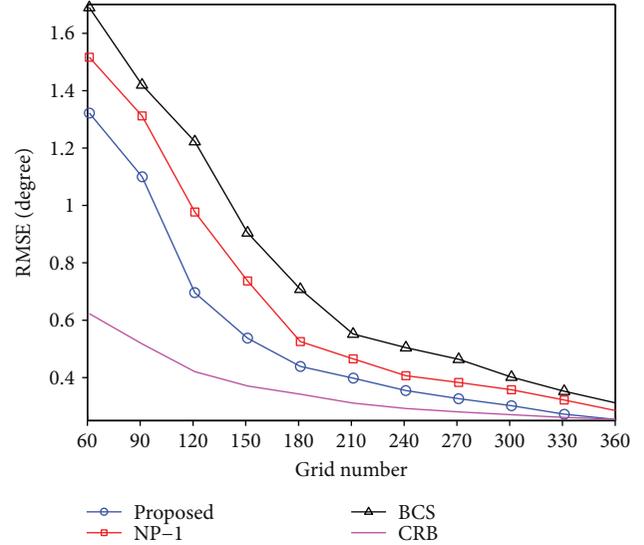


FIGURE 8: RMSE versus grid number with the fixed SNR 10 dB and number of snapshots 500.

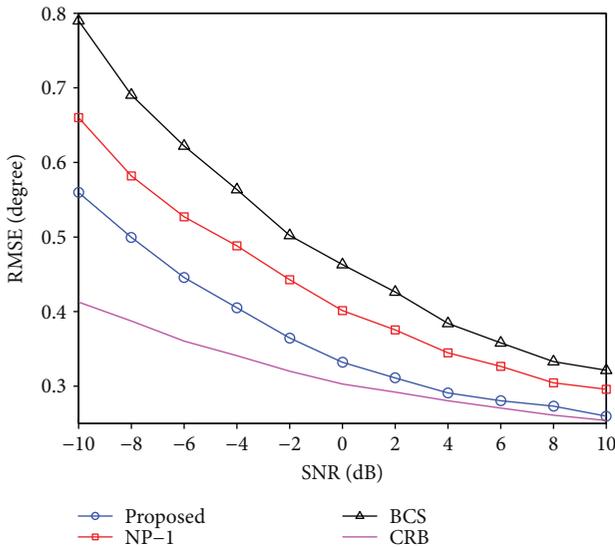


FIGURE 7: RMSE versus SNR with the fixed number of snapshots 200.

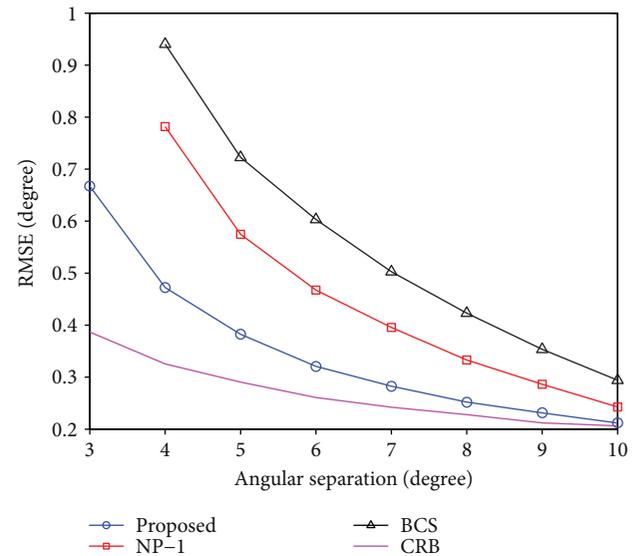


FIGURE 9: RMSE versus angular separation with the fixed SNR 10 dB and number of snapshots 500 for coherent sources.

proposed algorithm yields the best estimation accuracy as compared to the other two algorithms. This is because the proposed Gauss-Exp- χ^2 prior is a three-layer sparsity-encouraging prior, which can help to improve the accuracy of source reconstruction.

The third simulation investigates the RMSE versus grid number. The simulation settings are the same as those of the second simulation, except that grid number in this simulation is ranged from 61 to 361. Figure 8 plots the RMSE curves versus grid number with fixed SNR 10 dB and number of snapshots 500.

It is observed that the RMSE curves of the three algorithms rapidly decrease as grid number increases from 61 to 181, and these curves tend to slowly decrease with the

increase of grid number when grid number is beyond 181. In addition, the proposed algorithm has minimum estimation errors among the three algorithms. Note that the three algorithms recover sparse sources from a Bayesian perspective; thus, they are not restricted to the restricted isometry property. This implies that the grid number can continue to increase to further improve the estimation accuracy and the computational complexity would increase accordingly. Therefore, for the purpose of balancing estimation accuracy and computational complexity, the reasonable range of grid number is from 181 to 361.

In the last simulation, we test the angular resolution of the three algorithms. Consider two coherent sources from

10.1° and 10.1° + Δθ with the angular separation Δθ ranging from 3° to 10°. The fading coefficients are [1, 0.4469 – 0.7696j], and the grid number is set to be 361. Figure 9 depicts the RMSE versus angular separation with the fixed SNR 10 dB and number of snapshots 500. The results from Figure 9 show that both the BCS and NP-1 algorithms fail to work when the angular separation is less than 4°, whereas the proposed algorithm can still provide accurate DOA estimation as long as the angular separation is no less than 3°. Moreover, the proposed algorithm has smaller RMSE than the BCS and NP-1 algorithms at the same angular separation. Thus, we can say that the proposed algorithm has a higher angular resolution than the BCS and NP-1 algorithms.

6. Conclusion

In this paper, we develop a DOA estimation algorithm for a mixture of uncorrelated and coherent sources using SBI. In the Bayesian framework, a Gauss-Exp-Chi² prior is introduced to promote sparse solutions, and the corresponding three-layer hierarchical Bayesian model is established. Then, the model parameters are estimated via variational Bayesian approximations. Finally, we form the source power spectra by using the estimated model parameters, from which the number and the locations of the highest peaks are extracted to achieve source number and DOA estimates. Simulation results show that the proposed algorithm can effectively estimate the DOAs of mixed sources and has better estimation performance than the state-of-the-art BCS algorithm and NP-1 algorithm in terms of estimation accuracy, success rate, and angular resolution.

Appendix

A. The Derivation of the Probability Density Function for Gauss-Exp-Chi² Prior

In this Appendix, we prove that (12) holds.

$$\begin{aligned}
 p(\mathbf{H}; \tau, \nu) &= \prod_{t=1}^T \prod_{j=1}^J p(h_j(t); \tau, \nu) \\
 &= \prod_{t=1}^T \prod_{j=1}^J \int_0^{+\infty} p(h_j(t)|\alpha_j) p(\alpha_j|\tau, \nu) d\alpha_j \\
 &= \prod_{t=1}^T \prod_{j=1}^J \sqrt{\frac{\tau}{\pi}} \frac{\Gamma(\nu/2 + 1) \Gamma(\nu/2 + 1/2)}{\Gamma(\nu/2)} U_{chf} \\
 &\quad \cdot \left(\frac{\nu}{2} + \frac{1}{2}, \frac{1}{2}, h_j^2(t)\tau \right).
 \end{aligned} \tag{A1}$$

The expression on the right hand side of (12) can be rewritten as

$$\begin{aligned}
 p(h_j(t); \tau, \nu) &= \int_0^{+\infty} p(h_j(t)|\alpha_j) \\
 &\quad \cdot \left(\int_0^{+\infty} p(\alpha_j|\tau, \eta) p(\eta|\nu) d\eta \right) d\alpha_j.
 \end{aligned} \tag{A2}$$

Thus, we first calculate the integral over η

$$\begin{aligned}
 &\int_0^{+\infty} p(\alpha_j|\tau, \eta) p(\eta|\nu) d\eta \\
 &= \int_0^{+\infty} \tau \eta \exp(-\tau \eta \alpha_j) \frac{2^{-\nu/2}}{\Gamma(\nu/2)} \eta^{\nu/2-1} \exp\left(-\frac{\eta}{2}\right) d\eta \\
 &= \frac{2^{-\nu/2} \tau}{\Gamma(\nu/2)} \int_0^{+\infty} \eta^{\nu/2} \exp\left(-\tau \eta \alpha_j - \frac{\eta}{2}\right) d\eta \\
 &= \frac{\Gamma(\nu/2 + 1) 2^{-\nu/2} \tau}{\Gamma(\nu/2)} \left(\tau \alpha_j + \frac{1}{2} \right)^{-(\nu/2+1)}.
 \end{aligned} \tag{A3}$$

By substituting (A3) into (A2), we have

$$\begin{aligned}
 p(h_j(t); \tau, \nu) &= \int_0^{+\infty} \int_0^{+\infty} p(h_j(t)|\alpha_j) p(\alpha_j|\tau, \eta) p(\eta|\nu) d\alpha_j d\eta \\
 &= \bar{\omega}(\tau, \nu) \int_0^{+\infty} \frac{1}{\sqrt{2\pi\alpha_j}} \exp\left(-\frac{h_j^2(t)}{2\alpha_j}\right) \\
 &\quad \cdot \left(\tau \alpha_j + \frac{1}{2} \right)^{-(\nu/2+1)} d\alpha_j,
 \end{aligned} \tag{A4}$$

where $\bar{\omega}(\tau, \nu) = \Gamma(\nu/2 + 1) 2^{-\nu/2} \tau / \Gamma(\nu/2)$. Let $\beta_j = 1/\alpha_j$ (A4) can be expressed as

$$\begin{aligned}
 p(h_j(t); \tau, \nu) &= \frac{\bar{\omega}(\tau, \nu)}{\sqrt{2\pi}} \int_0^{+\infty} \beta_j^{3/2} \exp\left(-\frac{h_j^2(t)}{2}\beta_j\right) \\
 &\quad \cdot \left(\tau \beta_j^{-1} + \frac{1}{2} \right)^{-(\nu/2+1)} d\beta_j \\
 &= \frac{\bar{\omega}(\tau, \nu)}{\sqrt{2\pi}} \tau^{1/2} 2^{(\nu/2-1/2)} \int_0^{+\infty} \tilde{\beta}_j^{(\nu/2-1/2)} \\
 &\quad \cdot \left(1 + \tilde{\beta}_j \right)^{-(\nu/2+1)} \exp\left(-h_j^2(t)\tau \tilde{\beta}_j\right) d\tilde{\beta}_j,
 \end{aligned} \tag{A5}$$

where $\tilde{\beta}_j = \beta_j/2\tau$. According to the definition of the confluent hypergeometric function, (A5) can be further simplified to

$$\begin{aligned}
 p(h_j(t); \tau, \nu) &= \sqrt{\frac{\tau}{\pi}} \frac{\Gamma(\nu/2 + 1) \Gamma(\nu/2 + 1/2)}{\Gamma(\nu/2)} U_{chf} \\
 &\quad \cdot \left(\frac{\nu}{2} + \frac{1}{2}, \frac{1}{2}, h_j^2(t)\tau \right).
 \end{aligned} \tag{A6}$$

Thus, we have

$$\begin{aligned}
 p(\mathbf{H}; \tau, \nu) &= \prod_{t=1}^T \prod_{j=1}^J p(h_j(t); \tau, \nu) \\
 &= \prod_{t=1}^T \prod_{j=1}^J \sqrt{\frac{\tau}{\pi}} \frac{\Gamma(\nu/2 + 1) \Gamma(\nu/2 + 1/2)}{\Gamma(\nu/2)} U_{chf} \\
 &\quad \cdot \left(\frac{\nu}{2} + \frac{1}{2}, \frac{1}{2}, h_j^2(t)\tau \right).
 \end{aligned} \tag{A7}$$

Conflicts of Interest

The authors declare that they have no competing interests.

Authors' Contributions

Pinjiao Zhao proposed the main idea and conceived the proposed approach. Pinjiao Zhao and Weijian Si discussed and designed the proposed algorithm. Pinjiao Zhao performed the experiments and wrote the paper. Guobing Hu and Liwei Wang reviewed and revised the manuscript. All authors read and approved the manuscript.

Acknowledgments

This work was financially supported by the High Level Talent Research Starting Project of Jinling Institute of Technology (Grant no. jit-b-201724), the National Natural Science Foundation of China (Grant no. 61671168), and the Natural Science Foundation of the Jiangsu Province (Project no. BK20161104) and the Six Talent Peaks Project of the Jiangsu Province (Project no. DZXX-022).

References

- [1] H. Krim and M. Viberg, "Two decades of array signal processing research: the parametric approach," *IEEE Signal Processing Magazine*, vol. 13, no. 4, pp. 67–94, 1996.
- [2] W. Si, P. Zhao, Z. Qu, and Y. Wang, "Real-valued DOA estimation for a mixture of uncorrelated and coherent sources via unitary transformation," *Digital Signal Processing*, vol. 58, pp. 102–114, 2016.
- [3] T. Meng, M. Wu, and N. Yuan, "DOA estimation of noncircular signals using quaternions," *International Journal of Antennas and Propagation*, vol. 2017, Article ID 9848169, 6 pages, 2017.
- [4] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Transactions on Antennas and Propagation*, vol. 34, no. 3, pp. 276–280, 1986.
- [5] R. Roy and T. Kailath, "ESPRIT-estimation of signal parameters via rotational invariance techniques," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 37, no. 7, pp. 984–995, 1989.
- [6] W. Si, P. Zhao, Z. Qu, and L. Wang, "Computationally efficient angle and polarization estimation in the presence of multipath propagation using dual-polarization vector sensor array," *International Journal of Antennas and Propagation*, vol. 2016, Article ID 7537160, 15 pages, 2016.
- [7] T.-J. Shan, M. Wax, and T. Kailath, "On spatial smoothing for direction-of-arrival estimation of coherent signals," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 33, no. 4, pp. 806–811, 1985.
- [8] F.-M. Han and X.-D. Zhang, "An ESPRIT-like algorithm for coherent DOA estimation," *IEEE Antennas and Wireless Propagation Letters*, vol. 4, pp. 443–446, 2005.
- [9] X. Xu, Z. Ye, and J. Peng, "Method of direction-of-arrival estimation for uncorrelated, partially correlated and coherent sources," *IET Microwaves, Antennas & Propagation*, vol. 1, no. 4, pp. 949–954, 2007.
- [10] Z. Ye, Y. Zhang, and C. Liu, "Direction-of-arrival estimation for uncorrelated and coherent signals with fewer sensors," *IET Microwaves, Antennas & Propagation*, vol. 3, no. 3, pp. 473–482, 2009.
- [11] X. Xu, Z. Ye, Y. Zhang, and C. Chang, "A deflation approach to direction of arrival estimation for symmetric uniform linear array," *IEEE Antennas and Wireless Propagation Letters*, vol. 5, no. 1, pp. 486–489, 2006.
- [12] Z. Ye and X. Xu, "DOA estimation by exploiting the symmetric configuration of uniform linear array," *IEEE Transactions on Antennas and Propagation*, vol. 55, no. 12, pp. 3716–3720, 2007.
- [13] N. Yilmazer, J. Koh, and T. K. Sarkar, "Utilization of a unitary transform for efficient computation in the matrix pencil method to find the direction of arrival," *IEEE Transactions on Antennas and Propagation*, vol. 54, no. 1, pp. 175–181, 2006.
- [14] G. Karabulut, T. Kurt, and A. Yongaçoglu, "Estimation of directions of arrival by matching pursuit (EDAMP)," *EURASIP Journal on Wireless Communications and Networking*, vol. 2005, no. 2, 205 pages, 2005.
- [15] D. Malioutov, M. Çetin, and A. S. Willsky, "A sparse signal reconstruction perspective for source localization with sensor arrays," *IEEE Transactions on Signal Processing*, vol. 53, no. 8, pp. 3010–3022, 2005.
- [16] X. Xu, X. Wei, and Z. Ye, "DOA estimation based on sparse signal recovery utilizing weighted l_1 -norm penalty," *IEEE Signal Processing Letters*, vol. 19, no. 3, pp. 155–158, 2012.
- [17] M. E. Tipping, "Sparse Bayesian learning and the relevance vector machine," *Journal of Machine Learning Research*, vol. 1, no. 6, pp. 211–244, 2001.
- [18] 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.
- [19] Z. Zhang and B. D. Rao, "Sparse signal recovery with temporally correlated source vectors using sparse Bayesian learning," *IEEE Journal of Selected Topics in Signal Processing*, vol. 5, no. 5, pp. 912–926, 2011.
- [20] Z. Sadeghigol, M. H. Kahaei, and F. Haddadi, "Model based variational Bayesian compressive sensing using heavy tailed sparse prior," *Signal Processing: Image Communication*, vol. 41, pp. 158–167, 2016.
- [21] W. Si, X. Qu, Z. Qu, and P. Zhao, "Off-grid DOA estimation via real-valued sparse Bayesian method in compressed sensing," *Circuits, Systems, and Signal Processing*, vol. 35, no. 10, pp. 3793–3809, 2016.
- [22] D. J. C. MacKay, "Bayesian interpolation," *Neural Computation*, vol. 4, no. 3, pp. 415–447, 1992.
- [23] D. Tzikas, A. Likas, and N. Galatsanos, "The variational approximation for Bayesian inference," *IEEE Signal Processing Magazine*, vol. 25, no. 6, pp. 131–146, 2008.
- [24] S. D. Babacan, S. Nakajima, and M. N. Do, "Bayesian group-sparse modeling and variational inference," *IEEE Transactions on Signal Processing*, vol. 62, no. 11, pp. 2906–2921, 2014.
- [25] S. Ji, Y. Xue, and L. Carin, "Bayesian compressive sensing," *IEEE Transactions on Signal Processing*, vol. 56, no. 6, pp. 2346–2356, 2008.
- [26] S. D. Babacan, R. Molina, and A. K. Katsaggelos, "Bayesian compressive sensing using Laplace priors," *IEEE Transactions on Image Processing*, vol. 19, no. 1, pp. 53–63, 2010.
- [27] Z. Zhou, K. Liu, and J. Fang, "Bayesian compressive sensing using normal product priors," *IEEE Signal Processing Letters*, vol. 22, no. 5, pp. 583–587, 2015.

- [28] Z. Sabetsarvestani and H. Amindavar, "Sparse estimation with generalized beta mixture and the horseshoe prior," arXiv preprint arXiv:1411.2405, 2014.
- [29] N. Pedersen, D. Shutin, C. Manchón, and B. H. Fleury, "Sparse estimation using Bayesian hierarchical prior modeling for real and complex models," *IEEE Transactions on Signal Processing*, 2011.
- [30] R. G. Baraniuk, V. Cevher, M. F. Duarte, and C. Hegde, "Model-based compressive sensing," *IEEE Transactions on Information Theory*, vol. 56, no. 4, pp. 1982–2001, 2010.
- [31] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions: with Formulas, Graphs, and Mathematical Tables*, Courier Corporation, 1964.
- [32] Y. Gu and N. A. Goodman, "Information-theoretic compressive sensing kernel optimization and Bayesian Cramér-Rao bound for time delay estimation," *IEEE Transactions on Signal Processing*, vol. 65, no. 17, pp. 4525–4537, 2017.



Hindawi

Submit your manuscripts at
www.hindawi.com

