TSINGHUA SCIENCE AND TECHNOLOGY
ISSN 1007-0214 10/21 pp1069–1081
DOI: 10.26599/TST.2023.9010114
Volume 29, Number 4, August 2024

Combined UAMP and MF Message Passing Algorithm for Multi-Target Wideband DOA Estimation with Dirichlet Process Prior

Shanwen Guan, Xinhua Lu, Ji Li, Rushi Lan, and Xiaonan Luo*

Abstract: When estimating the direction of arrival (DOA) of wideband signals from multiple sources, the performance of sparse Bayesian methods is influenced by the frequency bands occupied by signals in different directions. This is particularly true when multiple signal frequency bands overlap. Message passing algorithms (MPA) with Dirichlet process (DP) prior can be employed in a sparse Bayesian learning (SBL) framework with high precision. However, existing methods suffer from either high complexity or low precision. To address this, we propose a low-complexity DOA estimation algorithm based on a factor graph. This approach introduces two strong constraints via a stretching transformation of the factor graph. The first constraint separates the observation from the DP prior, enabling the application of the unitary approximate message passing (UAMP) algorithm for simplified inference and mitigation of divergence issues. The second constraint compensates for the deviation in estimation angle caused by the grid mismatch problem. Compared to state-of-the-art algorithms, our proposed method offers higher estimation accuracy and lower complexity.

Key words: wideband direction of arrival (DOA) estimation; sparse Bayesian learning (SBL); unitary approximate message passing (UAMP) algorithm; Dirichlet process (DP)

1 Introduction

Wideband signals, characterized by their high information content and interference resistance, are widely employed in technologies such as radar, sonar, and wireless communication[1, 2]. One important aspect of working with these signals is determining their direction of arrival (DOA)[3]. As a result, the problem of DOA estimation for wideband signals has been extensively studied in the literature for various localization applications[4]. Compared to narrowband DOA estimation, the steering matrices for wideband DOA estimation increase with frequency, making the estimation process more complex.

In order to apply traditional narrowband DOA estimation methods to broadband scenarios, the received signal is first decomposed into narrowband signals using either a filter bank or the discrete Fourier transform (DFT). Subsequently, traditional methods for narrowband DOA estimation can be applied to each subband corresponding to different frequencies. The beamforming method estimates the angle accurately by using the weighted sum of array measurements. However, this method is susceptible to the aliasing effect and low resolution[5]. To counter these shortcomings, various alternatives have been proposed for wideband DOA estimation, including incoherent/coherent signal subspace methods (ISSMs/CSSMs)[6–8]

• Shanwen Guan, Ji Li, Rushi Lan, and Xiaonan Luo are with Guangxi Key Laboratory of Image and Graphic Intelligent Processing, Guilin University of Electronic Technology, Guilin 541004, China. E-mail: sguan691@163.com; Liji@guet.edu.cn; rslan@guet.edu.cn; luoxn@guet.edu.cn.
• Ji Li is also with Guilin Huigu Institute of AI Industrial Technology, Guilin 541004, China.
• Xinhua Lu is with School of Information Engineering, Nanyang Institute of Technology, Nanyang 473000, China. E-mail: ieluxinhua@sina.com.
* To whom correspondence should be addressed.
Manuscript received: 2023-03-16 ; revised: 2023-09-21; accepted: 2023-10-06

and variants of the CSSM such as the two-sided correlation transformation (TCT) algorithm[9], the weighted average of signal subspaces (WAVES) method[10], the signal subspace transformation (SST)[11] algorithm, and the test of orthogonality of projected subspace (TOPS)[12] method. These methods have been widely used in scenarios with multiple measurements to estimate the covariance matrix of subband signals. The CSSM outperforms the ISSM due to signal coherence in different subbands. These subspace-based methods and the prominent nonlinear least square methods based on maximum likelihood estimation have limitations in harsh scenarios, for example, when the number of sources is unknown, there are few snapshots, and the sources are highly correlated. These methods require the source number, which is difficult to know in advance, and are sensitive to source correlations. Furthermore, these methods require sufficient observations to yield accurate estimations, which can be difficult to obtain in certain scenarios.

Recent advances in compressed sensing (CS) theory have introduced sparse representation as a promising technology for broadband DOA estimation[13–15]. Sparse representation leverages a limited number of snapshots to achieve higher resolution and reduce aliasing effects[16–19]. Various methods based on sparse representation, including convex programming[20], greedy algorithms[21, 22], and sparse Bayesian learning[15, 23–25], have been proposed to improve the accuracy of broadband DOA estimation. However, these methods have their limitations, such as unreliable performance under low signal-to-noise ratio (SNR) for greedy algorithms and additional computational burden for parameter adjustment in convex optimization based methods. Compressed sensing techniques generally necessitate a densely sampled spatial grid to enhance the estimation performance of sparse signals. But the representation of the continuous domain via discrete domains leads to the problem of grid mismatch[26–28]. Moreover, the use of excessively dense partitions may lead to column-correlated overcomplete dictionaries that violates the condition for the sparse signal recovery. To settle these problems, many Bayesian algorithms for DOA estimation based on the off-grid model have been proposed[29–33]. However, these methods assume that different subband signals share the same bandwidth, which may not hold true in practice.

Many studies have proposed probabilistic models based on Dirichlet process (DP) prior to cluster different subbands in order to improve the accuracy of wideband DOA estimation[34–37]. Wang et al.[34] and Lu et al.[35] mainly inferred the hidden variables on mean field (MF) theory under the Bayesian networks. Lu et al.[36] proposed a combined belief propagation (BP) and MF message passing algorithm which can compute exact marginals to improve the channel estimation performance. Li et al.[37] proposed a novel DOA estimation method based on combined BP and MF message passing algorithm to address the off-grid problem of wideband DOA estimation. However, these methods suffer from high computational complexity. Real-valued transformations reduce DOA estimation algorithm complexity by approximately a factor of four by converting complex multiplications to real multiplications[38–42]. However, they do not address the exponential complexity issue caused by sparse coefficient matrix inversion during sparse signal recovery. To improve estimation accuracy and reduce computational complexity, a combined mean field and approximate message passing (MF-AMP) approach has been proposed[43]. Nevertheless, the AMP[44] algorithm may encounter difficulties with generic matrices, leading to divergence and poor performance[45]. To address this issue, variants of the AMP algorithm have been proposed, including damped AMP[45], memory AMP[46], and unitary approximate message passing (UAMP)[47]. However, these methods are unable to effectively capture the correlation between frequency points and angles of each subband, leading to limited estimation accuracy. As a result, it is necessary to develop more efficient methods that can reduce the computational burden while ensuring high estimation accuracy for wideband DOA estimation.

In this paper, we address the challenges of representing the Bayesian model for wideband DOA estimation by introducing the factor graph approach, which offers a more effective representation compared to the Bayesian network model[48, 49]. We recognize that the complexity arises from the direct relationship between the received signals and the DP prior in the probability model. To overcome this problem, we propose a novel factor graph that solves this issue by utilizing the factor graph transformation method within the probability graph model[50–52]. Additionally, we present a new and efficient method for wideband DOA estimation based on variational Bayesian inference (VBI) on the proposed factor graph. Our contributions

can be summarized as follows.

- **Factor graph representation of the off-grid model:** We have devised a factor graph to represent the off-grid model, incorporating factor and variable nodes. This provides a more intuitive depiction of the model's details.
- **A stretched factor graph with two additional hard constraints:** We have stretched the factor graph to include two new virtual variable nodes, each corresponding to a stringent constraint condition. One node facilitates the separation of the DP prior from the received signals, allowing the UAMP messages to propagate across the factor graph, while the other node addresses the grid mismatch issue, thereby enhancing estimation accuracy.
- **A novel broadband DOA estimation algorithm based on UAMP and MF:** We have integrated the MF rule and the UAMP algorithm to update the variables on the factor graph, leading to a novel inference algorithm that we refer to as UAMP-DP. Compared to variational Bayesian expectation maximization (VBEM) and Bayesian probabilistic matrix factorization (BPMF)[34, 36], the proposed approach significantly reduces complexity while maintaining or even surpassing their performance.

The remainder of this paper is organized as follows. Section 2 presents the system model for off-grid DOA estimation after unitary transformation, and represents it in the sparse Bayesian framework with the DP prior. Section 3 addresses the challenges of high complexity and off-grid problems by introducing two strong constraints through stretching the factor graph. We also derive efficient inferences for the parameters and hidden variables using variational Bayesian inference. Section 4 compares the proposed methods with previous approaches to demonstrate the performance improvements. Conclusions are drawn in Section 5.

2 Problem Formulation

Notation: Sets, matrices, and vectors are represented by flower body, boldface uppercase, and lowercase letters. Transpose and conjugate transposition of a matrix are denoted as $(\cdot)^{\mathrm{T}}$ and $(\cdot)^{\mathrm{H}}$. diag$(\cdot)$ denotes a diagonal matrix with its vector argument on the maindiagonal, and Tr$(\cdot)$ denotes the trace operator. The expectation operation with respect to a function $q(x)$ is represented by $\langle\cdot\rangle_{q(x)}$. Define $n_{x_i\to\alpha}(x_i)$ as the message from variable node x_i to factor node α, $m_{\alpha\to x_i}(x_i)$ as the message from factor node α to variable node x_i, and $b(x_i)$ as the belief of x_i node. The updated estimate is denoted by $\hat{(\cdot)}$.

2.1 System model

We consider U stationary sources impinging on an M-element uniform linear array (ULA) with an element spacing of d. The DFT of the received signals can be decomposed into F narrowband measurements $\boldsymbol{y}_f \in \mathbb{C}^{M\times 1}$ as

$$\boldsymbol{y}_f = \boldsymbol{A}_f\boldsymbol{x}_f + \boldsymbol{w}_f,\ f = 1,2,\ldots,F \tag{1}$$

where $\boldsymbol{x}_f \in \mathbb{C}^{N\times 1}$ is the sparse coefficient vector of length N at frequency f. $\boldsymbol{w}_f \in \mathbb{C}^{M\times 1}$ is complex Gaussian noise. $\boldsymbol{A}_f = [\boldsymbol{a}_f^{\mathrm{T}}(\theta_1), \boldsymbol{a}_f^{\mathrm{T}}(\theta_2), \ldots, \boldsymbol{a}_f^{\mathrm{T}}(\theta_N)]$, $\boldsymbol{A}_f \in \mathbb{C}^{M\times N}$, $N \gg M$, is an overcomplete dictionary. To mitigate the DOA estimation error, a straightforward approach is to use more dense grids to cover the detection space. This method results in a significant increase in computational complexity and coherence between adjacent atoms in the overcomplete dictionary, thereby violating the restricted isometry property (RIP) and leading to direction mismatch. Suppose the true DOA set is $\{\bar{\theta}_u\}_{u=1:U}$ with $U \ll N$. Without loss of generality, for some $\bar{\theta}_u \notin \{\theta_1\}_{n=1:N}$, $u \in [1,U]$, θ_{n_u} is the grid point closest to $\bar{\theta}_u$, where $n_u \in [1,N]$. So we apply a linear approximation to derive the off-grid model. The linear approximate of the true steering vector can be obtained by the first-order Taylor expansion method[21] as

$$\boldsymbol{A}_f(\bar{\theta}_u) \approx \boldsymbol{A}_f(\theta_{n_u}) + \boldsymbol{B}_f(\theta_{n_u})\beta_{f,n} \tag{2}$$

where the off-grid deviation is

$$\beta_{f,n} = \begin{cases} \bar{\theta}_u - \theta_{n_u}, & \text{for } n = n_u; \\ 0, & \text{others} \end{cases} \tag{3}$$

where $\boldsymbol{A}_f(\theta_{n_u})$ is the θ_{n_u} column of $\boldsymbol{A}_f$, and $\boldsymbol{B}_f(\theta_{n_u})$ represents the first-order derivative of $\boldsymbol{A}_f(\theta_{n_u})$ with respect to ϑ_{n_u}. $\boldsymbol{\beta}_f = [\beta_{f,1}, \beta_{f,2}, \ldots, \beta_{f,N}]^{\mathrm{T}}$ and $\boldsymbol{\Delta}_f = \mathrm{diag}(\boldsymbol{\beta}_f)$ are the deviation vector and the deviation matrix, respectively. $\boldsymbol{\Delta}_f$ follows the uniform distribution $\prod_{n=1}^{N} U\left(\left[-\frac{\Delta\theta}{2}, \frac{\Delta\theta}{2}\right]\right)$, where $\Delta\theta$ is the grid interval. The true dictionary can be given as $\boldsymbol{D}_f = \boldsymbol{A}_f + \boldsymbol{B}_f\boldsymbol{\Delta}_f$. The sparse model for each frequency point in the off-grid problem can be expressed as

$$\boldsymbol{y}_f = \boldsymbol{D}_f\boldsymbol{x}_f + \boldsymbol{w}_f,\ f = 1,2,\ldots,F \tag{4}$$

According to Ref. [47], as any $\boldsymbol{D}_f$ has a singular value decomposition (SVD) $\boldsymbol{D}_f = \boldsymbol{U}_f\boldsymbol{\Lambda}_f\boldsymbol{V}_f$ to enable the use of UAMP, the following unitary model can be

expressed as

$$\boldsymbol{r}_f = \boldsymbol{\Phi}_f \boldsymbol{x}_f + \boldsymbol{n}_f \tag{5}$$

where $\boldsymbol{r}_f = \boldsymbol{U}_f^{\mathrm{H}} \boldsymbol{y}_f$, $\boldsymbol{\Phi}_f = \boldsymbol{U}_f^{\mathrm{H}} \boldsymbol{D}_f$, and $\boldsymbol{n}_f = \boldsymbol{U}_f^{\mathrm{H}} \boldsymbol{w}_f$.

2.2 Sparse Bayesian learning framework

2.2.1 Noise model

Suppose that $\boldsymbol{n}_f$ is a complex Gaussian noise with precision α_0. Then the likelihood function of the signal model Eq. (5) can be expressed as

$$p(\boldsymbol{r}_f|\boldsymbol{x}_f, \boldsymbol{\Phi}_f, \alpha_0) = \mathrm{CN}(\boldsymbol{r}_f; \boldsymbol{\Phi}_f \boldsymbol{x}_f, \alpha_0^{-1} \boldsymbol{I}_M) \triangleq f_{\boldsymbol{r}_f}(\boldsymbol{\Phi}_f, \boldsymbol{x}_f, \alpha_0) \tag{6}$$

α_0 follows a Gamma prior as

$$p(\alpha_0|c,d) = \Gamma(\alpha_0; a, b) \triangleq f_{\alpha_0}(\alpha_0) \tag{7}$$

where the parameters a and b are selected to be small values to yield non-informative priors.

2.2.2 Sparse signal model

The $\boldsymbol{x}_f$ follows a complex Gaussian distribution as

$$p(\boldsymbol{x}_f|\boldsymbol{\alpha}_f) = \mathrm{CN}(\boldsymbol{x}_f; \boldsymbol{0}, \boldsymbol{\Lambda}_f^{-1}) \tag{8}$$

where $\boldsymbol{\alpha}_f^{\mathrm{T}} = [\alpha_{f,1}, \alpha_{f,2}, \ldots, \alpha_{f,N}]$ and $\boldsymbol{\Lambda}_f = \mathrm{diag}(\boldsymbol{\alpha}_f)$ are the precision vector and the precision matrix, respectively. $\boldsymbol{\alpha}_f$ follows a Gamma distribution with the parameters c and d as

$$p(\boldsymbol{\alpha}_f|c,d) = \prod_{n=1}^{N} \Gamma(\alpha_n|c,d) \triangleq f_{\boldsymbol{\alpha}_f}(\boldsymbol{\alpha}_f) \tag{9}$$

All $\boldsymbol{x}_f$ for $f = 1, 2, \ldots, F$ are independent and share the same $\boldsymbol{\Lambda}_k = \mathrm{diag}(\boldsymbol{\alpha}_k)$ to induce a joint sparsity in sparse Bayesian learning (SBL)[34].

2.2.3 DP prior on $\boldsymbol{\alpha}_f$

DP with excellent classification performance is used as the α_f prior. Assuming $\{\alpha_f\}_{1:F}$ are F random samples drawn from G, where G is a random measure drawn from $\mathrm{DP}(\gamma, G_0)$ with concentration parameter γ and base distribution G_0. We have the following expressions as

$$\boldsymbol{\alpha}_f \sim G, f = 1, 2, \ldots, F, G \sim \mathrm{DP}(\gamma, G_0) \tag{10}$$

As it is not possible to obtain G directly, a stick-breaking representation is used to express G as

$$G = \sum_{k=1}^{+\infty} w_k \delta_{\boldsymbol{\alpha}_k^*} \tag{11}$$

with

$$\alpha_k^* \sim G_0, w_k = \pi_k \prod_{i=1}^{k-1} (1 - \pi_i) \tag{12}$$

and

$$p(\boldsymbol{\pi}|\gamma) = \prod_{k=1}^{K} \mathrm{Beta}(\pi_k; 1, \gamma) \triangleq f_\pi(\boldsymbol{\pi}, \gamma) \tag{13}$$

where $\{\alpha_k^*\}_{1:+\infty}$ are atoms of G, and $\{w_k\}_{1:+\infty}$ are the weights of each atom which are attained by the stick-breaking method. In practical applications, the number of atoms is usually set to a larger number K. $\boldsymbol{\pi} = [\pi_1, \pi_2, \ldots, \pi_K]$ follows a Beta distribution with parameter γ. The parameter γ governs the number of clusters in the DP. γ is assumed to follow a Gamma distribution with the parameters e and f, which can be expressed as

$$p(\gamma|e,h) = \Gamma(\gamma; e, h) \triangleq f_\gamma(\gamma) \tag{14}$$

$\boldsymbol{z}_f$ is introduced as the assignment vector sampled from a multinomial distribution

$$p(\boldsymbol{z}_f|\{w_k\}_{k=1:K}) = \mathrm{Mult}(\{w_k\}_{k=1:K}) \tag{15}$$

Using Eq. (12), the conditional distribution of $\boldsymbol{z}_f$ can be expressed as

$$p(\boldsymbol{z}_f|\boldsymbol{\pi}) = \prod_{k=1}^{K} \left(\pi_k \prod_{i=1}^{k-1} (1 - \pi_i) \right)^{1[z_f = k]} \triangleq f_{\boldsymbol{z}_f}(\boldsymbol{z}_f, \boldsymbol{\pi}) \tag{16}$$

where $\Phi_{f,k} \triangleq 1[z_f = k]$ is defined as a structure where all its elements are zero expect for the k-th element, which is used to allocate $\boldsymbol{x}_f$ to the cluster k.

So the probabilistic model of $\boldsymbol{x}_f$ can be given as

$$p(\boldsymbol{x}_f|\boldsymbol{z}_f, \{\boldsymbol{\alpha}_k^*\}_{k=1:K}) = \prod_{k=1}^{K} \{\mathrm{CN}(\boldsymbol{x}_f|\boldsymbol{0}, \boldsymbol{\Lambda}_{z_f}^{-1})\}^{1\{z_f = k\}} \triangleq f_{\boldsymbol{x}_f}(\boldsymbol{x}_f, \boldsymbol{z}_f, \{\boldsymbol{\alpha}_k^*\}_{k=1:K}) \tag{17}$$

where $\boldsymbol{\Lambda}_{z_f} = \mathrm{diag}(\boldsymbol{\alpha}_{z_f}^*)$.

2.2.4 Uniform prior on $\boldsymbol{\beta}_f$

$\boldsymbol{\beta}_f$ follows a uniform prior which is only used to limit the boundness of $\boldsymbol{\beta}_f$ as

$$\boldsymbol{\beta}_f \sim \prod_{n=1}^{N} U\left(\left[-\frac{\Delta\theta}{2}, \frac{\Delta\theta}{2} \right] \right) \tag{18}$$

Because $\{\boldsymbol{\beta}_f\}_{f=1:F}$ can be classified into K clusters, the method of substituting $\{\boldsymbol{\beta}_k\}_{k=1:K}$ for $\{\boldsymbol{\beta}_f\}_{f=1:F}$ can make full use of the relevant information of the measurements in the same cluster k and eliminate the interference of measurements in other clusters.

So joint probability density function (PDF) of the

signal model can be expressed as

$$p(\{\boldsymbol{r}_f\}_{f=1:F},\boldsymbol{\Theta}) = \prod_{f=1}^{F} p(\boldsymbol{r}_f|\boldsymbol{x}_f,z_f,\{\boldsymbol{\beta}_k\}_{k=1:K},\alpha_0)\times p(\boldsymbol{x}_f|z_f,\{\boldsymbol{\alpha}_k^*\}_{k=1:K})p(z_f|\boldsymbol{\pi})p(\boldsymbol{\pi}|\gamma)p(\gamma)p(\alpha_0) \tag{19}$$

where $\Theta = \{\alpha_0,\gamma,\boldsymbol{\pi},\{z_f,\boldsymbol{x}_f\}_{f=1:F},\{\boldsymbol{\alpha}_k^*,\boldsymbol{\beta}_k\}_{k=1:K}\}$.

3 Proposed UAMP-DP Method

In this part, combined message passing algorithm is applied to calculate the approximate posterior on the factor graph, update the messages between the factor nodes and the variable nodes, and finally get the UAMP-DP method.

3.1 Factor graph model

Most variables in this model obey exponential distribution, and the expected function of these variables can be easily obtained by MF rules. But UAMP method has higher estimation accuracy than MF rule in VBI. Therefore, in order to compensate for the error of MF rules and grid division, we introduce two auxiliary variables, $\boldsymbol{h}_f \triangleq \boldsymbol{\Phi}_f\boldsymbol{x}_f$ and $\bar{\boldsymbol{A}}_f \triangleq \sum_{k=1}^{K}\Phi_{f,k}(\boldsymbol{A}_f+\boldsymbol{B}_f\Delta_k)$, which can be respectively expressed as

$$p(\bar{\boldsymbol{A}}_f|z_f,\{\beta_k\}_{k=1:K}) \triangleq f_{\bar{\boldsymbol{A}}_f}(\bar{\boldsymbol{A}}_f,\{\beta_k\}_{k=1:K},z_f) = \delta(\bar{\boldsymbol{A}}_f - \sum_{k=1}^{K}\Phi_{f,k}(\boldsymbol{A}_f+\boldsymbol{B}_f\Delta_k)) \tag{20}$$

and

$$p(\boldsymbol{h}_f|\boldsymbol{\Phi}_f,\boldsymbol{x}_f) = \delta(\boldsymbol{h}_f-\boldsymbol{\Phi}_f\boldsymbol{x}_f) \triangleq f_{\boldsymbol{h}_f}(\boldsymbol{h}_f,\boldsymbol{x}_f,\boldsymbol{\Phi}_f) \tag{21}$$

where $f_{\boldsymbol{r}_f}(\boldsymbol{h}_f,\alpha_0) = \mathrm{CN}(\boldsymbol{r}_f;\boldsymbol{h}_f,\alpha_0^{-1}\boldsymbol{I})$, the joint PDF in Eq. (19) can be rewritten as

$$p(\{\boldsymbol{r}_f,\boldsymbol{h}_f,\boldsymbol{\Phi}_f\}_{f=1:F},\boldsymbol{\Theta}) = \prod_{f=1}^{F} f_{\boldsymbol{r}_f}(\boldsymbol{h}_f,\alpha_0)f_{\boldsymbol{h}_f}(\boldsymbol{h}_f,\boldsymbol{x}_f,\boldsymbol{\Phi}_f)\times f_{\boldsymbol{\Phi}_f}(\bar{\boldsymbol{A}}_f,\{\beta_k\}_{k=1:K},z_f)f_{\boldsymbol{x}_f}(\boldsymbol{x}_f,z_f,\{\boldsymbol{\alpha}_k^*\}_{k=1:K})\times f_{z_f}(z_f,\boldsymbol{\pi})f_{\boldsymbol{\pi}}(\boldsymbol{\pi},\gamma)f_{\gamma}(\gamma)f_{\alpha_0}(\alpha_0) \tag{22}$$

As shown in Fig. 1, the variables $\{z_f\}_{f=1:F}$ And the observation variables $\{\boldsymbol{r}_f\}_{f=1:F}$ are separated by the auxiliary variables $\{\boldsymbol{h}_f\}_{f=1:F}$, which enables both UAMP and MF algorithms to be applied simultaneously on the factor graph to obtain an effective message passing algorithm. And the auxiliary variables $\bar{\boldsymbol{A}}_f$ can well approximate the real dictionary.

The factor graph Fig. 1 of the probability model proposed in this paper can be divided into four function blocks, labelled by Block(a)–Block(d), which are represented as the estimation block of DP prior parameter, DP mixture components, sparse signals and noise precision, and off-grid deviation, respectively.

3.2 Messages calculation of DP prior block

According to Eq. (17), we can get the MF message from $f_{\boldsymbol{x}_f}$ to z_f as

$$m_{f_{\boldsymbol{x}_f}}^{\mathrm{MF}} \to z_f\,(z_f) = \exp\left\{\langle \ln f_{\boldsymbol{x}_f}\rangle_{b(\boldsymbol{x}_f)\prod_{k=1}^{k} b(\boldsymbol{\alpha}_k^*)}\right\} = \exp\{\Phi_{f,k}\boldsymbol{g}_{f,k}\} \tag{23}$$

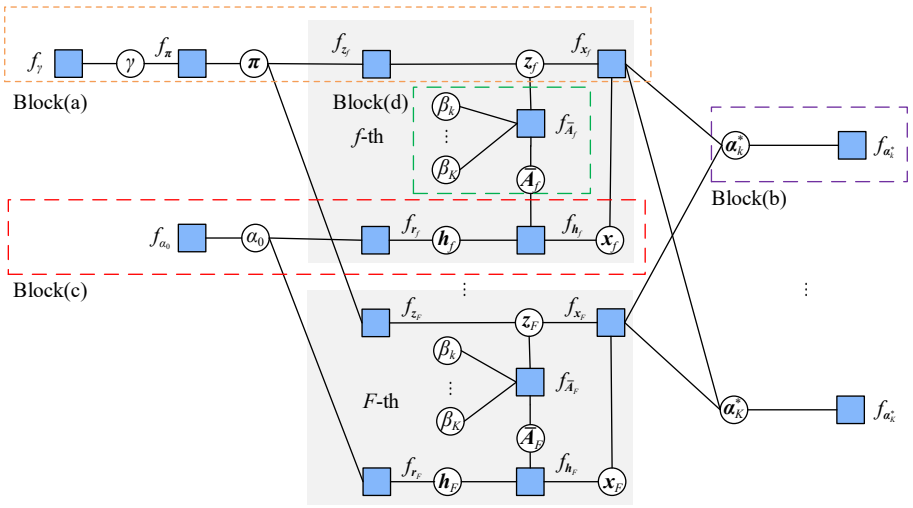


Fig. 1 Factor graph representation of joint probability $p(\{\boldsymbol{r}_f,\boldsymbol{h}_f,\boldsymbol{\Phi}_f\}_{f=1:F},\boldsymbol{\Theta})$.

where

$$\boldsymbol{g}_{f,k} \triangleq \langle \ln(\boldsymbol{\Lambda}_k)\rangle_{b(\boldsymbol{\alpha}_k^*)} - \mathrm{tr}(\mathrm{diag}(\boldsymbol{v}_{\boldsymbol{x}_f})\langle \boldsymbol{\Lambda}_k\rangle_{b(\boldsymbol{\alpha}_k^*)}) - \boldsymbol{x}_f^{\mathrm{H}}\langle \boldsymbol{\Lambda}_k\rangle_{b(\boldsymbol{\alpha}_k^*)}\boldsymbol{x}_f.$$

$\langle \ln(\boldsymbol{\Lambda}_k)\rangle_{b(\boldsymbol{\alpha}_k^*)}$ and $\langle \boldsymbol{\Lambda}_k\rangle_{b(\boldsymbol{\alpha}_k^*)}$ will be given later.

Substituting Eq. (16) and $b(\boldsymbol{\pi})$ into the message from $f_{z_f}(\boldsymbol{z}_f,\boldsymbol{\pi})$ to $\boldsymbol{z}_f$ yields

$$\ln m^{\mathrm{MF}}_{f_{z_f}\to \boldsymbol{z}_f}(\boldsymbol{z}_f) = \left\{\ln\langle f_{z_f}\rangle_{b(\boldsymbol{\pi})}\right\} = \left\{1[\boldsymbol{z}_f = k]\left[\langle \ln\pi_k\rangle_{b(\boldsymbol{\pi})} + \sum_{i=1}^{k-1}\langle \ln(1-\boldsymbol{\pi}_i)\rangle_{b(\boldsymbol{\pi})}\right]\right\} \tag{24}$$

Using Eqs. (23), (24), and (31), the belief of $\boldsymbol{z}_f$ which is the approximation PMF of $\boldsymbol{\Phi}_{f,k}$ can be obtained as

$$b(\boldsymbol{z}_f) = m^{\mathrm{MF}}_{f_{x_f}\to \boldsymbol{z}_f}(\boldsymbol{z}_f) m^{\mathrm{MF}}_{f_{z_f}\to \boldsymbol{z}_f}(\boldsymbol{z}_f) = \exp\{\boldsymbol{\Phi}_{f,k}\varepsilon_{f,k}\} \tag{25}$$

where

$$\varepsilon_{f,k} = \langle \ln\pi_k\rangle_{b(\pi)} + \sum_{i=1}^{k-1}\langle \ln(1-\pi_i)\rangle_{b(\pi)} + \boldsymbol{g}_{f,k} \tag{26}$$

After normalization, $\boldsymbol{\Phi}_{f,k}$ can be updated by

$$\hat{\boldsymbol{\Phi}}_{f,k} = \frac{\exp\{\varepsilon_{f,k}\}}{\sum_{k=1}^{K}\exp\left\{\varepsilon_{f,k}\right\}} \tag{27}$$

According to Eq. (27), MF message from f_{z_f} to $\boldsymbol{\pi}$ can be calculated by

$$m^{\mathrm{MF}}_{f_{z_f}\to\boldsymbol{\pi}}(\boldsymbol{\pi}) = \exp\left\{\langle \ln f_{z_f}(\boldsymbol{z}_f)\rangle_{b(z_f)}\right\} = \exp\left\{\hat{\boldsymbol{\Phi}}_{f,k}\ln\pi_k + \sum_{i=k+1}^{K}\hat{\boldsymbol{\Phi}}_{f,i}\ln(1-\pi_k)\right\} \tag{28}$$

Using Eqs. (13) and (34), message $m^{\mathrm{MF}}_{f_\pi\to\boldsymbol{\pi}}(\boldsymbol{\pi})$ can be obtained as

$$m^{\mathrm{MF}}_{f_\pi\to\boldsymbol{\pi}}(\boldsymbol{\pi}) = \exp\left\{\langle \ln f_{\boldsymbol{\pi}}\rangle_{b(\boldsymbol{\gamma})}\right\} = \exp\left\{(\hat{\gamma}-1)\sum_{k=1}^{K}\ln(1-\pi_k)\right\} \tag{29}$$

So the belief of $\boldsymbol{\pi}$ can be computed by

$$b(\boldsymbol{\pi}) = m^{\mathrm{MF}}_{f_\pi\to\boldsymbol{\pi}}(\boldsymbol{\pi})\prod_{f=1}^{F} m^{\mathrm{MF}}_{f_{z_f}\to\boldsymbol{\pi}}(\boldsymbol{\pi}) = \exp\left\{\left(\sum_{f=1}^{F}\sum_{l=k+1}^{K}\hat{\boldsymbol{\Phi}}_{f,l}+\hat{\gamma}-1\right)\ln(1-\pi_k)+\sum_{f=1}^{F}\hat{\boldsymbol{\Phi}}_{f,k}\ln\pi_k\right\} \tag{30}$$

Equation (30) shows that π_k follows a Beta distribution with two parameters $\tau_{1,k} \triangleq \sum_{f=1}^{F}\hat{\boldsymbol{\Phi}}_{f,k}+1$ and $\tau_{2,k} \triangleq \sum_{f=1}^{F}\sum_{l=k+1}^{K}\hat{\boldsymbol{\Phi}}_{f,l}+\hat{\gamma}$. Then the expectations of $\ln\pi_k$ and $\ln(1-\pi_k)$ can be computed by

$$\begin{aligned}\langle \ln\pi_k\rangle_{b(\pi)} &= \Psi(\tau_{1,k}) - \Psi(\tau_{1,k}+\tau_{2,k}),\\ \langle \ln(1-\pi_k)\rangle_{b(\pi)} &= \Psi(\tau_{2,k}) - \Psi(\tau_{1,k}+\tau_{2,k})\end{aligned} \tag{31}$$

where $\Psi(\cdot)$ denotes the diagmma function, then message from f_γ to γ can be given by

$$m^{\mathrm{MF}}_{f_\pi\to\gamma}(\gamma) = \exp\{\langle \ln f_{\boldsymbol{\pi}}\rangle_{b(\boldsymbol{\pi})}\} = \gamma^K\exp\left\{(\gamma-1)\sum_{k=1}^{K}\langle \ln(1-\pi_k)\rangle_{b(\boldsymbol{\pi})}\right\} \tag{32}$$

With the prior $f_\gamma(\gamma)$, the belief of γ is obtained as

$$b(\gamma) = m^{\mathrm{MF}}_{f_\pi\to\gamma}(\gamma) f_\gamma(\gamma) \propto \gamma^{e+K-1}\exp\left\{-\left(h - \sum_{k=1}^{K}\langle \ln(1-\pi_k)\rangle_{b(\boldsymbol{\pi})}\right)\gamma\right\} \tag{33}$$

and the update rule $\hat{\gamma}$ of γ is

$$\langle\gamma\rangle_{b(\gamma)} = \frac{\hat{e}}{\hat{h}} \triangleq \hat{\gamma} \tag{34}$$

where $\hat{e} = e + K - 1$ and $\hat{h} = h - \sum_{k=1}^{K}\langle \ln(1-\pi_k)\rangle$.

3.3 Messages calculation of DP mixture components block

With the beliefs of $b(\boldsymbol{z}_f)$ and $b(\boldsymbol{x}_f)$, the message from $\boldsymbol{x}_f$ to $\boldsymbol{\alpha}_k^*$ reads

$$m^{\mathrm{MF}}_{f_{\boldsymbol{x}_f}\to\boldsymbol{\alpha}_k^*}(\boldsymbol{\alpha}_k^*) = \exp\left\{\left\langle \ln f_{\boldsymbol{x}_f}(\boldsymbol{x}_f,\boldsymbol{z}_f)\right\rangle_{b(\boldsymbol{z}_f)b(\boldsymbol{x}_f)}\right\} \propto \exp\left\{\hat{\boldsymbol{\Phi}}_{f,k}\sum_{i=1}^{N}\left[\ln\alpha_{k,i}^* - \alpha_{k,i}^*(\hat{x}_{f,i}^2+\delta_{f,i}^2)\right]\right\} \tag{35}$$

With the prior of $f_{\boldsymbol{\alpha}_k^*}(\boldsymbol{\alpha}_k^*) = \prod_{n=1}^{N}\Gamma(\alpha_{k,i}^*|c,d)$, the belief of $\boldsymbol{\alpha}_k^*$ is given as

$$b(\boldsymbol{\alpha}_k^*) = m^{\mathrm{MF}}_{f_{\boldsymbol{\alpha}_k^*}\to\boldsymbol{\alpha}_k^*}(\boldsymbol{\alpha}_k^*)\prod_{f=1}^{F} m^{\mathrm{MF}}_{f_{\boldsymbol{x}_k}\to\boldsymbol{\alpha}_k^*}(\boldsymbol{\alpha}_k^*) \propto \prod_{i=1}^{N}\boldsymbol{\alpha}_{k,i}^{*\,\hat{c}_{k,i}-1}\exp\left\{-\boldsymbol{\alpha}_{k,i}^*\hat{d}_{k,i}\right\} \tag{36}$$

where $\hat{d}_{k,i} = d + \sum_{f=1}^{F}\sum_{k=1}^{K}\hat{\boldsymbol{\Phi}}_{f,k}\left(\hat{x}_{f,i}^2+\delta_{f,i}^2\right)$ and $\hat{c}_{k,i} = c + \sum_{f=1}^{F}\hat{\boldsymbol{\Phi}}_{f,k}$, then the expectations of $\alpha_{k,i}^*$ and $\ln\alpha_{k,i}^*$ can be updated as

$$\langle\alpha_{k,i}^*\rangle_{b(\alpha_{k,i}^*)} = \frac{\hat{c}_{k,i}}{\hat{d}_{k,i}} \triangleq \hat{\alpha}_{k,i}^* \tag{37}$$

and

$$\langle \ln\alpha_{k,i}^*\rangle_{b(\alpha_{k,i}^*)} = \Psi(\hat{c}_{k,i}) - \ln(\hat{d}_{k,i}) \tag{38}$$

3.4 UAMP messages calculation of sparse signals and noise precision block

With the beliefs of $\boldsymbol{z}_f$, $\{\boldsymbol{\alpha}_k^*\}_{k=1:K}$ and the true dictionary $\bar{\boldsymbol{A}}_f$, the MF message from $f_{\boldsymbol{x}_f}$ to $\boldsymbol{x}_f$ can be computed by

$$m_{f_{\boldsymbol{x}_f}\to\boldsymbol{x}_f}^{\mathrm{MF}}(\boldsymbol{x}_f)=\exp\{\langle\ln f_{\boldsymbol{x}_f}\rangle_{b(z_f)\prod_{k=1}^{K}b(\boldsymbol{\alpha}_k^*)}\}\propto$$
$$\exp\left\{\sum_{k=1}^{k}\hat{\Phi}_{f,k}(-\boldsymbol{x}_f^{\mathrm{H}}\langle\boldsymbol{\Lambda}_k\rangle_{b(\boldsymbol{\alpha}_k^*)}\boldsymbol{x}_f+\langle\ln\boldsymbol{\Lambda}_k\rangle)_{b(\boldsymbol{\alpha}_k^*)}\right\} \tag{39}$$

According to the UAMP method[47], message $m_{f_{\boldsymbol{h}_f}\to\boldsymbol{x}_f}(\boldsymbol{x}_f)$ can be obtained as

$$\boldsymbol{q}(\boldsymbol{x}_f)=\mathrm{CN}(\boldsymbol{v}_{\boldsymbol{q}_f};\hat{\boldsymbol{q}}_f,\boldsymbol{v}_{\boldsymbol{q}_f}) \tag{40}$$

where

$$1./\boldsymbol{v}_{\boldsymbol{q}_f}=|\bar{\boldsymbol{A}}_f^{\mathrm{H}}|^{\cdot 2}\boldsymbol{v}_s \tag{41}$$

and

$$\hat{\boldsymbol{q}}_f=\boldsymbol{v}_{\boldsymbol{q}_f}(\bar{\boldsymbol{A}}_f^{\mathrm{H}}\hat{\boldsymbol{s}})+\boldsymbol{x}_f \tag{42}$$

where $\hat{\boldsymbol{s}}$ and $\boldsymbol{v}_s$ are defined as[47]

$$\boldsymbol{v}_s=1./(\boldsymbol{v}_{\boldsymbol{p}_f}+\boldsymbol{v}_{\hat{\boldsymbol{\epsilon}}_f}) \tag{43}$$

and

$$\hat{\boldsymbol{s}}=\boldsymbol{v}_s(\hat{\boldsymbol{\epsilon}}_f-\hat{\boldsymbol{p}}_f) \tag{44}$$

$\hat{\boldsymbol{p}}_f$ and $\boldsymbol{v}_{\boldsymbol{p}_f}$ will be given later. The belief of $\boldsymbol{x}_f$ is updated as

$$b(\boldsymbol{x}_f)=m_{f_{\boldsymbol{x}_f}\to\boldsymbol{x}_f}^{\mathrm{MF}}(\boldsymbol{x}_f)\boldsymbol{q}(\boldsymbol{x}_f)\propto\mathrm{CN}(\boldsymbol{x}_f;\hat{\boldsymbol{x}}_f,\boldsymbol{v}_{\boldsymbol{x}_f}) \tag{45}$$

where

$$\boldsymbol{v}_{\boldsymbol{x}_f}=1./(\sum_{k=1}^{K}(\langle\boldsymbol{\alpha}_k^*\rangle_{b(\boldsymbol{\alpha}_k^*)})\mathrm{diag}(\hat{\boldsymbol{\Phi}}_f)+1./\boldsymbol{v}_{\boldsymbol{q}_f}) \tag{46}$$

and

$$\hat{\boldsymbol{x}}_f=\boldsymbol{v}_{\boldsymbol{x}_f}\odot(\hat{\boldsymbol{q}}_f./\boldsymbol{v}_{\boldsymbol{q}_f}) \tag{47}$$

Then message $m_{f_{\boldsymbol{h}_f}\to\boldsymbol{h}_f}=\mathrm{CN}(\boldsymbol{h}_f;\hat{\boldsymbol{p}}_f,\boldsymbol{v}_{\boldsymbol{p}_f})$ can be found in Ref. [44] as

$$\boldsymbol{v}_{\boldsymbol{p}_f}=|\bar{\boldsymbol{A}}_f|^{\cdot 2}\boldsymbol{v}_{\boldsymbol{x}_f} \tag{48}$$

and

$$\hat{\boldsymbol{p}}_f=\bar{\boldsymbol{A}}_f\hat{\boldsymbol{x}}_f-\hat{\boldsymbol{s}}\odot\boldsymbol{v}_{\boldsymbol{p}_f} \tag{49}$$

Then the message $m_{f_{\boldsymbol{r}_f}\to\hat{\boldsymbol{h}}_f}^{\mathrm{UAMP}}$ can be updated as

$$m_{f_{\boldsymbol{r}_f}\to\hat{\boldsymbol{h}}_f}^{\mathrm{UAMP}}=\exp\left\{\langle\ln f_{\boldsymbol{r}_f}\rangle_{b(\alpha_0)}\right\}\triangleq\mathrm{CN}(\boldsymbol{h}_f;\hat{\boldsymbol{\epsilon}}_f,\boldsymbol{v}_{\hat{\boldsymbol{\epsilon}}_f}) \tag{50}$$

where

$$\hat{\boldsymbol{\epsilon}}_f=\hat{\boldsymbol{r}}_f,\boldsymbol{v}_{\hat{\boldsymbol{\epsilon}}_f}=1/\hat{\alpha}_0 \tag{51}$$

The belief of $\boldsymbol{h}_f$ as

$$b(\boldsymbol{h}_f)=m_{f_{\boldsymbol{r}_f}\to\boldsymbol{h}_f}^{\mathrm{UAMP}}m_{f_{\boldsymbol{h}_f}\to\boldsymbol{h}_f}\propto\mathrm{CN}(\boldsymbol{h}_f;\hat{\boldsymbol{h}}_f,\boldsymbol{v}_{\boldsymbol{h}_f}) \tag{52}$$

where

$$\boldsymbol{v}_{\boldsymbol{h}_f}=1./(1./\boldsymbol{v}_p+\hat{\alpha}_0\boldsymbol{I}) \tag{53}$$

and

$$\hat{\boldsymbol{h}}_f=\boldsymbol{v}_{\boldsymbol{h}_f}\odot(\hat{\boldsymbol{p}}_f./\boldsymbol{v}_{\boldsymbol{p}_f}+\hat{\alpha}_0\boldsymbol{y}_f) \tag{54}$$

Noise precision $\hat{\alpha}_0$ can be obtained in Eq. (57), then message $m_{f_{\boldsymbol{r}_f}\to\alpha_0}$ can be computed as

$$m_{f_{\boldsymbol{r}_f}\to\alpha_0}=\exp\left\{\langle\ln f_{\boldsymbol{r}_f}(\boldsymbol{h}_f,\alpha_0)\rangle_{b(\boldsymbol{h}_f)}\right\}\propto$$
$$\alpha_0^M-\exp\left\{-\alpha_0\langle\|\boldsymbol{r}_f-\boldsymbol{h}_f\|^2\rangle_{b(\boldsymbol{h}_f)}\right\} \tag{55}$$

With the prior f_{α_0}, the belief of α_0 can be computed by

$$b(\alpha_0)=f_{\alpha_0}\prod_{f=1}^{F}m_{f_{\boldsymbol{y}_f}\to\alpha_0}\propto$$
$$\alpha^{a+\mathrm{MF}-1}\exp\left\{-(b+\sum_{f=1}^{F}\langle\|\boldsymbol{r}_f-\boldsymbol{h}_f\|^2\rangle_{b(\boldsymbol{h}_f)})\alpha_0\right\} \tag{56}$$

So the expectation of α_0 can be updated by

$$\hat{\alpha}_0=\frac{a+FM}{b+\sum_{f=1}^{F}\langle\|\boldsymbol{r}_f-\boldsymbol{h}_f\|^2\rangle_{b(\boldsymbol{h}_f)}} \tag{57}$$

3.5 Messages calculation of off-grid deviation block

According to Ref. [31], the off-grid deviation $\{\boldsymbol{\beta}\}_{k=1:K}$ are computed by

$$\{\hat{\boldsymbol{\beta}}\}_{k=1:K}=\max_{\{\hat{\boldsymbol{\beta}}\}_{k=1:K}}\langle\prod_{f=1}^{F}f_{\boldsymbol{r}_f}(\bar{\boldsymbol{A}}_f,\boldsymbol{x}_f,\alpha_0)\rangle_{b(\alpha_0)\prod_{f=1}^{F}b(z_f)b(\boldsymbol{x}_f)} \tag{58}$$

then $\hat{\boldsymbol{\beta}}_k$ can be updated by

$$\hat{\boldsymbol{\beta}}_k=\boldsymbol{\Gamma}_{\beta_k}^{-1}\sum_{f=1}^{F}\Phi_{f,k}\boldsymbol{v}_f^{\mathrm{T}} \tag{59}$$

where

$$\boldsymbol{\Gamma}_{\beta_k}=\hat{\alpha}_0\sum_{f=1}^{F}\hat{\Phi}_{f,k}[\overline{\boldsymbol{B}_f^{\mathrm{T}}\boldsymbol{B}_f}\odot(\hat{\boldsymbol{x}}_f\hat{\boldsymbol{x}}_f^{\mathrm{H}}+\mathrm{diag}(\boldsymbol{v}_{\boldsymbol{x}_f})] \tag{60}$$

$$\boldsymbol{v}_f=\mathrm{re}\{\mathrm{diag}(\hat{\boldsymbol{x}}_f)\boldsymbol{B}_f^{\mathrm{H}}(\boldsymbol{y}_f-\boldsymbol{A}_f\hat{\boldsymbol{x}}_f)-$$
$$\mathrm{diag}(\boldsymbol{B}_f^{\mathrm{H}}\boldsymbol{A}_f\mathrm{diag}(\boldsymbol{v}_{\boldsymbol{x}_f}))\} \tag{61}$$

The message $m^{\mathrm{MF}}_{f_{\boldsymbol{h}_f}\to\hat{\boldsymbol{A}}_f}$, which satisfies the deterministic constraints of $f_{\boldsymbol{h}_f}(\boldsymbol{h}_f,\boldsymbol{x}_f)$, is always equal to one. So the belief of $\bar{\boldsymbol{A}}_f$ can be computed by

$$b(\hat{\boldsymbol{A}}_f)=m^{\mathrm{MF}}_{f_{\hat{\boldsymbol{A}}_f}\to\bar{\boldsymbol{A}}_f}\,m^{\mathrm{MF}}_{f_{\boldsymbol{h}_f}\to\bar{\boldsymbol{A}}_f}=\exp\left\{\langle\ln f_{\bar{\boldsymbol{A}}_f}\rangle_{b(\mathrm{z}_f)\prod_{k=1}^{K}b(\boldsymbol{\beta}_k)}\right\}=\delta(\bar{\boldsymbol{A}}_f-\sum_{k=1}^{K}\hat{\boldsymbol{\Phi}}_{f,k}(\boldsymbol{A}_f+\boldsymbol{B}_f\hat{\boldsymbol{\Delta}}_k)) \tag{62}$$

Then the expectation of $\bar{\boldsymbol{A}}_f$ can be updated as

$$\hat{\bar{\boldsymbol{A}}}_f=\sum_{k=1}^{K}\hat{\boldsymbol{\Phi}}_{f,k}(\boldsymbol{A}_f+\boldsymbol{B}_f\hat{\boldsymbol{\Delta}}_k) \tag{63}$$

The proposed schedule of broadband off-grid DOA estimation is summarized in Algorithm 1. Variables and messages are updated iteratively until converging to the threshold Thr or reaching the maximum number of iterations $\mathrm{It}_{\max}$.

3.6 Computational complexity

According to Algorithm 1, the computational complexity of DP process is mainly dominated by Eqs. (26) and (37), $O(FKN)$ in Block(a) and Block(b). The main cost of Block(c), generated by Eqs. (42) and (47), is $O(FMN)$ in each iteration. In Block(d), the maximum cost is from the inverse of $\boldsymbol{\Gamma}_{\boldsymbol{\beta}_k}$. According to the component of variable $\{\boldsymbol{\alpha}^*_k\}_{k=1:K}$, we only calculate the $\boldsymbol{\beta}_k$ on the grids where users are located, so the computational complexity of $\boldsymbol{\Gamma}^{-1}_{\boldsymbol{\beta}_k}$ is $O(KFU^2)$. Thus the computational complexity of the proposed algorithm is $O(FMN)$, $N\gg\{K,U,F\}$.

It can be shown in Table 1 that the proposed algorithm has a lower complexity compared to other algorithms.

4 Simulation Result

In this section, we present the results of our simulations, which were conducted on a personal computer equipped with an Intel Core i5 @ 3 GHz processor and 8 GB LPDDR4 @ 3200 MHz, using MATLAB R2021a. Our experiments adhere to the settings described in Ref. [34]. The incident signals are formulated as the superposition of several harmonics:

$$s(t)=\sum_{u=1}^{U}\sum_{i=1}^{F_u}A_{u,i}\exp(-\mathrm{j}2\pi i f_u t) \tag{64}$$

Algorithm 1 Combined UAMP and MF message passing algorithm of broadband off-grid DOA estimation

Input: Received broadband signals $\{\boldsymbol{y}\}_{f=1:F}$;

Dictionary matrix $\{\boldsymbol{A}\}_{f=1:F}$;

Derivative matrix of dictionary $\{\boldsymbol{B}\}_{f=1:F}$

1: Set: Thr and $\mathrm{It}_{\max}$;

2: Initialize: a,b,c,d,e,h, $\boldsymbol{\Theta}$, $\hat{\boldsymbol{x}}_f$, $\boldsymbol{v}_{\boldsymbol{x}_f}$, and $\hat{\boldsymbol{s}}$

3: **while** $\|\hat{\boldsymbol{x}}^t_f-\hat{\boldsymbol{x}}^{t-1}\|^2_2\geqslant$ Thr or $t\leqslant\mathrm{It}_{\max}$ **do**

4: $\forall f$: update $\hat{\boldsymbol{p}}_f$ and $\boldsymbol{v}_{\boldsymbol{p}_f}$ by Eqs. (48) and (49);

5: $\forall f$: update $\hat{\boldsymbol{h}}_f$ and $\boldsymbol{v}_{\boldsymbol{h}_f}$ by Eqs. (53) and (54);

6: Update $\hat{\boldsymbol{s}}$ and $\hat{\boldsymbol{v}}_{\boldsymbol{s}}$ by Eqs. (43) and (44);

7: $\forall f$: update $\hat{\boldsymbol{q}}_f$ and $\boldsymbol{v}_{\boldsymbol{q}_f}$ by Eqs. (41) and (42);

8: $\forall f$: update $\hat{\boldsymbol{x}}_f$ and $\boldsymbol{v}_{\boldsymbol{x}_f}$ by Eqs. (46) and (47);

9: $\forall k,i$: update $\langle\alpha^*_{k,i}\rangle_{b(\alpha^*_{k,i})}$ and $\langle\ln\alpha^*_{k,i}\rangle_{b(\alpha^*_{k,i})}$ by Eqs. (37) and (38);

10: $\forall f,k$: update $\hat{\boldsymbol{\Phi}}_{f,k}$ by Eq. (27);

11: $\forall k$: update $\langle\ln\pi_k\rangle_{b(\boldsymbol{\pi})}$ and $\langle\ln(1-\pi_k)\rangle_{b(\boldsymbol{\pi})}$ by Eq. (31);

12: Update $\hat{\gamma}$ by Eq. (34);

13: Update $\hat{\alpha}_0$ by Eq. (57);

14: $\forall k$: update $\hat{\boldsymbol{\beta}}_k$ by Eq. (59);

15: $\forall f$: update $\hat{\bar{\boldsymbol{A}}}_f$ by Eq. (63);

16: $\forall f$: update $\hat{\boldsymbol{\Phi}}_f$ by Eq. (5);

17: $\forall f$: update $\hat{\boldsymbol{\epsilon}}_f$ and $\boldsymbol{v}_{\hat{\boldsymbol{\epsilon}}_f}$ by Eq. (51);

18: $t=t+1$;

19: **end while**

Table 1 Computational complexity of different algorithms.

Algorithm	Computational complexity
OMP	$O(UMN)$
OGSBI	$O(MN^2)$
VBEM	$O(N^3)$
BPMF	$O(N^3)$

Note: OMP: orthogonal matching pursuit; OGSBI: off-grid sparse Bayesian inference.

We consider an ULA serving U users, where $A_{u,i}=1$, for each user-sensor pair $(u,\ i)$. The frequencies of the subbands are $f_1=114$ Hz, $f_2=159$ Hz, and $f_3=197$ Hz, with $F_1=F_2=F_3=10$. The ULA operates at a design frequency of $f_{\mathrm{req}}=500$ Hz. The dictionary grid size is $N=91$, and the number of snapshots observed is $T=526$. We use $F=35$ subbands with high energy and $K=35$ classes for the DP prior. The performance of our proposed algorithm is then compared with various existing techniques, including beamforming[5], multiple signal classification (MUSIC)[7], OMP[17], OGSBI[31],

VBEM[34], and BPMF[37].

4.1 Joint space-frequency information reconstruction

To evaluate the performance of our proposed algorithm for incident signal reconstruction, we assume the incidence of three broadband signals $U = 3$ from the directions of $\bar{\theta}_1 = 121.99°$, $\bar{\theta}_2 = 89.75°$, and $\bar{\theta}_3 = 34.28°$ with SNR = 2.5 dB. Figure 2 displays the angular domain frequency information and spatial spectrum recovered by different algorithms, which contributes to more accurate DOA estimation and user discrimination.

Figure 2a presents the ground truth of the signal. Figures 2b and 2c depict the estimation results from classical beamforming and MUSIC, respectively, both of which suffer from serious spatial aliasing leading to extremely low angular resolution. Figures 2d and 2e show the results from OMP and OGSBI, displaying a relatively clear sparse structure. But there are still instances of false peaks that lead to discrepancies with the true DOA. Clearer frequency points are produced by VBEM and BPMF, as seen in Figs. 2f and 2g. Figure 2h demonstrates that our proposed method yields an extremely accurate result, comparing to VBEM and BPMF. The spatial spectrum of the

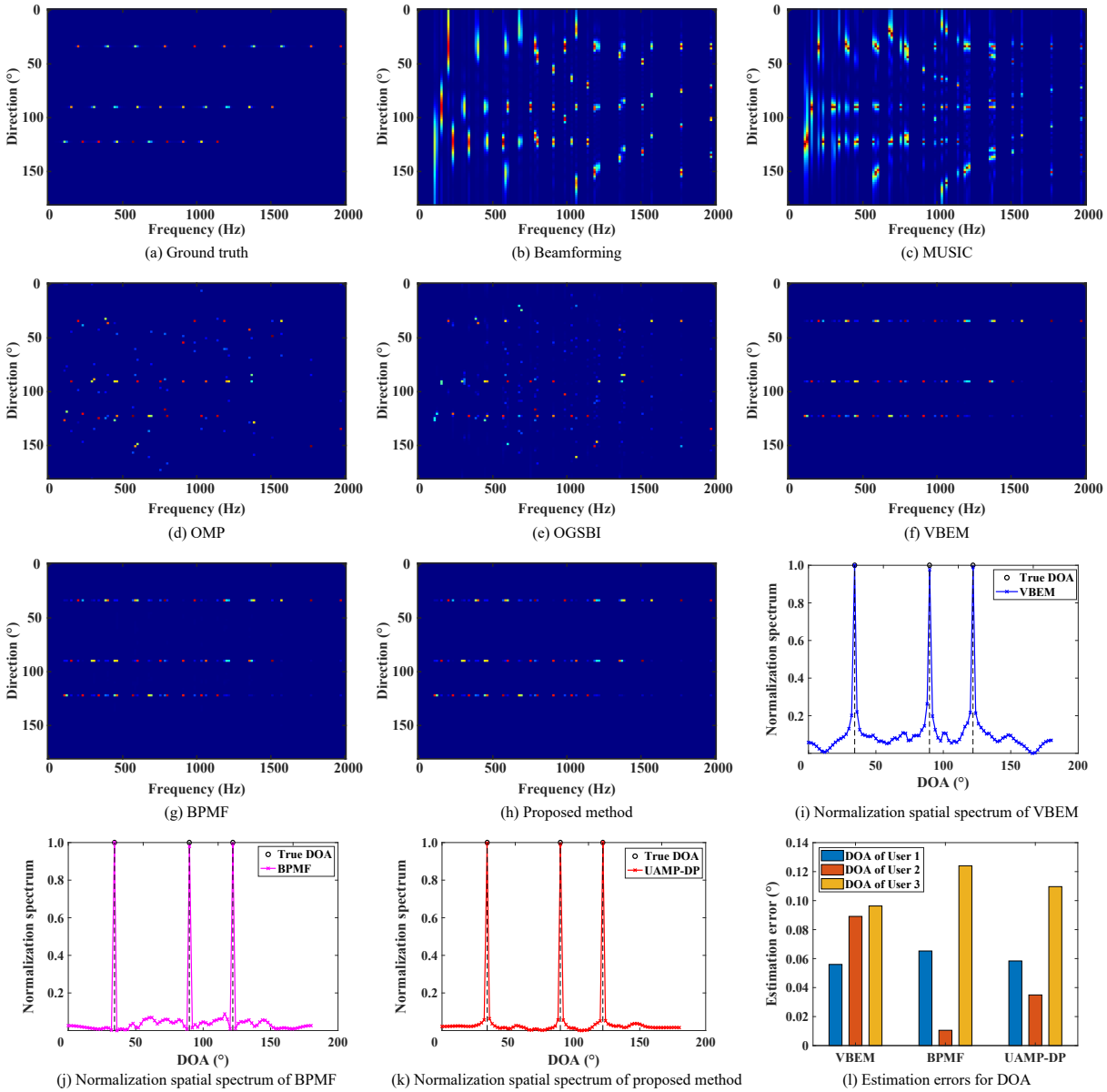


Fig. 2 Estimation results of different algorithms.

UAMP-DP method in Fig. 2k has more concentrated and smoother peaks than those of VBEM in Fig. 2i and BPMF in Fig. 2j. Figure 2l shows that our proposed method achieves lower DOA estimation errors than both VBEM and BPMF. The root mean square errors (RMSEs) of the UAMP-DP, VBEM, and BPMF methods are 0.0674°, 0.0758°, and 0.0716°, respectively.

4.2 RMSE performance versus number of subbands

This experiment employs 30 sets of random angles to explore the impact of subband selection on the accuracy of the DOA estimation under an SNR = 2.5 dB. Figure 3 shows the variations in the RMSE of DOA estimation across different algorithms with respect to the number of subbands. We observe that our proposed method provides the most accurate estimation results with a limited number of subbands. However, as the number of subbands increases, an excess of noise-only subbands comes into play, leading to a decline in performance and significant consumption of computing resources. For practical applications, it is beneficial to select an appropriate number of high-energy subbands that the signals occupy for DOA estimation, thereby enhancing the accuracy of the estimation while minimizing computational time. Therefore, this paper selects 35 high-energy subbands for DOA estimation.

4.3 RMSE performance versus grid interval

The grid interval of the dictionary directly impacts the accuracy of the DOA estimation, with finer grids yielding higher estimation accuracy. Nonetheless, the high dimensionality of the dictionary leads to an exponential increase in computational complexity.

To assess the performance of our proposed method, we replicate the scenario from Section 4.2 and conduct 30 Monte Carlo experiments under varying grid intervals with SNR = 2.5 dB. We then compare the RMSE of various algorithms. As illustrated in Fig. 4, we observe that (1) the estimation accuracy of all algorithms diminishes as the number of grids increases; (2) when the grid spacing ranges between 5°–7°, our proposed method aligns closely with the BPMF method. Yet, as the grid interval continues to expand, the RMSE of our proposed method proves smaller than that of other methods. Our proposed method continues to deliver robust DOA estimation performance, even with increased grid intervals.

4.4 RMSE performance versus computational time

In this instance, we compared the proposed method to the VBEM and BPMF in terms of RMSE versus running time to further illustrate its performance. We consider the same scenario as outlined in Section 4.1. As depicted in Fig. 5, we observed that: (1) the single iteration time of the VBEM and BPMF algorithms is similar but significantly longer than that of the UAMP-DP algorithm, due to the high-dimensional variance matrix inversion required in each iteration of the VBEM and BPMF methods; (2) the UAMP-DP algorithm converges to the optimal value in roughly 20 s, while the BP method and MF method converge to the optimal solution in about 38 s and 72 s, respectively, as shown in Fig. 5a; (3) once all the algorithms have fully converged, the RMSE of the proposed algorithm is smaller than those of the VBEM and BPMF methods, as indicated in Fig. 5b. These findings suggest that our proposed method boasts the advantages of low complexity and high estimation accuracy.

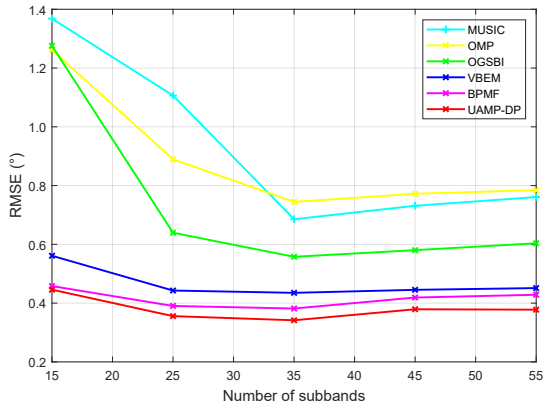


Fig. 3 RMSE versus number of subbands.

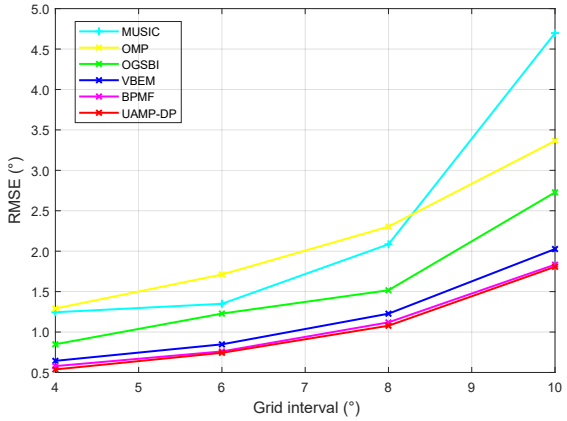


Fig. 4 RMSE versus grid interval.

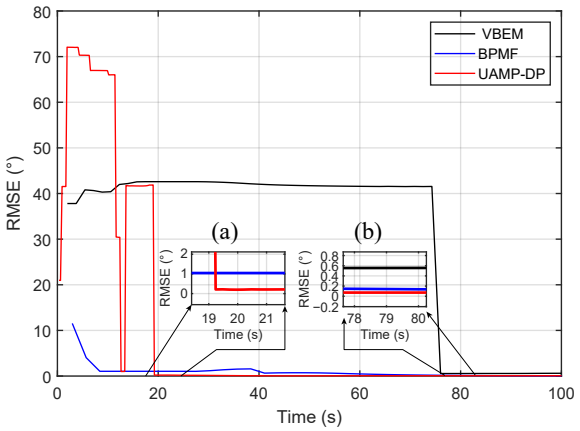


Fig. 5 RMSE versus computational time.

4.5 RMSE performance versus SNR

In this study, we examine the impact of SNR on the RMSE performance. This analysis involves the same scenario as discussed in Section 4.2 and includes 30 Monte Carlo simulations conducted under varying SNR conditions. The results, as displayed in Fig. 6, suggest that: (1) the VBI-based method nearly delivers the best DOA estimation performance; and (2) the UAMP-DP method is significantly closer to the Cramer-Rao lower bound (CRLB)[53] than the other methods under comparison.

5 Conclusion

In this study, we have undertaken an exploration of off-grid based algorithms for wideband DOA estimation using the DP prior. The existing sparse Bayesian model presents a high degree of coupling between observation and DP prior, which escalates the complexity of the algorithm.

To address this issue, we have introduced a novel factor graph by employing the stretching factor graph approach, incorporating two hard constraints into the design. This approach has enabled us to derive a unified UAMP-MF message passing algorithm to be executed on the factor graph. Simulation results validate that our proposed algorithm surpasses existing wideband DOA estimators employing DP prior, in terms of complexity. Remarkably, it achieves this while maintaining estimation accuracy on par with the state-of-the-art.

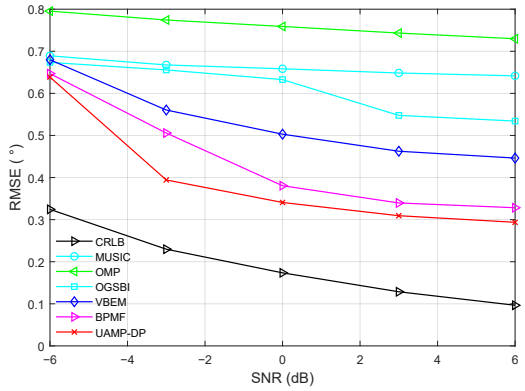


Fig. 6 RMSRs of incident signals with different SNRs.

Acknowledgment

This work was supported in part by the National Natural Science Foundation of China (Nos. 6202780103 and 62033001), the Innovation Key Project of Guangxi Province (No. AA22068059), the Key Research and Development Program of Guilin (No. 2020010332), the Natural Science Foundation of Henan Province (No. 222300420504), and Academic Degrees and Graduate Education Reform Project of Henan Province (No. 2021SJGLX262Y).

References

[1] Y. Hu and G. Leus, Robust differential received signal strength-based localization, *IEEE Trans. Signal Process.*, vol. 65, no. 12, pp. 3261–3276, 2017.

[2] B. Zhou, A. Liu, and V. Lau, Successive localization and beamforming in 5G mmWave MIMO communication systems, *IEEE Trans. Signal Process.*, vol. 67, no. 6, pp. 1620–1635, 2019.

[3] Z. Xiao and Y. Zeng, An overview on integrated localization and communication towards 6G, *Sci. China Inf. Sci.*, vol. 65, no. 3, pp. 1–46, 2021.

[4] Z. Q. He, Z. P. Shi, L. Huang, and H. C. So, Underdetermined DOA estimation for wideband signals using robust sparse covariance fitting, *IEEE Signal Process. Lett.*, vol. 22, no. 4, pp. 435–439, 2015.

[5] M. A. Doron, A. J. Weiss, and H. Messer, Maximum-likelihood direction finding of wide-band sources, *IEEE Trans. Signal Process.*, vol. 41, no. 1, p. 411, 1993.

[6] H. Wang and M. Kaveh, Coherent signal-subspace processing for the detection and estimation of angles of arrival of multiple wide-band sources, *IEEE Trans. Acoust., Speech, Signal Process*, vol. 33, no. 4, pp. 823–831, 1985.

[7] R. Schmidt, Multiple emitter location and signal parameter estimation, *IEEE Trans. Antennas Propagat.*, vol. 34, no. 3, pp. 276–280, 1986.

[8] Y. Bai, J. Li, Y. Wu, Q. Wang, and X. Zhang, Weighted incoherent signal subspace method for DOA estimation on wideband colored signals, *IEEE Access*, vol. 7, pp. 1224–1233, 2019.

[9] S. Valaee and P. Kabal, Wideband array processing using a two-sided correlation transformation, *IEEE Trans. Signal Process.*, vol. 43, no. 1, pp. 160–172, 1995.

[10] E. D. D. Claudio and R. Parisi, WAVES: Weighted average of signal subspaces for robust wideband direction finding, *IEEE Trans. Signal Process.*, vol. 49, no. 10, pp. 2179–2191, 2001.

[11] M. A. Doron and A. J. Weiss, On focusing matrices for wide-band array processing, *IEEE Trans. Signal Process.*, vol. 40, no. 6, pp. 1295–1302, 1992.

[12] J. Zhang, J. Dai, and Z. Ye, An extended TOPS algorithm based on incoherent signal subspace method, *Signal Process.*, vol. 90, no. 12, pp. 3317–3324, 2010.

[13] S. Ji, D. Dunson, and L. Carin, Multitask compressive sensing, *IEEE Trans. Signal Process.*, vol. 57, no. 1, pp. 92–106, 2009.

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

[15] K. Liu, X. Xi, Z. Xu, and S. Wang, A piecewise linear programming algorithm for sparse signal reconstruction, *Tsinghua Science and Technology*, vol. 22, no. 1, pp. 29–41, 2017.

[16] D. L. Donoho, Compressed sensing, *IEEE Trans. Inform. Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[17] S. Ganguly, I. Ghosh, R. Ranjan, J. Ghosh, P. K. Kumar, and M. Mukhopadhyay, Compressive sensing based off-grid DOA estimation using OMP algorithm, in *Proc. 2019 6th Int. Conf. Signal Processing and Integrated Networks* (*SPIN*), Noida, India, 2019, pp. 772–775.

[18] C. Liu, Y. V. Zakharov, and T. Chen, Broadband underwater localization of multiple sources using basis pursuit de-noising, *IEEE Trans. Signal Process.*, vol. 60, no. 4, pp. 1708–1717, 2012.

[19] Z. Tang, G. Blacquiere, and G. Leus, Aliasing-free wideband beamforming using sparse signal representation, *IEEE Trans. Signal Process.*, vol. 59, no. 7, pp. 3464–3469, 2011.

[20] E. Candes and T. Tao, The Dantzig selector: Statistical estimation when p is much larger than n, *Ann. Statist.*, vol. 35, no. 6, pp. 2313–2351, 2007.

[21] D. P. Wipf and B. D. Rao, An empirical Bayesian strategy for solving the simultaneous sparse approximation problem, *IEEE Trans. Signal Process.*, vol. 55, no. 7, pp. 3704–3716, 2007.

[22] R. Wang, J. Zhang, S. Ren, and Q. Li, A reducing iteration orthogonal matching pursuit algorithm for compressive sensing, *Tsinghua Science and Technology*, vol. 21, no. 1, pp. 71–79, 2016.

[23] D. Malioutov, M. Cetin, and A. S. Willsky, A sparse signal reconstruction perspective for source localization with sensor arrays, *IEEE Trans. Signal Process.*, vol. 53, no. 8, pp. 3010–3022, 2005.

[24] 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 Trans. Wirel. Commun.*, vol. 11, no. 10, pp. 3607–3617, 2012.

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

[26] J. Li, Y. Li, and X. Zhang, Two-dimensional off-grid DOA estimation using unfolded parallel coprime array, *IEEE Commun. Lett.*, vol. 22, no. 12, pp. 2495–2498, 2018.

[27] J. Yuan, G. Zhang, H. Leung, and S. Ma, Off-grid DOA estimation for noncircular signals via block sparse representation using extended transformed nested array, *IEEE Signal Process. Lett.*, vol. 30, pp. 130–134, 2023.

[28] Z. M. Liu, Z. T. Huang, Y. Y. Zhou, and J. Liu, Direction-of-arrival estimation of noncircular signals via sparse representation, *IEEE Trans. Aerosp. Electron. Syst.*, vol. 48, no. 3, pp. 2690–2698, 2012.

[29] J. Yang and Y. Yang, Sparse Bayesian DOA estimation using hierarchical synthesis lasso priors for off-grid signals, *IEEE Trans. Signal Process.*, vol. 68, pp. 872–884, 2020.

[30] Z. M. Liu, Z. T. Huang, and Y. Y. Zhou, Sparsity-inducing direction finding for narrowband and wideband signals based on array covariance vectors, *IEEE Trans. Wireless Commun.*, vol. 12, no. 8, pp. 1–12, 2013.

[31] Z. Yang, L. Xie, and C. Zhang, Off-grid direction of arrival estimation using sparse Bayesian inference, *IEEE Trans. Signal Process.*, vol. 61, no. 1, pp. 38–43, 2013.

[32] A. Das and T. J. Sejnowski, Narrowband and wideband off-grid direction-of-arrival estimation via sparse Bayesian learning, *IEEE J. Oceanic Eng.*, vol. 43, no. 1, pp. 108–118, 2018.

[33] A. Das, Deterministic and Bayesian sparse signal processing algorithms for coherent multipath directions-of-arrival (DOAs) estimation, *IEEE J. Oceanic Eng.*, vol. 44, no. 4, pp. 1150–1164, 2019.

[34] L. Wang, L. Zhao, G. Bi, C. Wan, L. Zhang, and H. Zhang, Novel wideband DOA estimation based on sparse Bayesian learning with dirichlet process priors, *IEEE Trans. Signal Process.*, vol. 64, no. 2, pp. 275–289, 2016.

[35] X. Lu, C. N. Manchon, and Z. Wang, Collapsed VBI-DP based structured sparse channel estimation algorithm for massive MIMO-OFDM, *IEEE Access*, vol. 7, pp. 16665–16674, 2019.

[36] X. Lu, C. Zhang, and Z. Wang, Combined belief propagation-mean field message passing algorithm for dirichlet process mixtures, *IEEE Signal Process. Lett.*, vol. 26, no. 7, pp. 1041–1045, 2019.

[37] M. Li, X. Lu, J. Li, and S. Guan, A combined BP-MF algorithm for wideband off-grid DOA estimation with DP prior, *Electron. Lett.*, vol. 58, no. 25, pp. 1002–1005, 2022.

[38] A. Das, Real-valued sparse Bayesian learning for off-grid direction-of-arrival (DOA) estimation in ocean acoustics, *IEEE J. Oceanic Eng.*, vol. 46, no. 1, pp. 172–182, 2021.

[39] J. Dai, X. Xu, and D. Zhao, Direction-of-arrival estimation via real-valued sparse representation, *Antennas Wirel. Propag. Lett.*, vol. 12, pp. 376–379, 2013.

[40] 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 Trans. Antennas Propagat.*, vol. 54, no. 1, pp. 175–181, 2006.

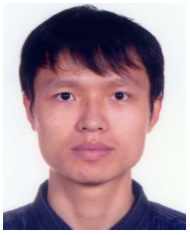
[41] K. C. Huarng and C. C. Yeh, A unitary transformation method for angle-of-arrival estimation, *IEEE Trans. Signal Process.*, vol. 39, no. 4, pp. 975–977, 1991.

[42] J. Dai and H. C. So, Real-valued sparse Bayesian learning for DOA estimation with arbitrary linear arrays, *IEEE Trans. Signal Process.*, vol. 69, pp. 4977–4990, 2021.
[43] Z. Yuan, C. Zhang, Q. Guo, Z. Wang, X. Lu, and S. Wu, Combined message passing based SBL with dirichlet process prior for sparse signal recovery with multiple measurement vectors, *IEEE Access*, vol. 6, pp. 13181–13190, 2018.
[44] S. Rangan, Generalized approximate message passing for estimation with random linear mixing, in *Proc. 2011 IEEE Int. Symp. Information Theory Proceedings*, St. Petersburg, Russia, 2011, pp. 2168–2172.
[45] S. Rangan, P. Schniter, A. K. Fletcher, and S. Sarkar, On the convergence of approximate message passing with arbitrary matrices, *IEEE Trans. Inform. Theory*, vol. 65, no. 9, pp. 5339–5351, 2019.
[46] L. Liu, S. Huang, and B. M. Kurkoski, Memory approximate message passing, in *Proc. 2021 IEEE Int. Symp. Information Theory* (*ISIT*), Melbourne, Australia, 2021, pp. 1379–1384.
[47] M. Luo, Q. Guo, M. Jin, Y. C. Eldar, D. Huang, and X. Meng, Unitary approximate message passing for sparse Bayesian learning, *IEEE Trans. Signal Process.*, vol. 69, pp. 6023–6039, 2021.
[48] F. R. Kschischang, B. J. Frey, and H. A. Loeliger, Factor graphs and the sum-product algorithm, *IEEE Trans. Inform. Theory*, vol. 47, no. 2, pp. 498–519, 2001.
[49] J. Yi, D. Mirza, C. Schurgers, and R. Kastner, Joint time synchronization and tracking for mobile underwater systems, in *Proc. 8th Int. Conf. Underwater Networks & Systems*, Kaohsiung, China, 2013, pp. 1–8.
[50] C. Zhang, Z. Yuan, Z. Wang, and Q. Guo, Low complexity sparse Bayesian learning using combined belief propagation and mean field with a stretched factor graph, *Signal Process.*, vol. 131, pp. 344–349, 2017.
[51] G. Colavolpe and G. Germi, On the application of factor graphs and the sum–product algorithm to ISI channels, *IEEE Trans. Commun.*, vol. 53, no. 5, pp. 818–825, 2005.
[52] J. Dauwels, On variational message passing on factor graphs, in *Proc. 2007 IEEE Int. Symp. Information Theory*, Nice, France, 2007, pp. 2546–2550.
[53] P. Stoica, E. G. Larsson, and A. B. Gershman, The stochastic CRB for array processing: A textbook derivation, *IEEE Signal Process. Lett.*, vol. 8, no. 5, pp. 148–150, 2001.

Shanwen Guan received the BEng degree in electronic information engineering from Wuhan Institute of Technology, Wuhan, China in 2016. He is currently pursuing the PhD degree at Guangxi Key Laboratory of Image and Graphic Intelligent Processing, School of Computer Science and Information Security, Guilin University of Electronic Technology, Guilin, China. His research interests include variational Bayesian inference, sparse signal recovery techniques, and indoor positioning techniques.

Xinhua Lu received the BEng, MEng, and PhD degrees from Zhengzhou University, Zhengzhou, China in 2003, 2007, and 2019, respectively. Now he is a lecturer at School of Information Engineering, Nanyang Institute of Technology, Nanyang, China. His main research interests include machine learning, variational Bayesian inference, and wireless communication.

Ji Li received the PhD degree from Université de Montréal, Canada in 2007. Since 2020, he has been a professor at Guangxi Key Laboratory of Image and Graphic Intelligent Processing, School of Computer Science and Information Security, Guilin University of Electronic Technology, China. His main research fields include indoor positioning, wireless communication, artificial intelligence, palm vein recognition, and multi-source data fusion.



Rushi Lan received the BS degree in information and computing science and the MS degree in applied mathematics from Nanjing University of Information Science and Technology, Nanjing, China in 2008 and 2011, respectively, and the PhD degree in software engineering from University of Macau, Macau, China in 2016. He is currently a professor at Guangxi Key Laboratory of Image and Graphic Intelligent Processing, School of Computer Science and Information Security, Guilin University of Electronic Technology, Guilin, China. His research interests include image classification, image denoising, and metric learning.

Xiaonan Luo was previously the director of National Engineering Research Center of Digital Life at Sun Yat-sen University, Guangzhou, China. He is currently a professor at Guangxi Key Laboratory of Image and Graphic Intelligent Processing, School of Computer Science and Information Security, Guilin University of Electronic Technology, Guilin, China. He has also received the National Science Fund for Distinguished Young Scholars, which was granted by the National Natural Science Foundation of China. His current research interests include computer graphics, indoor and outdoor location, and pattern recognition.