

## Article

# An Off-Grid DOA Estimation Method via Fast Variational Sparse Bayesian Learning

Xin Tong, Yuzhuo Chen, Zhongliang Deng  and Enwen Hu \* 

School of Electronic Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China; tongxin0110@bupt.edu.cn (X.T.); chenyzhuo@bupt.edu.cn (Y.C.); dengzhl@bupt.edu.cn (Z.D.)

\* Correspondence: owen.hu@bupt.edu.cn

## Abstract

In practical array signal processing applications, direction-of-arrival (DOA) estimation often suffers from degraded accuracy under low signal-to-noise ratio (SNR) and limited snapshot conditions. To address these challenges, we propose an off-grid DOA estimation method based on Fast Variational Bayesian Inference (OGFVBI). Within the variational Bayesian framework, we design a fixed-point criterion rooted in root-finding theory to accelerate the convergence of hyperparameter learning. We further introduce a grid fission and adaptive refinement strategy to dynamically adjust the sparse representation, effectively alleviating grid mismatch issues in traditional off-grid approaches. To address frequency dispersion in wideband signals, we develop an improved subspace focusing technique that transforms multi-frequency data into an equivalent narrowband model, enhancing compatibility with subspace DOA estimators. We demonstrate through simulations that OGFVBI achieves high estimation accuracy and resolution while significantly reducing computational time. Specifically, our method achieves more than 37.6% reduction in RMSE and at least 28.5% runtime improvement compared to other methods under low SNR and limited snapshot scenarios, indicating strong potential for real-time and resource-constrained applications.

**Keywords:** DOA estimation; focusing algorithm; fast variational Bayesian inference; off-grid model



Academic Editor: Reza K. Amineh

Received: 28 May 2025

Revised: 7 July 2025

Accepted: 8 July 2025

Published: 10 July 2025

**Citation:** Tong, X.; Chen, Y.; Deng, Z.; Hu, E. An Off-Grid DOA Estimation Method via Fast Variational Sparse Bayesian Learning. *Electronics* **2025**, *14*, 2781. <https://doi.org/10.3390/electronics14142781>

**Copyright:** © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Direction-of-arrival (DOA) estimation, as a crucial branch of array signal processing, has been extensively studied [1] and plays a pivotal role in various fields such as radar, sonar, mobile communications, and autonomous driving.

### 1.1. Subspace Methods

Given that modern wireless communication systems, such as 4G and 5G, predominantly use broadband signals to enhance data transmission rates and spectrum efficiency, this paper focuses on DOA estimation methods based on 5G OFDM signals. In this context, a broadband signal refers to one whose bandwidth is not negligible compared to its carrier frequency, typically characterized by a fractional bandwidth (i.e., the ratio of signal bandwidth to center frequency) exceeding 5–10%, whereas signals with fractional bandwidth below 3–5% are generally treated as narrowband in array processing literature [2]. This distinction is essential because traditional narrowband DOA models rely on the assumption that all signal frequency components experience nearly identical phase shifts across the

array—a condition no longer valid in broadband scenarios [3]. Existing broadband DOA estimation methods primarily fall into two categories: Maximum Likelihood (ML) [4], which suffers from high computational complexity due to multidimensional search, and subspace-based methods, which offer lower complexity by extending narrowband techniques and is more suitable for wideband array signal processing applications.

For broadband DOA estimation, the Coherent Signal Subspace Method (CSSM) [5] outperforms incoherent approaches by aligning multi-frequency data via a focusing matrix, enabling narrowband processing while handling coherent sources. Although CSSM reduces computational complexity and maintains robustness under low SNR [6,7], its performance depends heavily on the initial angle estimation and still faces complexity challenges.

Over the past few decades, narrowband DOA estimation methods have matured significantly, ushering in the era of high-resolution DOA estimation. Subspace-based methods such as MUSIC [8,9] and ESPRIT [10,11] achieve high-resolution DOA estimation through signal–noise subspace decomposition. However, they require prior knowledge of source number, high SNR, and sufficient snapshots for accurate covariance estimation, and fail with coherent sources due to multipath effects [12,13].

### 1.2. Sparse Bayesian Methods

Recent advances in compressed sensing have transformed array DOA estimation by utilizing signal sparsity in spatial or transformed domains [14]. This approach reformulates the problem as Sparse Signal Recovery (SSR) [15,16]. Unlike traditional subspace methods, sparse reconstruction performs better with low SNR and limited snapshots while being more noise-resistant. The key idea is to extract sparse signals from an overcomplete dictionary by identifying key components. While traditional SBL methods work well for these problems, their performance suffers from slow hyperparameter convergence and spatial discretization issues in low-SNR conditions, particularly with parameter circular dependency and grid mismatch.

To solve the parameter circular dependency problem, Luessi et al. [17] developed a sparse Bayesian algorithm for multi-measurement vectors. Their method uses iterative maximization within a variational Bayesian framework for accurate hyperparameter estimation. Despite this improvement, block-sparse models still face slow convergence and high computational costs. Subsequent work by Tipping et al. [18] accelerated convergence through fast marginal likelihood maximization, while Fang et al. [19] improved complex signal handling with mode-coupled learning. However, achieving both fast and accurate reconstruction remains challenging in practice.

While the aforementioned improvements address computational challenges, recent advances in SBL further enhance its capabilities. For example, Refs. [20–23] further focused on reducing computational load by incorporating various real-valued transformations within the compressed sensing framework; Ref. [24] introduces a generalized double Pareto (GDP) prior—sparser than the Laplace prior—for improved signal modeling, while [25] proposes a spatial alternating method to accelerate convergence. Although these methods show promise, DOA estimation accuracy remains limited under single-snapshot or coarse-grid conditions. Building on these developments, variational sparse Bayesian inference methods [26–29] further refine the SBL framework for enhanced estimation performance.

### 1.3. On-Grid vs. Off-Grid Challenges

The off-grid problem arises because, within the SSR framework, it is assumed that the true signal DOA lies on some predefined fixed grid points. However, since signal sources are continuously distributed in space, the true DOA always falls outside the grid points, regardless of how fine the grid is. Therefore, when the DOA does not exactly align with the

predefined sampling grid points, the performance of these methods degrades due to the off-grid problem. The off-grid problem also occurs in DOA estimation problems involving arrays with special structures, including sparse arrays [30] and coprime planar arrays [31].

To address this issue, Yang et al. [32] proposed an off-grid DOA estimation method based on Sparse Bayesian Learning (OGSBI). Although the OGSBI method effectively achieves off-grid DOA estimation, its performance is less than ideal when using coarse grids. Dai et al. [33] introduced a root-off-grid sparse Bayesian DOA estimation method (RootSBL), which improves the accuracy of coarse grid DOA estimation while reducing computational complexity. Subsequently, in [34], a second-order Taylor approximation was introduced to reduce approximation errors, addressing the issue of low estimation accuracy with first-order Taylor approximations under coarse grids, though the computational complexity remained relatively high. Additionally, there are off-grid SSR methods that do not rely on grid constraints, such as the method based on atomic norm minimization (ANM) [35].

#### 1.4. Deep Learning Methods

Additionally, with the development of machine learning technologies, some studies have begun to explore the application of deep learning techniques to DOA estimation. Currently, most deep learning-based DOA estimation methods adopt classification models. For example, in [36], the DeepMUSIC framework is proposed, where the DOA region is divided into subregions, and a CNN is trained for each subregion. In [37], under low SNR conditions, the real part, imaginary part, and phase of the covariance matrix are used as three-channel inputs. In [38], a residual network is used to model array defects robustly. However, these methods typically require large amounts of training data, and their performance may be affected when the data distribution changes.

Table 1 provides a comparative summary of the above existing DOA estimation methods, highlighting their respective advantages and limitations in practical DOA estimation scenarios.

**Table 1.** Comparison of advantages and disadvantages of DOA estimation methods.

Method	Algorithm	Advantages	Disadvantages
Subspace Methods	MUSIC-type algorithms ESPRIT-type algorithms	High resolution under high SNR and sufficient snapshots	Poor performance under low SNR and limited snapshots; cannot handle coherent signals
Compressed Sensing-based	Convex Relaxation Optimization	Strong robustness; high resolution; can handle coherent signals	High computational complexity; requires parameter tuning
	Sparse Bayesian Learning	Strong robustness; no prior sparse information needed; can handle coherent signals	High computational complexity
Deep Learning-based	CNN/ResNet architectures	Data-driven adaptability; robust to array imperfections; potential for real-time inference	Requires massive training data; sensitive to distribution shifts; black-box interpretation challenges

#### 1.5. Practical Applications and Implementations of DOA Estimation

DOA estimation is widely used in practical systems, especially in adaptive beamforming. It helps improve spatial resolution and reduce interference. A typical example is the Capon beamformer, also known as the Minimum Variance Distortionless Response (MVDR) algorithm. It minimizes the output power while keeping a distortionless response in the desired direction [39]. Compared to traditional beamforming, it provides better

angular resolution. However, its performance may drop when snapshots are limited or model mismatch occurs.

Furthermore, DOA estimation techniques have been extended to real-time beamsteering hardware. For instance, adaptive DDS-PLL (Direct Digital Synthesizer–Phase-Locked Loop) architectures have been proposed for real-time AoA estimation and dynamic beam control in phased arrays [40]. Reliable and fast DOA estimation is essential in these systems, highlighting the importance of robust DOA algorithms in practical deployments.

In addition to software-based methods, DOA estimation has also been implemented in hardware. Phase interferometry-based techniques are often used because of their low complexity. These methods allow DOA estimation to run in real time with high efficiency. Reconfigurable digital architectures have also been developed. These can execute DOA algorithms directly in hardware, which is suitable for embedded platforms with strict latency and power limits [41].

This paper proposes an off-grid DOA estimation method based on Fast Variational Bayesian Inference (OGFVBI) to address the problems of grid mismatch, slow convergence, and high complexity in existing sparse Bayesian approaches. To the best of our knowledge, this work is the first to combine fixed-point acceleration and dynamic dual-stage grid evolution within a unified variational Bayesian framework. This integration leads to improved estimation accuracy and efficiency, particularly under low SNR and limited snapshot conditions.

The main contributions of this work are threefold:

- (1) A fixed-point update scheme is designed within the variational Bayesian inference framework, avoiding slow EM-type updates and improving convergence stability. This addresses the main limitation of real-valued variational Bayesian inference (realVBI) in convergence speed.
- (2) A dual-stage grid fission mechanism is developed to dynamically expand and refine the grid using posterior information. This two-phase strategy improves resolution and robustness compared to the single-stage grid refinement in RootSBL.
- (3) A frequency-aligned subspace focusing technique is proposed for broadband OFDM signals, enabling covariance reconstruction and compatibility with narrowband DOA methods across subcarriers.

Together, these components form a unified framework that achieves higher accuracy, lower complexity, and faster convergence than previous off-grid Bayesian DOA methods. Simulation results show that the proposed method achieves high estimation accuracy and resolution, while maintaining low computational overhead, especially under low SNR and limited snapshot conditions.

The remainder of this paper is organized as follows. Section 2 describes the signal model. Section 3 presents the proposed improved SBL algorithm in detail. Section 4 provides simulation results and performance comparisons with existing methods. Finally, Section 5 concludes the paper.

## 2. Signal Model

Assume there are  $K$  distant signal sources arriving at an array formed by  $M$  elements, with the array arranged on a uniform linear array ( $K < M$ ). The element spacing is  $d = \lambda/2$ , where  $\lambda$  represents the wavelength and the signal's snapshot number is  $T$ . The angle of arrival of the signals is  $\theta_k$  ( $k = 1, 2, \dots, K$ ).

Considering the first array element as the reference, the array steering vector corresponding to this angle is

$$\mathbf{a}(\theta_k) = \left[ 1 \ e^{-j\frac{2\pi}{\lambda} \sin \theta_k} \dots e^{-j\frac{2\pi}{\lambda} d(M-1) \sin \theta_k} \right]^T \quad (1)$$

The signal received by the  $m$ -th array element can be expressed as

$$x_m(t) = \sum_{k=1}^K s_k(t) e^{-j \frac{2\pi}{\lambda} (m-1)d \sin \theta_k} + n_m(t) \quad (2)$$

where  $m = 1, 2, \dots, M$ ,  $s_k(t)$  ( $k = 1, 2, \dots, K$ ) is the signal from the  $k$ -th source arriving at the uniform linear array,  $n_m(t)$  is Gaussian white noise, and  $\theta_k$  is the angle of arrival of the  $k$ -th source.  $[\cdot]^T$  represents the transpose of the array.

The array data received in this configuration can be expressed in matrix form as

$$\mathbf{x}(t) = \mathbf{A}(\boldsymbol{\theta})\mathbf{s}(t) + \mathbf{n}(t) \quad (3)$$

The formula  $\mathbf{x}(t) = [x_1(t), x_2(t), \dots, x_M(t)]^T$  represents the received signal of the array, with  $t = 1, 2, \dots, T$ ;  $\mathbf{s}(t) = [s_1(t), s_2(t), \dots, s_K(t)]^T$  represents the incoming signal;  $\mathbf{n}(t)$  is Gaussian white noise with a mean of zero, having a high power spectral density, and is independent of the incoming signal; and  $\mathbf{A}(\boldsymbol{\theta}) = [\mathbf{a}(\theta_1), \mathbf{a}(\theta_2), \dots, \mathbf{a}(\theta_K)]$  represents the array steering vector, where the direction of the array is represented by (1).

The received signal from the array can be written in matrix form, as

$$\mathbf{X} = \mathbf{A}(\boldsymbol{\theta})\mathbf{S} + \mathbf{N} \quad (4)$$

When the signal sources and noise are uncorrelated, the covariance matrix of the data received by the array elements can be expressed as

$$\mathbf{R} = \mathbb{E}\{\mathbf{x}(t)\mathbf{x}^H(t)\} = \mathbf{A}\mathbf{R}_s\mathbf{A}^H + \sigma^2\mathbf{I}_M \quad (5)$$

where  $\mathbb{E}\{\cdot\}$  represents the expectation operator;  $\mathbf{R}_s$  is the signal's autocorrelation matrix;  $\sigma^2$  is the noise power; and  $\mathbf{I}_M$  is the  $M \times M$  identity matrix. By performing normalization on (5), we obtain

$$\mathbf{y} = \text{vec}(\mathbf{R}) = G(\boldsymbol{\theta})\mathbf{p} + \sigma^2\text{vec}(\mathbf{I}_M) \quad (6)$$

where  $\text{vec}(\cdot)$  represents the vectorization operation;  $G(\boldsymbol{\theta}) = \mathbf{A}^*(\boldsymbol{\theta}) \circ \mathbf{A}(\boldsymbol{\theta})$  is a  $M^2 \times K$  dimensional matrix;  $(\cdot)^*$  represents the conjugate transpose operation; and  $\circ$  denotes the Khatri–Rao product operation; and  $\mathbf{p} = [\sigma_1^2, \sigma_2^2, \dots, \sigma_K^2]^T$  is the energy vector of the signal sources.

Due to the assumption that the received signal fills the entire spatial domain with uniform distribution, the entire spatial domain is discretized into  $N$  grid points, where  $\bar{\boldsymbol{\theta}} = [\bar{\theta}_1, \bar{\theta}_2, \dots, \bar{\theta}_N]^T$  is the  $n$ -th grid point, and  $N \gg K$ . Each grid point represents a potential angle of arrival, thus forming a complete dictionary set

$$G(\bar{\boldsymbol{\theta}}) = \mathbf{A}^*(\bar{\boldsymbol{\theta}}) \circ \mathbf{A}(\bar{\boldsymbol{\theta}}) = [\mathbf{a}^*(\bar{\theta}_1) \otimes \mathbf{a}(\bar{\theta}_1), \mathbf{a}^*(\bar{\theta}_2) \otimes \mathbf{a}(\bar{\theta}_2), \dots, \mathbf{a}^*(\bar{\theta}_N) \otimes \mathbf{a}(\bar{\theta}_N)] \quad (7)$$

where  $\otimes$  represents the Kronecker product.

Equation (6) can thus be transformed into a fully sparse representation model:

$$\mathbf{y} = G(\bar{\boldsymbol{\theta}})\mathbf{p} + \sigma^2\text{vec}(\mathbf{I}_M) \quad (8)$$

Based on the actual situation, the signal source's position is unlikely to be placed neatly on a predefined discretized grid. Therefore, we consider introducing an off-grid model. Suppose  $r = \bar{\theta}_2 - \bar{\theta}_1$  is the uniform distribution between grid points, and  $\hat{\theta}_k$  represents the

actual angle of arrival. It does not belong to the grid set, i.e.,  $\tilde{\theta}_k \notin \{\bar{\theta}_1, \bar{\theta}_2, \dots, \bar{\theta}_N\}$ . In the case of the two grid points, a small step expansion is performed, and for  $\tilde{\theta}_k$  we have

$$\mathbf{a}(\tilde{\theta}_k) \approx \mathbf{a}(\tilde{\theta}_{nk}) + \mathbf{b}(\tilde{\theta}_{nk})(\tilde{\theta}_k - \bar{\theta}_{nk}) \approx \mathbf{a}(\bar{\theta}_{nk}) + \mathbf{b}(\bar{\theta}_{nk})\beta_{nk} \quad (9)$$

In the equation, the grid mismatch  $\beta_{nk} = \tilde{\theta}_k - \bar{\theta}_{nk}$  and  $\mathbf{b}(\bar{\theta}_{nk}) = \mathbf{a}'(\bar{\theta}_{nk})$  is the steering vector of  $\mathbf{a}(\bar{\theta}_{nk})$ , where the matrix  $\mathbf{B} = [\mathbf{b}(\bar{\theta}_1), \dots, \mathbf{b}(\bar{\theta}_N)]$  is constructed from  $\mathbf{b}(\bar{\theta}_{nk})$  and  $\boldsymbol{\beta} = [\beta_1, \dots, \beta_N]$ . The off-grid model in (9) adopts a first-order Taylor expansion, which is effective when the angular offset between the actual DOA and the nearest grid point is small. However, for coarse grids, this linear approximation may introduce non-negligible errors. In such cases, higher-order expansions (e.g., second-order Taylor series as discussed in [34]) can be considered to improve the fidelity of the steering vector approximation. Nevertheless, to maintain computational efficiency, the proposed method retains the first-order model while mitigating grid mismatch via adaptive grid evolution and refinement.

Substituting Equation (9) into Equation (7), the complete dictionary set of the off-grid model can be expressed as

$$\begin{aligned} \boldsymbol{\Psi}(\boldsymbol{\beta}) &= \dot{\mathbf{A}} + \dot{\mathbf{B}}\text{diag}(\boldsymbol{\beta}) = \mathbf{A}^*(\bar{\boldsymbol{\theta}}) \circ \mathbf{A}(\bar{\boldsymbol{\theta}}) + \dot{\mathbf{B}}\text{diag}(\boldsymbol{\beta}) \\ &= [\dot{\mathbf{a}}(\bar{\theta}_1), \dot{\mathbf{a}}(\bar{\theta}_2), \dots, \dot{\mathbf{a}}(\bar{\theta}_N)] + \dot{\mathbf{B}}\text{diag}(\boldsymbol{\beta}) \end{aligned} \quad (10)$$

where  $\dot{\mathbf{a}}(\tilde{\theta}_k) \approx \dot{\mathbf{a}}(\bar{\theta}_{nk}) + \dot{\mathbf{b}}(\bar{\theta}_{nk})(\tilde{\theta}_k - \bar{\theta}_{nk})$ ,  $\dot{\mathbf{b}}(\bar{\theta}_{nk}) = \mathbf{b}^*(\bar{\theta}_{nk}) \otimes \mathbf{a}(\bar{\theta}_{nk}) + \mathbf{a}^*(\bar{\theta}_{nk}) \otimes \mathbf{b}(\bar{\theta}_{nk})$ ,  $\dot{\mathbf{B}} = [\dot{\mathbf{b}}(\bar{\theta}_1), \dots, \dot{\mathbf{b}}(\bar{\theta}_N)]$ , bold lowercase letters (e.g.,  $\mathbf{a}$ ) denote vectors, and bold uppercase letters (e.g.,  $\mathbf{A}$ ) denote matrices. The operator  $\text{diag}(\cdot)$  denotes a diagonal matrix formed from a vector. The notation  $\mathbf{a}(\theta)$  represents the array steering vector at angle  $\theta$ , while  $\dot{\mathbf{a}}(\theta)$  denotes its derivative with respect to  $\theta$ .

In summary, the off-grid sparse representation model can be expressed as

$$\mathbf{y} = \boldsymbol{\Psi}(\boldsymbol{\beta})\mathbf{p} + \sigma^2\text{vec}(\mathbf{I}_M) \quad (11)$$

According to [33], when the signal source is a wideband high-frequency signal, the mismatch of the steering vector of the received signal follows a Gaussian distribution.  $\hat{\mathbf{R}} = \left( \sum_{t=1}^T \mathbf{x}(t)\mathbf{x}^H(t) \right) / T$  represents the estimated covariance matrix of the actual scenario. The matrix  $\mathbf{R}$  is the ideal covariance matrix, and the estimation error between the two estimates is defined as  $\boldsymbol{\varepsilon} = \text{vec}(\hat{\mathbf{R}}) - \text{vec}(\mathbf{R}) = \hat{\mathbf{y}} - \mathbf{y}$ , where the mismatch follows a Gaussian distribution.

$$\boldsymbol{\varepsilon} \sim \text{ASN}(0, \mathbf{W}) \quad (12)$$

where  $\text{ASN}(\cdot)$  represents the asymptotic normal distribution,  $\mathbf{W} = (\mathbf{R}^T \otimes \mathbf{R}) / T$ , and  $\mathbf{R}$  is the ideal covariance matrix. Since  $\mathbf{R}$  cannot be directly obtained, the covariance of  $\boldsymbol{\varepsilon}$  will be approximated by  $\hat{\mathbf{W}} = (\hat{\mathbf{R}}^T \otimes \hat{\mathbf{R}}) / T$ . Therefore, the normalized sample covariance matrix can be expressed as

$$\hat{\mathbf{y}} = \text{vec}(\hat{\mathbf{R}}) = \boldsymbol{\Psi}(\boldsymbol{\beta})\mathbf{p} + \sigma^2\text{vec}(\mathbf{I}_M) + \boldsymbol{\varepsilon} \quad (13)$$

Through sparse Bayesian learning, the main difficulty in directly processing Equation (13) is that  $\sigma^2$  is unknown and  $\boldsymbol{\varepsilon}$  follows a Gaussian distribution, which is not a standard SBL form. To solve this issue, we treat  $\sigma^2$  as an unknown part of the signal, i.e.,

$$\hat{\mathbf{y}} = \begin{bmatrix} \boldsymbol{\Psi}(\boldsymbol{\beta}) & \mathbf{I}_M \end{bmatrix} \begin{bmatrix} \mathbf{p} \\ \sigma^2 \end{bmatrix} + \boldsymbol{\varepsilon} \quad (14)$$



where  $\mathbf{I}_M = [e_1^T, e_2^T, \dots, e_M^T]^T$ ;  $e_m$  is a vector with all elements zero except the  $m$ -th element, which is one.

To facilitate sparse Bayesian analysis, we define  $\Phi(\beta) = [\Psi(\beta) \mathbf{I}_M]$  and  $s = [p \quad \sigma^2]^T$ , which can be expressed as Equation (14).

$$\hat{\mathbf{y}} = \Phi(\beta)s + \varepsilon \quad (15)$$

### 3. Methodology

#### 3.1. Traditional DOA Estimation Algorithms

Subspace-based DOA estimation methods have long been regarded as classical techniques in array signal processing, owing to their ability to achieve high-resolution direction estimation. Among them, the most representative is the Multiple Signal Classification (MUSIC) algorithm.

The MUSIC algorithm is based on the eigenvalue decomposition of the covariance matrix of the array received data. It separates the signal and noise subspaces and utilizes their orthogonality to construct a spatial spectrum function. Specifically, the covariance matrix of the data received by the array elements is shown in Equation (5).

After eigenvalue decomposition, the noise subspace  $U_N$  is obtained. The MUSIC spatial spectrum is then constructed as follows:

$$P_{\text{MUSIC}} = \frac{1}{a^H(\theta)U_N U_N^H a(\theta)} \quad (16)$$

DOA estimates are then determined by searching for the peaks of the MUSIC spatial spectrum, where the steering vector is most orthogonal to the noise subspace.

Although MUSIC achieves excellent resolution under ideal conditions (e.g., high SNR, sufficient snapshots), it has several limitations. It requires prior knowledge of the number of sources and performs poorly when signal sources are coherent or when the number of snapshots is limited. Moreover, it is sensitive to estimation errors in the sample covariance matrix.

In this work, MUSIC serves as a benchmark method for evaluating the performance of the proposed algorithm.

In contrast, SBL-based DOA estimation algorithms can fully utilize the discontinuous virtual element information discarded by subspace-based methods. Even when the number of signal sources to be estimated is unknown, they can still achieve high-precision DOA estimation with limited snapshots.

#### 3.2. Broadband Signal Processing: Focusing Algorithm

Since the goal of this paper is to implement DOA estimation for broadband OFDM signals, it is necessary to first apply a frequency focusing method to preprocess the broadband signal. We adopt an improved subspace-based focusing technique to align the frequency components across subcarriers and reconstruct an equivalent narrowband covariance matrix. This enables the application of the proposed OGFVBI algorithm to estimate DOAs from broadband observations. Compared with the typical Two-sided Correlation Transformation (TCT) method [7], the proposed focusing scheme is simpler in construction, thereby lowering computational complexity while maintaining spatial fidelity.

Focusing inevitably alters the signal structure and noise behavior. To ensure that the SNR remains consistent, we apply a normalization criterion such that the focusing matrix is unitary [4]. Under this condition, the output noise remains Gaussian, and the post-focusing SNR is preserved.

To enable the aforementioned focusing and subspace processing, we make several standard assumptions. First, the  $K$  signal sources are mutually uncorrelated and located in the far field, leading to planar wavefronts. The received signal comprises multiple sub-carrier components, each modeled as narrowband due to the sufficiently small subcarrier spacing. The additive noise at each antenna element is modeled as zero-mean, circularly symmetric complex Gaussian noise, which is temporally and spatially white with identical variance  $\sigma^2$ . The received signal can then be rewritten as

$$x(f_k, t) = A(f_k, \theta)s(f_k, t) + n(f_k, t) \quad (17)$$

where  $A(f_k, \theta)$  is the array steering matrix, and  $s(f_k, t)$  is the received signal data on subcarrier  $f_k$ . The covariance matrix of the received signal can be written as follows:

$$R_Y(f_k) = A(f_k, \theta)R_S(f_k)A^H(f_k, \theta) + \sigma^2 I \quad (18)$$

where  $R_S(f_k)$  is the covariance matrix of  $S(f_k)$ , and  $\sigma^2 I$  denotes the covariance matrix of spatially white complex Gaussian noise. The signal subspace of  $R_Y(f_k)$  can be obtained through eigenvalue decomposition:

$$R_Y(f_k) = V_k \Sigma_{s,k} V_k^H + G_k \Sigma_{w,k} G_k^H \quad (19)$$

where  $V_k$  is the signal subspace, and  $G_k$  is the noise subspace. The focusing matrix  $T(f_k)$  is constructed as

$$T(f_k) = V_0 V_k^H \quad (20)$$

where  $V_0$  is the signal subspace of  $R_Y(f_0)$  or  $A(f_0, \theta)R_S(f_0)A^H(f_0, \theta)$ , and the calculation formula for  $R_S(f_0)$  is as follows:

$$R_S(f_0) = \frac{1}{K} \sum_{k=1}^K \hat{R}_S(f_k) \quad (21)$$

The estimate of  $R_S(f_k)$ , denoted as  $\hat{R}_S(f_k)$ , is as follows:

$$\hat{R}_S(f_k) = \left( A_k^H A_k \right)^{-1} A_k^H \left[ \hat{R}_k - \hat{\sigma}_k^2 I \right] A_k \left( A_k^H A_k \right)^{-1} \quad (22)$$

where  $A_k$  and  $\hat{R}_k$  are abbreviations for  $A(f_k, \theta)$  and  $\hat{R}_Y(f_k)$ , respectively. The optimization problem for  $\theta$  can be expressed as

$$\min_w w^H R w \quad \text{s.t.} \quad w^H a(\theta) a^H(\theta) w + \beta \|w\|^2 = c \quad (23)$$

where  $a(\theta)$  is the array steering vector,  $R$  is the covariance matrix of the received signal,  $w$  is the beamforming weight vector, and  $c, \beta > 0$  are constants.

This optimization problem can be solved using the Lagrange multiplier method:

$$f(w) = w^H R w - \lambda \left( w^H a(\theta) a^H(\theta) w + \beta \|w\|_2^2 - c \right) \quad (24)$$

where  $\beta$  is determined by the following equation:

$$\beta = \frac{kM}{\xi_{M-1} - \xi_{M-1}}, \quad k \leq 1 \quad (25)$$



where  $M$  is the number of array elements and  $\zeta_{M-1}$  and  $\zeta_M$  are the smallest and second smallest eigenvalues of the array's received signal covariance matrix, respectively. Therefore, we calculate for each angle  $\theta$ :

$$\hat{B}(\theta) = \left( \beta^{-1} \mathbf{I} - \alpha \mathbf{a}(\theta) \mathbf{a}^H(\theta) \right) \mathbf{R} \quad (26)$$

where  $\alpha$  is determined by the following equation:

$$\alpha = \frac{\beta^{-2}}{1 + \beta^{-1} \|\mathbf{a}(\theta)\|_2^2} = \frac{\beta^{-2}}{1 + \beta^{-1} M} \quad (27)$$

where  $\mathbf{w}$  is the eigenvector corresponding to the smallest eigenvalue of  $\hat{B}(\theta)$ .

Thus, the spatial power spectrum is defined as

$$P(\theta) = \frac{1}{|\mathbf{w}^H \mathbf{a}(\theta)|^2} \quad (28)$$

For each angle this power spectrum is computed and the angle  $\theta$  corresponding to its peak is obtained, i.e., the angle and its number obtained by pre-estimation is used as an input to the subsequent algorithm.

Then, using the above calculations,  $T(f_k)$  is applied to focus the correlation matrix of the received data as

$$\mathbf{T}(f_k) \mathbf{R}_k \mathbf{T}^H(f_k) = \mathbf{V}_0 \mathbf{V}_k^H \mathbf{V}_k \Sigma_{s,k} \mathbf{V}_k^H \mathbf{V}_k \mathbf{V}_0^H + \sigma^2 \mathbf{V}_0 \mathbf{V}_k^H \mathbf{G}_k \mathbf{G}_k^H \mathbf{V}_k \mathbf{V}_0^H \quad (29)$$

Since the signal subspace and noise subspace are orthogonal, the focused correlation matrix for all subcarriers can be simplified as

$$\sum_{k=1}^K \mathbf{T}(f_k) \mathbf{R}_k \mathbf{T}^H(f_k) = \sum_{k=1}^K \mathbf{V}_0 \Sigma_{s,k} \mathbf{V}_0^H \quad (30)$$

where  $\sum_{k=1}^K \mathbf{T}(f_k) \mathbf{R}_k \mathbf{T}^H(f_k)$  is the array covariance matrix after focusing, which can replace  $\hat{\mathbf{R}}$  in Equation (13) in Section 2, and then proceed with the subsequent VBI processing.

When the number of sources is  $N$  and the number of array elements is  $M$ , the computational complexity of constructing the focusing matrix in the proposed method is  $\mathcal{O}(M \times N^2)$ , whereas that of the traditional TCT algorithm is  $\mathcal{O}(M^3)$ . Since  $N < M$  in most practical scenarios, the proposed method offers lower computational complexity and a simpler construction procedure.

In terms of estimation accuracy, the proposed method yields the same eigenvectors of the covariance matrix as the TCT algorithm after focusing. Therefore, it achieves comparable DOA estimation accuracy to that of the TCT method.

### 3.3. Variational Bayesian Inference

In this work, we adopt variational Bayesian inference (VBI) to estimate the DOA parameters. Compared to EM, which provides only point estimates and can suffer from slow or unstable convergence, VBI estimates full posteriors and improves robustness. While MCMC methods can be accurate, they are computationally intensive and unsuitable for real-time processing.

VBI offers a good balance between accuracy and efficiency by turning inference into an optimization problem with closed-form updates. In our framework, VBI is further

accelerated using a fixed-point update scheme and combined with adaptive grid evolution, making it well-suited for fast and scalable off-grid DOA estimation.

Following the variational Bayesian framework, we begin by approximating the posterior distribution of the signal component. Based on the joint likelihood in Equation (15), the update expression for the signal posterior can be derived as follows:

$$p(\tilde{\mathbf{y}} \mid \mathbf{s}; \boldsymbol{\beta}) = \text{CN}(\tilde{\mathbf{y}} \mid \boldsymbol{\Phi}(\boldsymbol{\beta})\mathbf{s}, I_N) \quad (31)$$

Construct a two-stage hierarchical prior probability density function. Assuming  $\mathbf{s}$  follows a Gaussian distribution, the Gaussian prior probability density function of  $\mathbf{s}$  is

$$p(\mathbf{s} \mid \boldsymbol{\delta}) = \text{CN}(\mathbf{s} \mid \mathbf{0}, \boldsymbol{\Lambda}) \quad (32)$$

In the equation,  $\boldsymbol{\Lambda} = \text{diag}(\boldsymbol{\delta})$ ,  $\boldsymbol{\delta} = [\delta_1 \ \delta_2 \ \dots \ \delta_{N+1}]^T$ , and the hyperparameter  $\boldsymbol{\delta}$  is further assumed to follow an independent Gamma distribution, i.e.,

$$p(\boldsymbol{\delta}) = \prod_{i=1}^{N+1} \Gamma(\delta_i; 1, \rho) \quad (33)$$

In the equation, the shape parameter is fixed at 1 to ensure a stronger shrinkage effect on the parameters, causing most of the estimated values to approach zero, thereby achieving sparsity.  $\rho$  is a small positive constraint, typically set to 0.01. The Gamma distribution is defined as  $\Gamma(x \mid a, b) = b^a / \Gamma(a) x^{a-1} \exp(-bx)$ .

Based on the assumptions of the above Bayesian hierarchical model, the joint distribution of the signal can be obtained:

$$p(\tilde{\mathbf{y}}, \mathbf{s}, \boldsymbol{\delta}) = p(\tilde{\mathbf{y}} \mid \mathbf{s}; \boldsymbol{\beta}) p(\mathbf{s} \mid \boldsymbol{\delta}) p(\boldsymbol{\delta}) \quad (34)$$

According to Bayesian theory, the maximum posterior probability of the parameters to be estimated can be obtained, but its calculation usually involves high-dimensional, complex integrals, which are difficult to solve. Therefore, we introduce variational inference to solve the problem of maximum posterior estimation.

In variational inference, the observed data are the array element received data  $\tilde{\mathbf{y}}$ , and the set of latent variables  $\boldsymbol{\xi} = \{\mathbf{s}, \boldsymbol{\delta}\}$  consists of the parameters to be estimated.

$$\ln p(\tilde{\mathbf{y}}) = L(q) + KL(q \parallel p) = \int q(\boldsymbol{\xi}) \ln \left\{ \frac{p(\tilde{\mathbf{y}}, \boldsymbol{\xi})}{q(\boldsymbol{\xi})} \right\} d\boldsymbol{\xi} - \int q(\boldsymbol{\xi}) \ln \left\{ \frac{p(\boldsymbol{\xi} \mid \tilde{\mathbf{y}})}{q(\boldsymbol{\xi})} \right\} d\boldsymbol{\xi} \quad (35)$$

In the equation,  $L(q)$  is the evidence lower bound (ELBO);  $KL(q \parallel p)$  is the KL divergence, which represents the degree of approximation between the probability distribution  $q$  and the posterior distribution  $p$ . The smaller the KL divergence, the higher the approximation. Variational inference aims to maximize the lower bound of  $\ln p(\tilde{\mathbf{y}})$  by finding the distribution of  $q(\boldsymbol{\xi})$  such that the lower bound of the ELBO is maximized. At this point,  $q(\boldsymbol{\xi})$  can be used to approximate the posterior probability distribution of the latent variables.

To maximize the lower bound, the probability distribution is decomposed as  $q(\boldsymbol{\xi}) = q(\mathbf{s})q(\boldsymbol{\delta})$ , and the general expression for the optimal approximate distribution is given.

$$\ln q_j^*(\xi_j) = E_{i \neq j}[\ln p(\tilde{\mathbf{y}}, \boldsymbol{\xi})] + \text{const} \quad (36)$$

In the equation,  $E_{i \neq j}[\cdot]$  is the conditional expectation of the parameter's approximate distribution  $q$  for  $\xi_i$ , conditioned on  $\xi_i, i \neq j$ .

### 3.3.1. Ignore the Terms Unrelated to the Signal

From Equation (36), the best approximate distribution of  $q(\xi)$  can be obtained.

$$\begin{aligned}\ln q^*(s) &= \langle \ln p(\tilde{y} | s, \beta) p(s | \delta) \rangle_{q(\delta)} \\ &= \left\langle -\frac{1}{2} (\tilde{y} - \tilde{\Phi}^H(\beta)s)^H (\tilde{y} - \tilde{\Phi}^H(\beta)s) - \frac{1}{2} s^H \delta s \right\rangle_{q(\delta)} \\ &= \frac{1}{2} \left\langle -s^H (\tilde{\Phi}^H(\beta) \tilde{\Phi}(\beta) + \delta) s + 2s^H \tilde{\Phi}^H(\beta) \tilde{y} \right\rangle_{q(\delta)}\end{aligned}\quad (37)$$

The approximate posterior of  $s$  follows a complex Gaussian distribution, expressed as

$$q^*(s) = \text{CN}(s | \mu, \Sigma) \quad (38)$$

where

$$\mu = \Sigma \tilde{\Phi}^H(\beta) \tilde{y} \quad (39)$$

$$\Sigma = \left( \tilde{\Phi}^H(\beta) \tilde{\Phi}(\beta) + \Lambda^{-1} \right)^{-1} \quad (40)$$

### 3.3.2. Ignore the Terms Unrelated to Precision

From Equation (36), the best approximate distribution of  $q(\xi)$  can be obtained.

$$\begin{aligned}\ln q^*(\delta) &= \langle \ln p(s | \delta) p(\delta) \rangle_{q(s)} \\ &= \left\langle \sum_{i=1}^{N+1} \ln \delta_i + \rho \delta_i - s^H \delta s \right\rangle_{q(s)} \\ &= \sum_{i=1}^{N+1} \ln \delta_i - (\langle d_i^2 \rangle + \rho) \delta_i\end{aligned}\quad (41)$$

Its approximate distribution is  $q^*(\delta) = \prod_i^{N+1} \Gamma(\delta_i | 1, \rho)$ , and the signal precision follows a multidimensional Gamma distribution. The superscript  $t$  represents the iteration number, and the iterative update expression for its precision is

$$\left( d_i^2 \right)^{(t+1)} = (\mu_i \mu_i^*)^{(t)} + \Sigma_{ii}^{(t)} \quad (42)$$

$$(\langle \delta_i \rangle)^{(t+1)} = \rho + \left( \langle d_i^2 \rangle \right)^{(t)} \quad (43)$$

In the equation,  $\mu_i^{(t)}$  represents the  $i$ -th element of the mean vector  $\mu^{(t)}$ , and  $\Sigma_{ii}^{(t)}$  represents the  $i$ -th diagonal element of the covariance matrix. Therefore, based on Equations (42) and (43), the update function  $f_i(\delta)$  for  $\delta_i$  can be obtained, i.e.,  $\delta_i^{(t+1)} = f_i(\delta_i^{(t)})$ , and its specific form is as follows:

$$f_i(\delta) = \frac{1}{\rho + \mu_i \mu_i^* + \Sigma_{ii}} \quad (44)$$

By iteratively updating Equations (39), (40), and (44) until convergence, the solutions for the relevant parameters can be obtained. However, due to the circular dependency of the hyperparameter  $\delta$  on  $\mu$  and  $\Sigma$ , this may lead to slower convergence of  $\delta$ , thus affecting the overall speed of the algorithm.

### 3.4. Fast Variational Bayesian Inference

To address the issue of parameter circular dependency in the iterative update, this paper incorporates the concept of Fast Variational Bayesian Inference (FVBI) during the parameter iteration phase to effectively accelerate the iterative process.

According to Equation (44), by repeatedly updating the hyperparameter  $\delta_i$ , an estimate sequence  $\{\delta_i^{[t]}\}_{t=1}^{\infty}$  can be obtained. If this update sequence converges, the limit must be a fixed point of this recursive relation:

$$\delta_i^* \in G_i = \{\delta_i \in \mathbb{R}_{\geq 0} : f_i(\delta_i) - \delta_i = 0\} \quad (45)$$

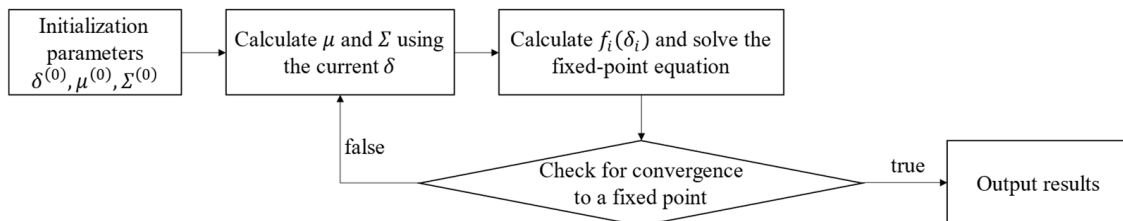
Thus, the convergent solution  $\delta_i^*$  of the hyperparameter  $\delta_i$  can be obtained by solving the equation  $f_i(\delta_i) - \delta_i = 0$ . According to reference [42], if the initial condition is  $\delta_i^{[0]} \geq 0$  and the first-order recursive relation  $f_i$  is strictly increasing, then when  $t \rightarrow \infty$ , the limit of  $\delta_i^{[t]}$  is as follows:

$$\lim_{t \rightarrow \infty} \delta_i^{[t]} = \begin{cases} \infty & \text{if } f_i(\delta_i^{[0]}) > \delta_i^{[0]} \text{ and } G_i^+ = \emptyset \\ \min G_i^+ & \text{if } f_i(\delta_i^{[0]}) > \delta_i^{[0]} \text{ and } G_i^+ \neq \emptyset \\ \max G_i^- & \text{if } f_i(\delta_i^{[0]}) \leq \delta_i^{[0]} \end{cases} \quad (46)$$

where  $G_i^+$  represents the fixed point greater than  $\delta_i^{[0]}$ ,  $G_i^-$  represents the fixed point less than or equal to  $\delta_i^{[0]}$ , and  $\emptyset$  represents the empty set.

Since  $d_i^2$  is a strictly decreasing monotonic function with respect to  $\delta_i$ , the update function  $f_i(\delta_i)$  for the hyperparameter  $\delta_i$  is strictly increasing, so the above conclusion can be applied.

In essence, the proposed Fast Variational Bayesian Inference reformulates the conventional slow, recursive variational EM update into a root-finding problem of a strictly increasing function, as illustrated in Figure 1, thereby enabling faster fixed-point iteration and improving computational efficiency under mild initialization conditions.



**Figure 1.** Fast Variational Bayesian Inference process diagram.

### 3.5. Grid Evolution

To reduce computational load, the initial grid is kept coarse. However, overly coarse grids may lead to multiple DOAs falling within the same interval, causing traditional methods to fail. This paper proposes a novel grid update method with two components: a fission process to ensure one DOA per grid interval, and an update process to approximate the true DOA. Alternating iterations of these processes refine the grid and improve estimation accuracy. Figure 2 shows a schematic diagram of the grid evolution process.

#### 3.5.1. Update Process

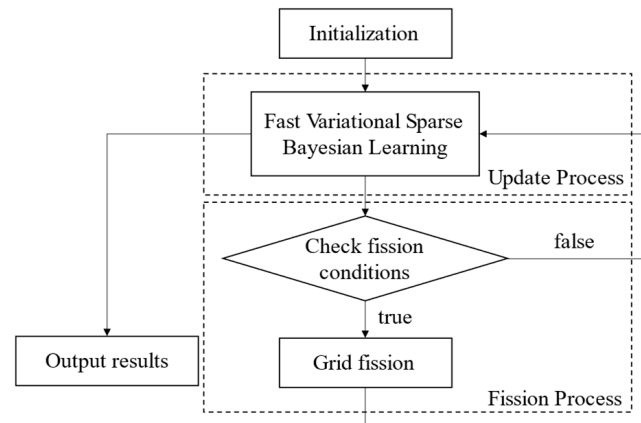
The update of grid refinement  $\beta$  is completed during the maximization phase of the Expectation Maximization (EM) algorithm, i.e.,  $\beta = \arg\max \langle \ln p(\tilde{\mathbf{y}} | \mathbf{s}, \beta) \rangle_{q(\mathbf{s})}$ .

$$\begin{aligned} \langle \ln p(\tilde{\mathbf{y}} | \mathbf{s}, \beta) \rangle_{q(\mathbf{s})} &= - \left\langle (\tilde{\mathbf{y}} - \tilde{\Phi}(\beta)\mathbf{s})^H (\tilde{\mathbf{y}} - \tilde{\Phi}(\beta)\mathbf{s}) \right\rangle_{q(\mathbf{s})} \\ &= - \|\tilde{\mathbf{y}} - \tilde{\Phi}(\beta)\boldsymbol{\mu}\|_2^2 - \text{tr} \left( \tilde{\Phi}(\beta) \Sigma \tilde{\Phi}^H(\beta) \right) + \text{const} \\ &= -\boldsymbol{\beta}^T \mathbf{P} \boldsymbol{\beta} + 2\mathbf{v}^T \boldsymbol{\beta} + \text{const} \end{aligned} \quad (47)$$

where

$$\mathbf{P} = \text{Re} \left\{ (\mathbf{B}_U)^T \mathbf{B}_U^* \odot \underline{\underline{\mu}} \underline{\underline{\mu}}^H + \underline{\underline{\Sigma}} \right\} \quad (48)$$

$$\mathbf{v} = \text{Re} \left\{ \text{diag}(\underline{\underline{\mu}}^*) (\mathbf{B}_U)^H (\hat{\mathbf{y}} - \mathbf{A}_U \underline{\underline{\mu}} - \underline{\underline{\mu}}_0 \mathbf{1}_U) \right\} - \text{Re} \left\{ \text{diag}(\mathbf{B}_U)^H \mathbf{A}_U \underline{\underline{\Sigma}} + \text{diag}(\gamma (\mathbf{B}_U)^T (\mathbf{1}_U)^*) \right\} \quad (49)$$



**Figure 2.** Grid evolution process diagram.

In the equation,  $\mathbf{A}_U = \mathbf{U}^H \hat{\mathbf{W}}^{-1/2} \mathbf{G}(\bar{\theta})$ ;  $\mathbf{B}_U = \mathbf{U}^H \hat{\mathbf{W}}^{-1/2} \mathbf{B}$ ;  $\mathbf{1}_U = \mathbf{U}^H \hat{\mathbf{W}}^{-1/2} \text{vec}\{\mathbf{I}_M\}$ ;  $\|\cdot\|_2$  denotes the 2-norm of the matrix;  $\text{tr}(\cdot)$  denotes the trace of the matrix;  $\odot$  represents the Hadamard product;  $\text{const}$  is a constant;  $\underline{\underline{\mu}}$  is the vector composed of the first  $N$  elements of  $\underline{\underline{\mu}}$ ;  $\underline{\underline{\mu}}_0$  is the last element of  $\underline{\underline{\mu}}$ ;  $\underline{\underline{\Sigma}}$  is the matrix composed of the first  $N$  rows and columns of  $\underline{\underline{\Sigma}}$ ; and  $\gamma$  is the vector composed of the first  $N$  columns of the  $N + 1$ -th row of  $\underline{\underline{\Sigma}}$ .

Differentiating Equation (44) and setting it to zero yields the update expression for  $\underline{\underline{\beta}}$ :

$$\underline{\underline{\beta}} = \mathbf{P}^{-1} \mathbf{v} \quad (50)$$

However, this equation requires that  $\mathbf{P}$  is invertible. In practical scenarios, especially under low SNR or when the dictionary is highly coherent, the matrix  $\mathbf{P}$  may become ill-conditioned or even singular. To ensure numerical stability, we adopt a regularized inversion strategy by adding a small damping term:

$$\underline{\underline{\beta}} = (\mathbf{P} + \varepsilon \mathbf{I})^{-1} \mathbf{v} \quad (51)$$

where  $\varepsilon$  is a small positive constant (e.g.,  $\varepsilon = 10^{-6}$ ) that ensures invertibility without significantly altering the solution. Alternatively, if the diagonal of  $\mathbf{P}$  is dominant, the simplified component-wise update

$$\beta_i = \frac{v_i}{P_{ii}} \quad i = 1, 2, \dots, N \quad (52)$$

can be used as a computationally efficient approximation.

The original off-grid method updates the measurement matrix  $\Phi(\underline{\underline{\beta}})$  using the obtained  $\underline{\underline{\beta}}$  and proceeds to the next iteration. While the updated matrix offers a better approximation of the true steering matrix, the distance between the actual  $\bar{\theta}_k$  and the nearest grid point  $\bar{\theta}_k$  remains unchanged. Thus, the modeling error from the off-grid gap in Equation (9) persists. To resolve this, we can treat the grid point positions as adjustable parameters and directly update them.

$$\bar{\theta}_{i,new} = \bar{\theta}_i + \beta_i \quad (53)$$

The overcomplete dictionary set is updated with the refined grid points. After several iterations, the updated grid points will converge towards the true DOA. In practice, candidates are selected based on the refined grid positions. If  $\bar{\theta}_i + \beta_i \in [\bar{\theta}_{i-1}, \bar{\theta}_{i+1}]$ , the new  $\bar{\theta}_{i,new}$  is accepted; otherwise, the original grid point is kept. The process continues until the stopping condition  $\frac{\|\mu^{(i+1)} - \mu^{(i)}\|_2}{\|\mu^{(i+1)}\|_2} > \tau$  or  $i < \text{Iter}_{\max}$  is met, where  $\tau$  is the error threshold,  $i$  is the iteration count, and  $\text{Iter}_{\max}$  is the maximum number of iterations.

### 3.5.2. Fission Process

#### 1. Fission Strategy

If the energy of the  $i$ -th row of  $\mathbf{s}$  is small, indicating the source is not at that position, the corresponding  $\theta_i$  remains unchanged. Therefore, not all grid points need to be updated in each iteration. The maximum number of targets resolvable by an array with  $M$  elements is  $M - 1$ . Thus, the grid points corresponding to the top  $M - 1$  largest energy values of  $\mathbf{s}_i$  are selected for updating. The average power of  $\mathbf{s}_i$  is as follows:

$$Q(i) = \sqrt{\sum_{t=1}^T |\mu_{ti}|^2}, i = 1, 2, \dots, N \quad (54)$$

The larger the  $Q(i)$ , the higher the likelihood of a DOA in that direction. Thus, only grid intervals with large  $Q(i)$  values require fission. We select the top  $M - 1$  largest  $Q(i)$  values as candidate grid points for fission. However, when a signal source has large energy, its side lobe energy may surpass the main lobe energy of other sources, potentially leading to the loss of grid points for smaller main lobes. Therefore, the accurate method is to choose the angle positions corresponding to local maxima in the top  $M - 1$  largest average power  $Q(i)$  values as fission grid points.

Fission generates two new grid points. When the  $i$ -th grid point is selected for fission, the new grid points are placed at the midpoints between the  $r(i - 1)$  and  $r(i)$  grid intervals. This process continues, progressively approaching the true DOA position.

#### 2. Iterative Update

When a grid point undergoes fission, new grid points are generated. According to the off-grid model (15), rows corresponding to  $\mathbf{s}, \beta$ , and the array manifold matrix for this grid point should be added for the next iteration. The grid interval, RRR, also needs to be recalculated as  $\mathbf{r} = \{\bar{\theta}_2 - \bar{\theta}_1, \bar{\theta}_3 - \bar{\theta}_2, \dots, \bar{\theta}_N - \bar{\theta}_{N-1}\}$ .

#### 3. Fission Termination Condition

As the fission process progresses, the number of grid points increases, raising computational complexity. Excessive fission depth can degrade estimation accuracy. Fission stops when the left and right intervals of a grid point III fall below a threshold, i.e.,  $r(i - 1) < \theta_\epsilon$  and  $r(i) < \theta_\epsilon$ . The fission termination condition  $\theta_\epsilon$  impacts DOA performance.

## 4. Simulation Results

This section evaluates the performance of the proposed OGFVBI method under wideband signal scenarios and compares it with methods such as MUSIC [8], RootSBL [33], and realVBI [20]. The Cramér-Rao Bound (CRB) is adopted as the theoretical benchmark for estimation accuracy variance [43].

The signal model is based on wideband signal simulation, considering multiple signal sources. The signals are modulated using 16-QAM and transmitted over an OFDM signal, with the following parameters: OFDM subcarrier size is 1024, and the number of symbols is 16. The experimental setup employs a uniform linear array (ULA) with  $M = 20$  elements

and an inter-element spacing of half a wavelength, covering an angular range from  $-90^\circ$  to  $90^\circ$ . The initial grid interval is set to  $1^\circ$ . The number of Monte Carlo simulation trials was set to  $N_{MC} = 100$ . In all relevant plots, shaded regions around the curves represent  $\pm 1$  standard deviation calculated over  $N_{MC}$  Monte Carlo trials, reflecting the robustness and statistical stability of each algorithm. All simulations were conducted in MATLAB R2022a on a computer equipped with an Intel i7-10710U CPU and 16 GB of RAM.

The CRB is calculated for uncorrelated narrowband sources under a known SNR. Let  $\mathbf{A} \in \mathbb{C}^{M \times K}$  denote the array manifold matrix and  $\mathbf{D} \in \mathbb{C}^{M \times K}$  its partial derivative with respect to DOA angles. The Fisher Information Matrix (FIM) is computed as follows:

$$\mathbf{F} = \frac{2N}{\sigma^2} \text{Re} \left( \mathbf{D}^H \left( \mathbf{I} - \mathbf{A} (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H \right) \mathbf{D} \right) \quad (55)$$

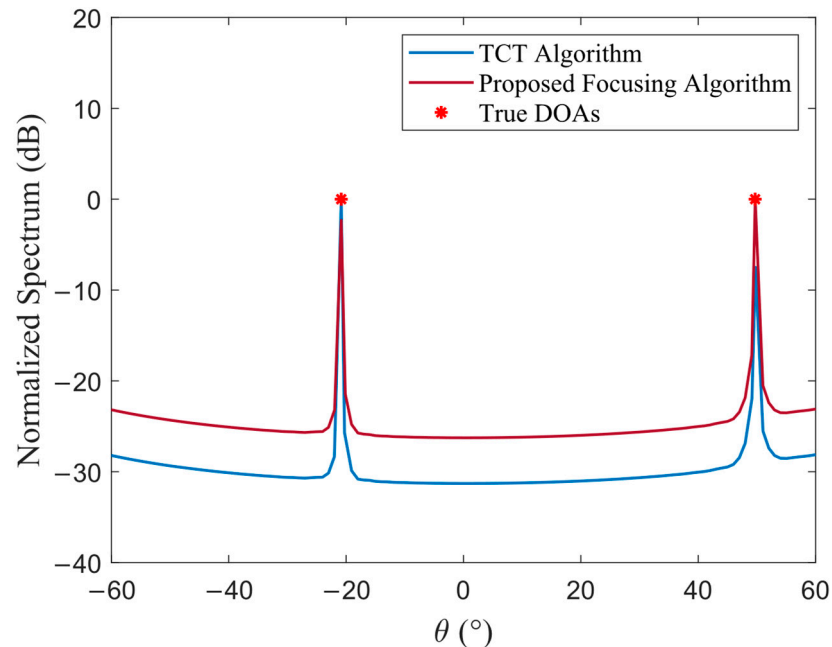
where  $\sigma^2$  is the noise power and  $N$  is the number of snapshots. The CRB for each DOA parameter  $\theta_k$  is obtained from the diagonal entries of the inverse FIM and converted to degrees:

$$\text{CRB}(\theta_k) = \sqrt{[\mathbf{F}^{-1}]_{k,k}} \cdot \frac{180}{\pi} \quad (56)$$

#### 4.1. Focusing Algorithm Estimation Performance Analysis

To further validate the effectiveness of the proposed focusing method, we compare it against the classical TCT-based focusing technique in terms of resulting spatial spectrum sharpness. Figure 3 shows the spatial power spectrum obtained using both methods under the same signal scenario.

The spatial spectra of the proposed algorithm and TCT show close agreement, indicating comparable performance. This equivalence occurs because both methods produce identical eigenvectors in the focused covariance matrix. However, the proposed approach achieves this with lower computational complexity.

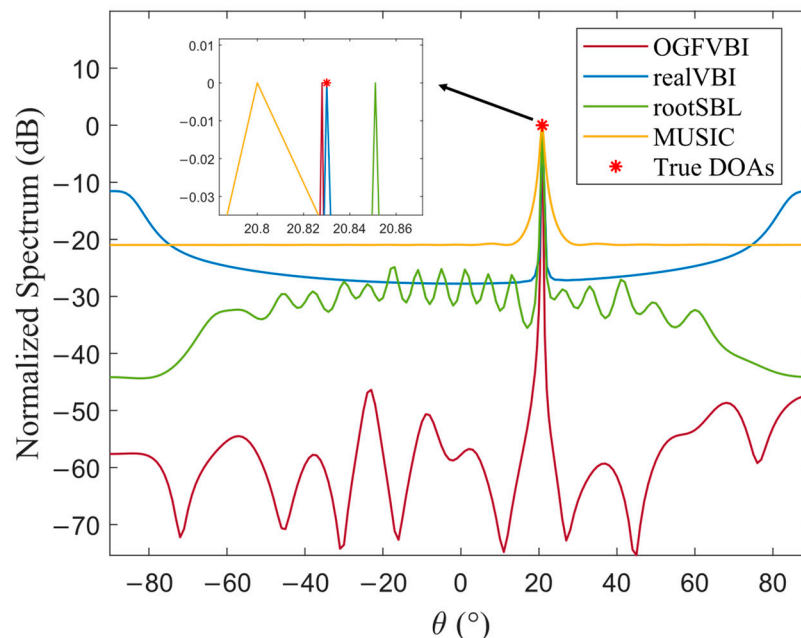


**Figure 3.** Spatial power spectrum: proposed focusing algorithm vs. TCT algorithm ( $SNR = 10$  dB,  $SNAP = 200$ ).

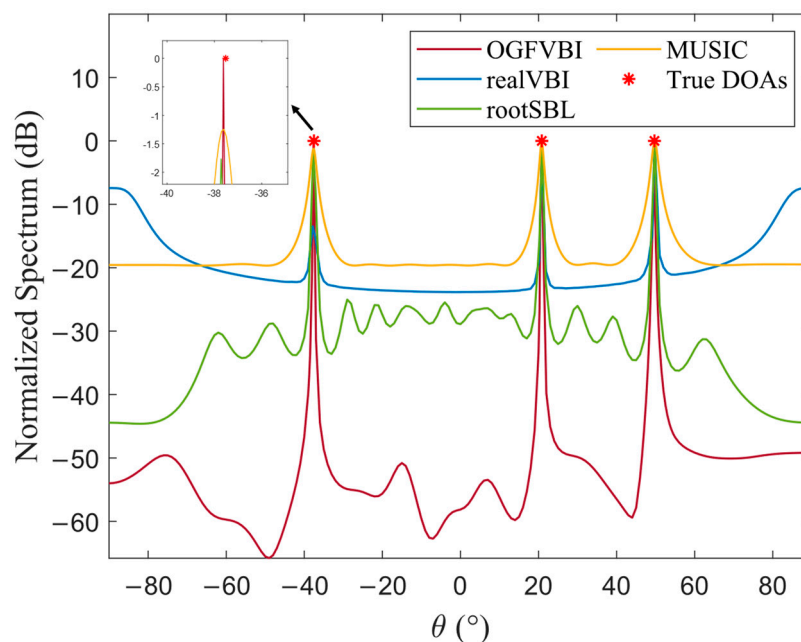


#### 4.2. Spatial Spectrum Analysis

In the experiments, the DOA of a single source is set to  $20.83^\circ$ , and the corresponding normalized spatial spectrum is shown in Figure 4. For the multiple-source scenario, the DOAs are set to  $-37.52^\circ$ ,  $20.83^\circ$ , and  $49.69^\circ$ , and the resulting normalized spatial spectrum is shown in Figure 5.



**Figure 4.** Normalized spatial spectrum of a single signal source ( $SNR = 10$  dB,  $SNAP = 200$ ).



**Figure 5.** Normalized spatial spectrum of multiple signal sources ( $SNR = 10$  dB,  $SNAP = 200$ ).

Figures 4 and 5, respectively, illustrate the normalized spatial spectrum comparisons of different DOA estimation algorithms under single-source and multi-source scenarios, clearly reflecting the performance differences in terms of localization accuracy and resolution capability. Although all four methods produce spectral peaks that approximately align with the true source locations, the zoomed-in views reveal that the proposed method consistently achieves precise alignment with the true DOAs in both scenarios. The spectral

peaks are sharp, with concentrated main lobes and strong sidelobe suppression, and the overall spectrum is smooth and stable—demonstrating superior estimation accuracy and spatial resolution.

In the single-source case shown in Figure 4, the proposed method remains precisely aligned with the true DOA even in the zoomed-in view, indicating its high sensitivity to small angular variations. In contrast, the MUSIC and RootSBL methods exhibit noticeable deviations. Although the realVBI method aligns with the main peak, it suffers from significant sidelobe fluctuations and contains several false peaks, which may lead to misjudgments.

Figure 5 further validates the robustness of the proposed method in multi-source environments. It clearly resolves all three sources, while other algorithms, such as MUSIC and realVBI, exhibit insufficient resolution due to high sidelobes or shifted main peaks.

These results demonstrate that the proposed method maintains high accuracy and resolution even under challenging conditions such as limited snapshots, low SNR, and coarse grids. This advantage stems from the adoption of Fast Variational Bayesian inference combined with a grid adaptation mechanism, which effectively overcomes the performance bottlenecks of traditional methods caused by grid mismatch or slow hyperparameter convergence.

#### 4.3. Resolution Analysis

To determine whether the estimation method can successfully resolve two closely spaced signal sources, the following two conditions must be satisfied:

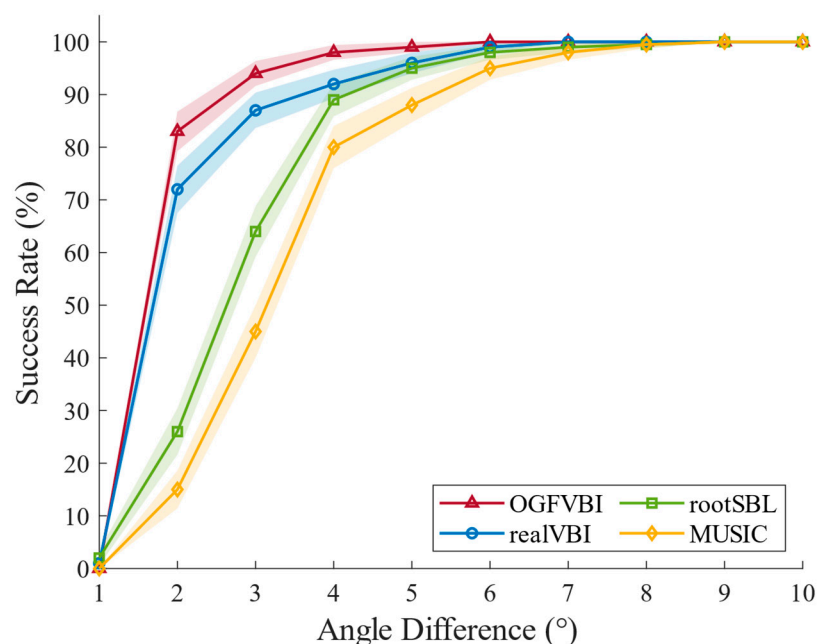
$$\begin{cases} \max\{|\bar{\theta}_1 - \theta_1|, |\bar{\theta}_2 - \theta_2|\} < \theta_c \\ \text{Success Rate} = N_{\text{success}} / N_{\text{MC}} \end{cases} \quad (57)$$

in the equation,  $\bar{\theta}_1$  and  $\bar{\theta}_2$  represent the estimated values corresponding to the true angles  $\theta_1$  and  $\theta_2$ , respectively.  $\theta_c$  is the set angle error threshold, which is set to  $0.5^\circ$  here.  $N_{\text{success}}$  denotes the number of successful resolutions in  $N_{\text{MC}}$  Monte Carlo trials where both incident sources are correctly resolved within the angular error threshold.

Consider two closely spaced, uncorrelated signals arriving at the array from directions  $\theta_1$  and  $\theta_2 = \theta_1 + \Delta$ , respectively, with  $\theta_1$  fixed at  $20.8^\circ$ . The angular separation  $\Delta$  is varied from  $1^\circ$  to  $10^\circ$  in  $1^\circ$  increments. The resulting resolution success rate as a function of angular separation is plotted in Figure 6.

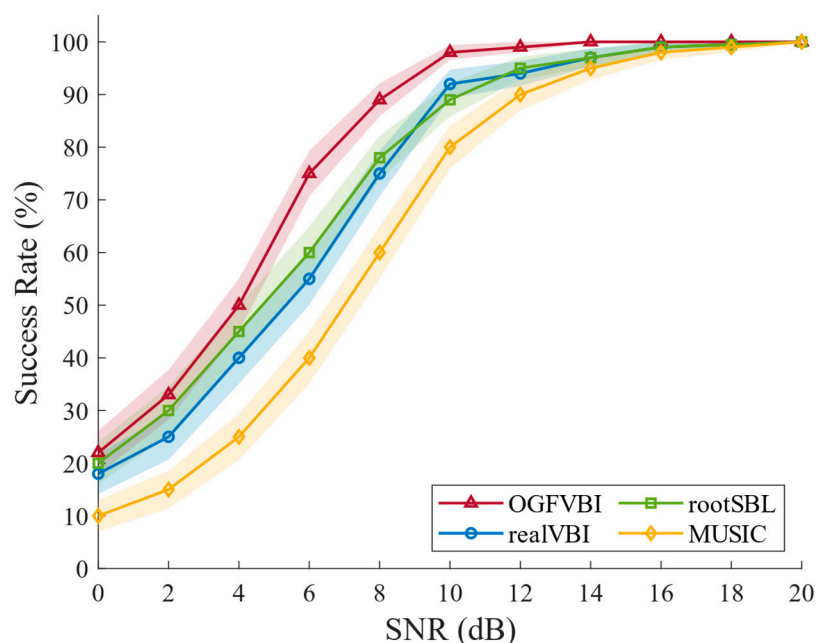
Figure 6 illustrates the trend of resolution success rate versus angular separation for four different algorithms under varying angle differences between two signal sources. For different incident signal grid spacings, the method proposed in this paper achieves a resolution success rate close to 100% when the angle interval is  $4^\circ$ . In contrast, the realVBI and RootSBL methods require a grid spacing of more than  $6^\circ$  to achieve a 100% resolution success rate. The MUSIC method, under the same conditions, requires the incident signal grid spacing to be increased to  $8^\circ$  to successfully resolve. It can be observed that the method proposed in this paper consistently maintains a significant performance advantage across the entire angle interval range, particularly in small angle intervals, where its resolution success rate is far higher than that of realVBI, RootSBL, and MUSIC, making it particularly suitable for complex real-world environments with low SNR and small angle differences between sources. The reason for this is that the method introduces a dynamic grid fission and update mechanism, allowing the algorithm to progressively refine the grid and approach the true signal source position even when starting with a coarse grid, significantly improving resolution. In contrast, traditional methods such as MUSIC heavily rely on the orthogonality of signal subspaces and a high number of snapshots, making them extremely sensitive to small angle differences and highly correlated signals, which can

lead to subspace leakage and resolution failure. Although RootSBL and realVBI somewhat address the grid discretization issue, they still depend on fixed grids, which may cause aliasing or spectral peak merging when dealing with closely spaced sources.



**Figure 6.** Resolution success rate (%) as a function of angular separation ( $^{\circ}$ ) ( $SNR = 10$  dB,  $SNAP = 200$ ).

Consider two closely spaced, uncorrelated signals impinging on the array from directions  $20.8^{\circ}$  and  $24.8^{\circ}$ , respectively. The resulting resolution success rate as a function of SNR is shown in Figure 7.



**Figure 7.** Resolution success rate (%) versus SNR (dB) ( $\Delta = 4^{\circ}$ ,  $SNAP = 200$ ).

Figure 7 shows the resolution success rate variation of four DOA estimation algorithms under different SNR conditions. It is clearly observed that as the SNR increases, the resolution capability of all algorithms improves. However, the method proposed in this

paper consistently outperforms the others across the entire SNR range, especially showing significantly higher resolution success rates in the low SNR segment.

It can be seen that even under low SNR conditions, where the signal is severely overwhelmed by noise, the proposed method can still reliably extract sparse signal features and gradually approach the true DOA position. In contrast, although the realVBI and RootSBL methods are also based on Bayesian modeling, they have a stronger dependence on hyperparameters and do not address the grid adaptation issue. As a result, they are prone to getting trapped in erroneous local estimates and fail to resolve correctly at low SNR. As the SNR increases further, the resolution of all methods gradually improves. However, the proposed method still maintains close to 100% success rate in high SNR conditions, demonstrating its excellent robustness and resolution ability across the entire SNR range. This shows that the method performs exceptionally well in environments with strong interference and is well-suited for complex scenarios with low SNR and closely spaced signal sources.

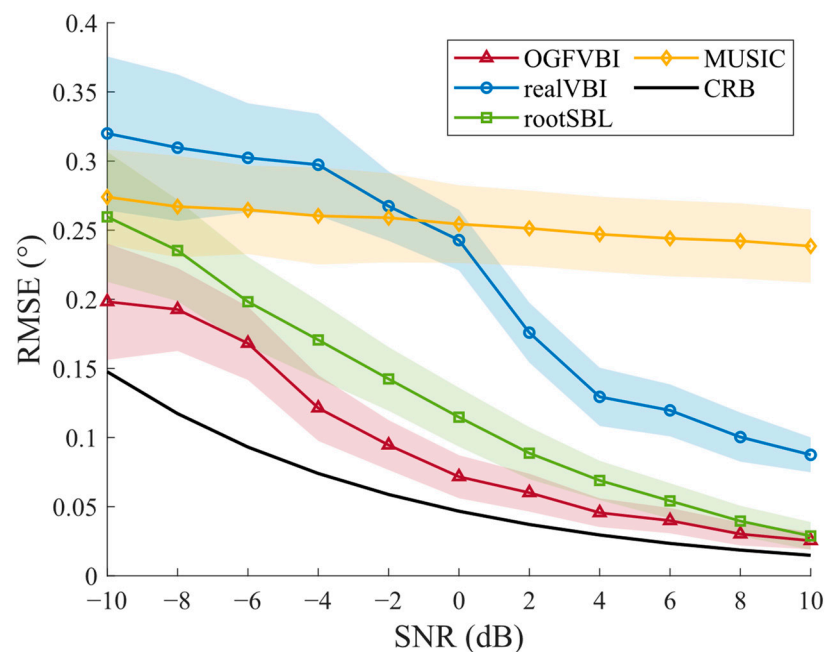
#### 4.4. Estimation Accuracy Analysis

The estimation errors of different methods are evaluated using the Root Mean Square Error (RMSE). The RMSE of the estimated directions of  $K$  incident signals is defined as follows:

$$\text{RMSE} = \sqrt{\frac{\sum_{k=1}^K \sum_{n_{\text{MC}}=1}^{N_{\text{MC}}} (\hat{\theta}_{kn_{\text{MC}}} - \hat{\theta}_k)^2}{K \times N_{\text{MC}}}} \quad (58)$$

where  $\hat{\theta}_{kn_{\text{MC}}}$  denotes the estimated direction of the  $k$ -th signal in the  $n_{\text{MC}}$ -th Monte Carlo simulation.

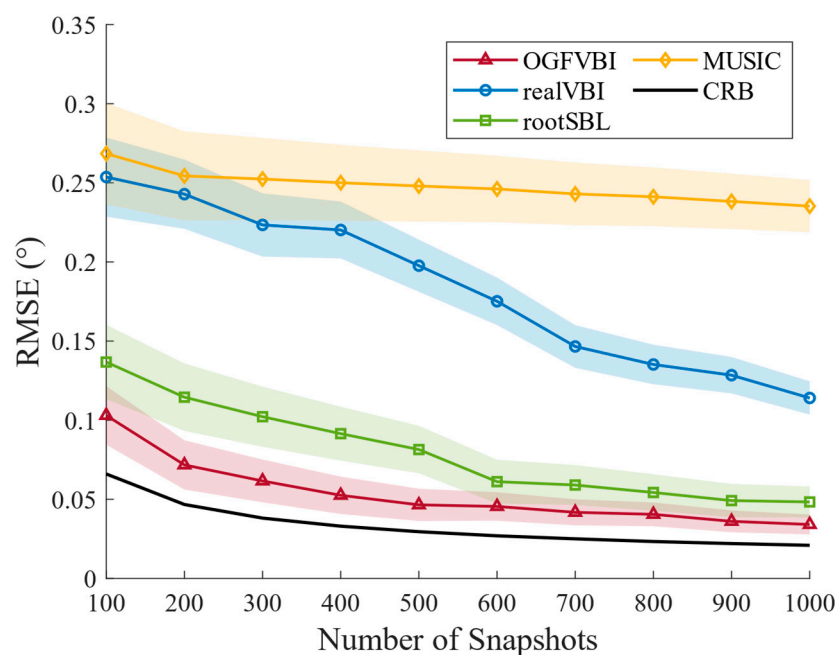
In the experiment, three signal sources are set at directions of  $-37.5^\circ$ ,  $20.8^\circ$ , and  $49.6^\circ$ . The RMSEs of different methods are calculated under varying SNR conditions, and the results are shown in Figure 8. It can be seen that the proposed method consistently maintains the lowest estimation error across the entire SNR range and gradually approaches the CRB curve as the SNR increases.



**Figure 8.** RMSE (°) vs. SNR (dB) comparisons of the various methods ( $SNAP = 200$ ).

For example, at an SNR of 0 dB, the RMSE of the proposed method is 0.0717, which improves by 70.5% compared to the 0.2428 of the realVBI method, 37.6% compared to the 0.1147 of the RootSBL method, and 71.8% compared to the 0.2544 of the MUSIC method. The standard deviation of the proposed method is also lower across all SNR levels, indicating more consistent estimation results. These results show that the proposed method quickly converges to the true DOA under low SNR and effectively avoids hyperparameter dependency, demonstrating strong robustness and high accuracy. In contrast, realVBI converges more slowly and often gets trapped in local minima, RootSBL suffers from bias in noisy conditions, and MUSIC is highly sensitive to covariance estimation errors, causing sharp performance drops at low SNR.

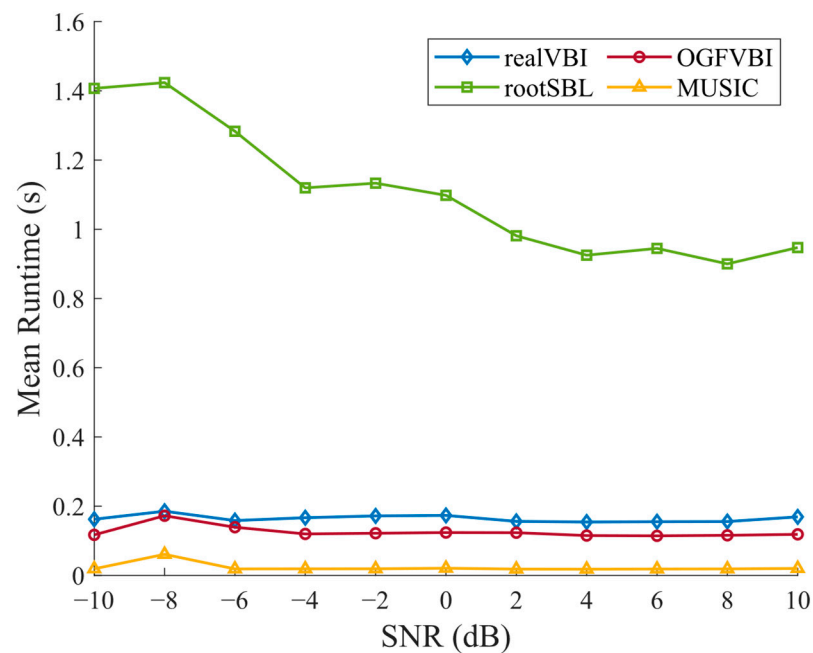
With all other experimental parameters held constant, Figure 9 shows the RMSE of different methods under varying snapshot numbers. The estimation errors of all methods decrease as the number of snapshots increases. However, the proposed method consistently achieves the lowest RMSE with smaller standard deviation, indicating more stable and reliable estimates. When snapshots reach 1000, the proposed method's RMSE is 0.0341, improving by 85.5% compared to MUSIC (0.2353), 70.1% compared to realVBI (0.1140), and 29.4% compared to RootSBL (0.0483). These results demonstrate that the proposed method efficiently uses limited snapshots for accurate DOA estimation, showing strong robustness and practical value.



**Figure 9.** RMSE (°) vs. number of snapshots comparisons of the various methods (SNR = 0 dB).

#### 4.5. Efficiency Analysis: Runtime and Convergence

Considering the impact of SNR on computational efficiency, we keep other conditions unchanged and vary the SNR from  $-10$  to  $10$  in steps of 2. The runtime of different methods under varying SNRs is recorded, as shown in Figure 10. It can be seen that the runtimes of OGFVBI, realVBI, and MUSIC remain relatively stable across different SNR levels, demonstrating good consistency. In contrast, the runtime of the RootSBL method is significantly higher than the others and slightly decreases as the SNR increases, showing larger overall fluctuations.



**Figure 10.** Mean runtime(s) vs. SNR (dB) comparisons of the various methods ( $SNAP = 200$ ).

Further analysis of the average iteration counts, runtimes, and their corresponding standard deviations for the methods is presented in Table 2.

**Table 2.** Computational efficiency and convergence metrics of compared methods ( $SNR = 0$  dB,  $SNAP = 200$ ).

Method	Runtime (s)	Standard Deviation (s)	Iterations
OGFVBI	0.1238	0.1255	18
RealVBI	0.1732	0.1643	35
RootSBL	1.0981	1.1058	52
MUSIC	0.0206	0.0228	—

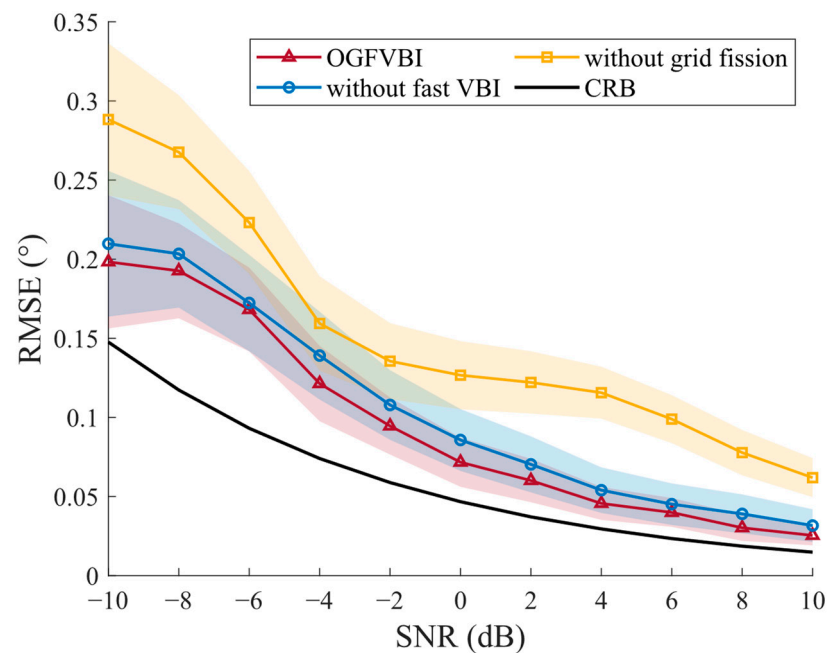
Compared to RootSBL and realVBI, our proposed OGFVBI reduces the average runtime by approximately 88.7% and 28.5%, respectively. It also converges with about 65.4% fewer iterations than RootSBL and 48.6% fewer iterations than realVBI. These results confirm the significant computational efficiency gained through the fast variational update and adaptive grid fission strategy used in OGFVBI. Moreover, the relatively low standard deviation in runtime indicates that OGFVBI maintains stable performance across different trials.

#### 4.6. Ablation Study

To evaluate the contributions of the key modules in the proposed OGFVBI algorithm, we performed ablation experiments by separately removing the Fast VBI module and the grid fission module. The results on RMSE, standard deviation, runtime, and runtime variability are jointly analyzed to illustrate their impact.

As shown in Figure 11 and Table 3, the complete OGFVBI model achieves the lowest RMSE values across all SNR levels, with results closely approaching the theoretical CRB bound, demonstrating high estimation accuracy and stability (low standard deviation). Correspondingly, it also attains the shortest average runtime of 0.1238 s with a standard deviation of 0.1255 s, indicating efficient and stable computations.





**Figure 11.** Ablation study of OGFVBI: impact of grid fission and Fast VBI on estimation accuracy ( $SNAP = 200$ ).

**Table 3.** Computational performance comparison of OGFVBI and ablated versions ( $SNR = 0$  dB,  $SNAP = 200$ ).

Method	Runtime (s)	Standard Deviation (s)	Iterations
OGFVBI	0.1238	0.1255	18
Without Fast VBI	0.1716	0.1828	26
Without grid fission	0.1378	0.1468	21

When the Fast VBI module is removed, the RMSE increases slightly across all SNRs, and the standard deviation also rises, indicating degraded accuracy and less stable convergence. Meanwhile, the average runtime significantly increases to 0.1716 s (about 39% longer), with higher runtime variability (standard deviation 0.1828 s). This confirms that the fixed-point acceleration scheme in Fast VBI effectively reduces iteration counts, stabilizes the hyperparameter update process, and thus improves both convergence speed and computational efficiency.

Removing the grid fission module causes a more pronounced increase in RMSE, especially at low SNR values, and larger variability in estimation error. Although the runtime increase is moderate (to 0.1378 s) with a slight rise in standard deviation (0.1468 s), the loss of adaptive grid refinement leads to higher grid mismatch errors, reducing the estimation resolution and accuracy.

In summary, the Fast VBI module primarily enhances convergence speed and runtime stability, while the grid fission module plays a crucial role in reducing grid mismatch and improving estimation precision. Both modules jointly enable OGFVBI to achieve its superior performance in wideband DOA estimation tasks.

#### 4.7. Limitations and Future Extensions

Despite the advantages of the proposed OGFVBI method in terms of accuracy, robustness, and computational efficiency, several limitations and open challenges remain to be addressed in future work.



First, the performance of the algorithm is influenced by the initial grid resolution used to discretize the angular domain. Although the proposed dual-stage grid fission strategy can adaptively refine the grid and alleviate the mismatch between true angles and grid points, the method may still exhibit sensitivity to poor initial settings, especially when the angular search space is large or densely populated with sources. Future improvements could focus on fully gridless approaches or grid initialization strategies guided by prior knowledge.

Second, the algorithm assumes that the number of sources is known and that the sources are mutually uncorrelated. In practical applications, however, these assumptions may not hold. The presence of coherent or closely spaced sources may degrade the performance of subspace-based focusing and reduce the effectiveness of the sparse Bayesian recovery. Similarly, when the number of active sources is over- or under-estimated, the variational updates may either converge slowly or produce biased solutions. An adaptive model selection mechanism or automatic relevance determination could help address such model mismatch issues.

Third, the method is derived under the assumption of white Gaussian noise. In real-world systems, noise may exhibit impulsive, colored, or non-Gaussian behavior due to hardware imperfections, environmental interference, or calibration errors. Under such conditions, the covariance estimation and the posterior inference steps may become biased. Incorporating robust noise modeling or outlier mitigation strategies is a potential direction to improve the applicability of the method.

Beyond these current limitations, the method also has room for extension and generalization. In its present form, the proposed algorithm supports one-dimensional DOA estimation using ULA. In many practical scenarios, however, two-dimensional or even three-dimensional DOA estimation is required—such as in MIMO radar, acoustic sensing, or mmWave vehicular systems. Extending OGFVBI to handle joint azimuth and elevation estimation using uniform rectangular arrays (URA) or uniform circular arrays (UCA) would significantly enhance its spatial resolution and application scope.

Additionally, the method currently assumes far-field signal propagation where planar wavefronts are valid. In near-field scenarios—such as indoor localization or short-range radar—the signal wavefronts are spherical, and the array manifold becomes range-dependent. Adapting the OGFVBI framework to near-field models would allow simultaneous estimation of both direction and distance, enabling integrated localization and tracking applications with a single station.

Finally, from a system implementation perspective, the low complexity and fast convergence properties of OGFVBI make it attractive for real-time embedded applications. Potential deployment scenarios include joint communication and sensing (ISAC) systems, autonomous navigation platforms, and mobile edge computing for intelligent vehicles. Future work will explore how to integrate the algorithm with hardware constraints, dynamic tracking, and low-latency signal processing requirements.

## 5. Conclusions

This paper addresses the issues of low estimation accuracy, slow convergence, and high computational complexity in existing DOA estimation algorithms under conditions of low SNR, few snapshots, and coarse grids. We propose an off-grid DOA estimation method based on Fast Variational Bayesian Inference (OGFVBI). The method is based on the framework of variational Bayesian inference and approximates the posterior distribution by maximizing the evidence lower bound. A fast fixed-point detection mechanism is introduced to efficiently update hyperparameters, significantly accelerating the algorithm's convergence and reducing its computational complexity. At the same time, grid fission

and adaptive updating strategies are combined to evolve the grid to gradually approach the true source direction, effectively alleviating the grid mismatch problem in traditional off-grid models. In addition, an improved frequency focusing method is designed, which approximates broadband signals as narrowband models, further expanding the algorithm's applicability in broadband conditions. Simulation results show that the proposed method outperforms existing mainstream methods in estimation accuracy, resolution, and computational efficiency, especially in harsh environments with low SNR and fewer snapshots, demonstrating significant advantages and strong practical applicability and potential for broader use.

**Author Contributions:** Conceptualization, Z.D. and E.H.; methodology, X.T., Z.D. and E.H.; software, Y.C.; validation, X.T. and Y.C.; formal analysis, Y.C.; investigation, X.T.; resources, E.H.; data curation, Y.C. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the National Key Research and Development Program of China (grant number: 2022YFB3904702) and the National Natural Science Foundation of China under Grant: 6220020330.

**Data Availability Statement:** The data presented in this study are available within the article.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Molaei, A.M.; Zakeri, B.; Andargoli, S.M.H.; Abbasi, M.A.B.; Fusco, V.; Yurduseven, O. A comprehensive review of direction-of-arrival estimation and localization approaches in mixed-field sources scenario. *IEEE Access* **2024**, *2*, 65883–65918. [\[CrossRef\]](#)
2. Stutzman, W.L.; Thiele, G. *Antenna Theory and Design*; John Wiley & Sons: Hoboken, NJ, USA, 2012.
3. Yang, Y.; Lu, R.; Gao, L.; Gong, S. 4.5 GHz lithium niobate MEMS filters with 10% fractional bandwidth for 5G front-ends. *J. Microelectromechanical Syst.* **2019**, *28*, 575–577. [\[CrossRef\]](#)
4. Stoica, P.; Sharman, K.C. Maximum likelihood methods for direction-of-arrival estimation. *IEEE Trans. Acoust. Speech Signal Process.* **1990**, *38*, 1132–1143. [\[CrossRef\]](#)
5. El-Khamy, S.E.; El-Shazly, A.M.; Eltrass, A.S. A new computationally efficient approach for high-resolution DOA estimation of wideband signals using compressive sensing. *IEEE* **2023**, *1*, 74–81.
6. Hung, H.; Kaveh, M. Focussing matrices for coherent signal-subspace processing. *IEEE Trans-Actions Acoust. Speech Signal Process.* **1988**, *36*, 1272–1281. [\[CrossRef\]](#)
7. Valaee, S.; Kabal, P. Wideband array processing using a two-sided correlation transformation. *IEEE Trans. Signal Process.* **1995**, *43*, 160–172. [\[CrossRef\]](#) [\[PubMed\]](#)
8. Schmidt, R. Multiple emitter location and signal parameter estimation. *IEEE Trans. Signal Process.* **1986**, *34*, 276–280. [\[CrossRef\]](#)
9. Wu, C.; Ye, C. DOA estimation for unfolded coprime arrays: Successive-MUSIC algorithm. *IOP Conf. Ser. Mater. Sci. Eng.* **2020**, *719*, 012035–012041. [\[CrossRef\]](#)
10. Zhang, W.; Han, Y.; Jin, M.; Li, X.S. An improved ESPRIT-like algorithm for coherent signals DOA estimation. *IEEE Commun. Lett.* **2019**, *24*, 339–343. [\[CrossRef\]](#)
11. Wang, R.; Wang, Y.; Li, Y.; Cao, W.; Yan, Y. Geometric algebra-based ESPRIT algorithm for DOA estimation. *Sensors* **2021**, *21*, 5933. [\[CrossRef\]](#)
12. Wax, M.; Adler, A. Direction of arrival estimation in the presence of model errors by signal subspace matching. *Signal Process.* **2021**, *181*, 107900–107913. [\[CrossRef\]](#)
13. Gao, W.; Zhu, S.; Li, X.; Wang, H.; Wang, L. Frequency-difference MUSIC: A method for DOA estimation in inhomogeneous media. *Signal Image Video Process.* **2024**, *18*, 7029–7040. [\[CrossRef\]](#)
14. Fang, Y.; Zhu, S.; Gao, Y.; Zeng, C. DOA estimation for coherent signals with improved sparse representation in the presence of unknown spatially correlated Gaussian noise. *IEEE Trans. Veh. Technol.* **2020**, *69*, 10059–10069. [\[CrossRef\]](#)
15. Liu, B.; Matsushita, S.; Xu, L. DOA estimation with small snapshots using weighted mixed norm based on spatial filter. *IEEE Trans. Veh. Technol.* **2020**, *69*, 16183–16187. [\[CrossRef\]](#)
16. Liu, Y.; Dong, N.; Zhang, X.; Zhao, X.; Zhang, Y.; Qiu, T. DOA estimation for massive MIMO systems with unknown mutual coupling based on block sparse Bayesian learning. *Sensors* **2022**, *22*, 8634. [\[CrossRef\]](#)

17. Luessi, M.; Babacan, S.D.; Molina, R.; Katsaggelos, A.K. Bayesian simultaneous sparse approximation with smooth signals. *IEEE Trans. Signal Process.* **2013**, *61*, 5716–5729. [\[CrossRef\]](#)
18. Tipping, M.E.; Faul, A.C. Fast Marginal Likelihood Maximisation for Sparse Bayesian Models. In Proceedings of the Ninth International Workshop on Artificial Intelligence and Statistics, Key West, FL, USA, 3–6 January 2003; pp. 276–283.
19. Fang, J.; Shen, Y.; Li, H.; Wang, P. Pattern-coupled sparse Bayesian learning for recovery of block-sparse signals. *IEEE Trans. Signal Process.* **2015**, *63*, 360–372. [\[CrossRef\]](#)
20. Dai, J.; So, H.C. Real-valued sparse Bayesian learning for DOA estimation with arbitrary linear arrays. *IEEE Trans. Signal Process.* **2021**, *69*, 4977–4990. [\[CrossRef\]](#)
21. Liu, D.; Zhao, Y. Real-valued sparse Bayesian learning algorithm for off-grid DOA estimation in the beamspace. *Digit. Signal Process.* **2022**, *121*, 103322–103328. [\[CrossRef\]](#)
22. Zeng, H.; Yue, H.; Cao, J.; Zhang, X. Real-valued direct position determination of quasi-stationary signals for nested arrays: Khatri–Rao Subspace and Unitary Transformation. *Sensors* **2022**, *22*, 4209. [\[CrossRef\]](#)
23. Wang, G.; Kang, Y.; Wang, H. Low-Complexity DOA Estimation Algorithm based on Real-Valued Sparse Bayesian Learning. *Circuits Syst. Signal Process.* **2024**, *43*, 4319–4338. [\[CrossRef\]](#)
24. Wang, Q.; Yu, H.; Li, J.; Ji, F.; Chen, F. Sparse Bayesian learning using generalized double Pareto prior for DOA estimation. *IEEE Signal Process. Lett.* **2021**, *28*, 1744–1748. [\[CrossRef\]](#)
25. Guo, Q.; Xin, Z.; Zhou, T.; Xu, S. Off-grid space alternating sparse Bayesian learning. *IEEE Trans. Instrum. Meas.* **2023**, *72*, 1–10. [\[CrossRef\]](#)
26. Zhang, Y.; Yang, Y.; Yang, L. Off-grid DOA estimation through variational Bayesian inference in colored noise environment. *Digit. Signal Process.* **2021**, *111*, 102967–102981. [\[CrossRef\]](#)
27. Wang, P.; Yang, H.; Ye, Z. An off-grid wideband DOA estimation method with the variational Bayes expectation-maximization framework. *Signal Process.* **2022**, *193*, 108423–108430. [\[CrossRef\]](#)
28. Wang, H.; Wang, X.; Huang, M.; Wan, L. A novel variational SBL approach for off-grid DOA detection under nonuniform noise. *Digit. Signal Process.* **2022**, *128*, 103622–103630. [\[CrossRef\]](#)
29. Fu, H.; Dai, F.; Hong, L. Off-grid error calibration for DOA estimation based on sparse Bayesian learning. *IEEE Trans. Veh. Technol.* **2023**, *72*, 16293–16307. [\[CrossRef\]](#)
30. Liu, Y.; Zhang, Z.; Zhou, C.; Yan, C.; Shi, Z. Robust variational Bayesian inference for direction-of-arrival estimation with sparse array. *IEEE Trans. Veh. Technol.* **2022**, *71*, 8591–8602. [\[CrossRef\]](#)
31. Zheng, H.; Zhou, C.; Shi, Z.; Yan, C.; Shi, Z. Coarray tensor direction-of-arrival estimation. *IEEE Trans. Signal Process.* **2023**, *71*, 1128–1142. [\[CrossRef\]](#)
32. Yang, Z.; Xie, L.; Zhang, C. Off-grid direction of arrival estimation using sparse Bayesian inference. *IEEE Trans. Signal Process.* **2012**, *61*, 38–43. [\[CrossRef\]](#)
33. Dai, J.; Bao, X.; Xu, W.; Chang, C. Root sparse Bayesian learning for off-grid DOA estimation. *IEEE Signal Process. Lett.* **2016**, *24*, 46–50. [\[CrossRef\]](#)
34. Huang, H.; So, H.C.; Zoubir, A.M. Off-grid direction-of-arrival estimation using second-order Taylor approximation. *Signal Process.* **2022**, *196*, 108513–108519. [\[CrossRef\]](#)
35. Wagner, M.; Park, Y.; Gerstoft, P. Gridless DOA estimation and root-MUSIC for non-uniform linear arrays. *IEEE Trans. Signal Process.* **2021**, *69*, 2144–2157. [\[CrossRef\]](#)
36. Elbir, A.M. DeepMUSIC: Multiple signal classification via deep learning. *IEEE Sens. Lett.* **2020**, *4*, 1–4. [\[CrossRef\]](#)
37. Papageorgiou, G.K.; Sellathurai, M.; Eldar, Y.C. Deep networks for direction-of-arrival estimation in low SNR. *IEEE Trans. Signal Process.* **2021**, *69*, 3714–3729. [\[CrossRef\]](#)
38. Li, J.; Shao, X.; Li, J.; Ge, L. Direction of arrival estimation of array defects based on deep neural network. *Circuits Syst. Signal Process.* **2022**, *41*, 4906–4927. [\[CrossRef\]](#)
39. Mestre, X.; Lagunas, M.A. Finite sample size effect on minimum variance beamformers: Optimum diagonal loading factor for large arrays. *IEEE Trans. Signal Process.* **2005**, *54*, 69–82. [\[CrossRef\]](#)
40. Florio, A.; Coviello, G.; Talarico, C.; Avitabile, G. Adaptive DDS-PLL Beamsteering Architecture based on Real-Time Angle-of-Arrival Estimation. In Proceedings of the 2024 IEEE 67th International Midwest Symposium on Circuits and Systems (MWSCAS), Springfield, MA, USA, 11–14 August 2024; pp. 628–631.
41. Florio, A.; Avitabile, G.; Talarico, C.; Coviello, G. A Reconfigurable Full-Digital Architecture for Angle of Arrival Estimation. *IEEE Trans. Circuits Syst. I Regul. Pap.* **2023**, *71*, 1443–1455. [\[CrossRef\]](#)

42. Möderl, J.; Leitinger, E.; Fleury, B.H.; Pernkopf, F.; Witrals, K. Fast variational block-sparse Bayesian learning. *arXiv* **2023**, arXiv:2306.00442.
43. Stoica, P.; Nehorai, A. MUSIC, maximum likelihood, and Cramer-Rao bound. *IEEE Trans. Acoust. Speech Signal Process.* **2002**, *37*, 720–741. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.